DTIC FILE COPY

(2)

AD-A229 629

[redacted]

DTIC
ELECTE
DEC 12 1990
S E D

DEPARTMENT OF THE AIR FORCE

AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

2

AFIT/GLM/LSM/90S-37

FORECASTING AIR FORCE LOGISTICS COMMAND SECOND DESTINATION TRANSPORTATION: AN APPLICATION OF MULTIPLE REGRESSION ANALYSIS AND NEURAL NETWORKS

THESIS

Kevin R. Moore, Captain, USAF

AFIT/GLM/LSM/90S-37

DTIC
1990
D

Approved for public release; distribution unlimited

The opinions and conclusions in this paper are those of the author and are not intended to represent the official position of the DOD, USAF, or any other government agency.

Accession For
NTIS GRA&I ☒
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

DTIC COPY INSPECTED 5

AFIT/GLM/LSM/90S-37

FORECASTING AIR FORCE LOGISTICS COMMAND SECOND DESTINATION TRANSPORTATION: AN APPLICATION OF MULTIPLE REGRESSION ANALYSIS AND NEURAL NETWORKS

THESIS

Presented to the Faculty of the School of Systems
and Logistics of the Air Force Institute of Technology
Air University
In Partial Fulfillment of the
Requirements for the Degree of
Master of Science in Logistics Management

Kevin R. Moore, B.S.
Captain, USAF

September 1990

Approved for public release; distribution unlimited

Acknowledgements

This research would not have been possible without the help from the DSXR staff. William Lyons, Charles Ramsey, and Martin Meeks went above and beyond to help complete this research. Special thanks go out to Marlene Black and Nina Stutler for compiling the aircraft flying hour data. Also, I would like to thank Lt Col Robert E. Trempe for his words of wisdom and knowledge about Second Destination Transportation and Lt Col Phillip E. Miller for teaching me the power of statistics. Finally, I would like to thank my thesis advisor, Lt Col James R. Holt, for his innovative spirit and AI technical knowledge, without which this thesis would not be addressing the capability of neural networks in a forecasting application.

K. R. Moore

Table of Contents

	Page
Acknowledgements	ii
List of Figures	viii
List of Tables	xiii
Abstract	xvi
I. Introduction	1
Overview	1
Definitions	3
Current Methodology	5
Specific Problem	8
Research Objectives	9
Scope of the Research	10
Plan of Analysis	10
II. Literature Review	11
Previous SDT Reports and Research	11
LMI Task 75-4 (1976)	11
Foster Report (1977)	12
Grayson Report (1977)	13
Lamb and Sarnacki Research (1978)	14
Abell Report (1982)	16
Strom Research (1989)	18
Navy and Army SDT Program	22
Navy SDT	22
Army SDT	24
Summary of the SDT Reports and the Army/Navy SDT	25
Neurocomputing	27
Neural Networks	28
Single Layer Perceptron	33
Multi-Layer Perceptron	37
Backpropagation	39
Types of Networks	41
Neural Network Applications	42
Neural Network Forecasting Applications	43
Neural Network vs. Regression (Bond Rating Problem)	43
Application of Neural Networks in Time Series Forecasting	45
Chapter Summary	47

Page

III. Methodology 48

Collection of the Data 48
Data Analysis Methodology 49
Plots 49
Trend and Seasonal Analysis 49
Business Cycle Analysis 50
Pattern Identification 50
DSXR Simple Regression Model Validation and Forecasting Evaluation 51
Multiple Regression Model Development, Validation and Forecasting Evaluation 54
Neural Network Model Development and Forecasting Evaluation 57
Methodology Summary 60

IV. MSC SDT Forecasting Results and Analysis 61

DSXR Simple Regression Model Validation and Forecasting Evaluation 61
DSXR PACAF MSC Model 61
DSXR USAFE MSC Model 66
Data Analysis 71
Plots 71
Trend and Seasonal Analysis 77
Pattern Identification 78
Business Cycle Analysis 79
Multiple Regression Models 81
Data 84
PACAF MSC Multiple Regression Model Development 84
Model Validation and Forecasting Evaluation 85
USAFE MSC Multiple Regression Model Development 92
Model Validation and Forecasting Evaluation 93
Neural Networks 99
Data 99
PACAF MSC Networks 100
Network Development 102
Forecasting Evaluation 104
USAFE MSC Networks 105
Network Development 107
Forecasting Evaluation 108
Chapter Summary 110

V. MAC SDT Forecasting Results and Analysis 111

Data Analysis 112
Plots 112

Page

Trend and Seasonal Analysis 115
Pattern Identification 115
DSXR Simple Regression Model Validation and Forecasting Evaluation 116
DSXR PACAF MAC Model 116
DSXR USAFE MAC Model 117
Multiple Regression Models 120
Data 120
PACAF MAC Multiple Regression Model Development 121
Model Validation and Forecasting Evaluation 121
USAFE MAC Multiple Regression Model Development 126
Model Validation and Forecasting Evaluation 127
Neural Networks 132
Data 132
PACAF MAC Networks 132
Network Development 133
Forecasting Evaluation 135
USAFE MAC Networks 136
Network Development 137
Forecasting Evaluation 138
Time Series Neural Networks 140
Data 141
PACAF and USAFE MAC Time Series Network Development. 141
PACAF MAC Time Series Network Forecasting Evaluation 145
USAFE MAC Time Series Network Forecasting Evaluation 146
Chapter Summary 148

VI. Research Conclusions and Findings 149

Research Conclusions. 149
Research Objective 1 149
Research Objective 2 151
Findings 152
Multivariable Approach 152
Aircraft Flying Hour Variables 153
Military Population Variables 156
Financial Implications 158
Programmed versus Actual Flying Hours. 160
Neural Network Findings 161
Further Research. 164
Neural Networks. 164
Combined Forecasting Techniques 164
Inventory Models 164

Page

Appendix A: DSXR PACAF MSC Model SAS Regression Output 166

Appendix B: DSXR USAFE MSC Model SAS Regression Output 174

Appendix C: Trend and Seasonal Analysis (PACAF, USAFE MSC and Flying Hours). 182

Appendix D: Time Series Analysis (PACAF and USAFE MSC) 190

Appendix E: Business Cycle Analysis (PACAF and USAFE MSC). . . 196

Appendix F: PACAF and USAFE Military Populations 198

Appendix G: PACAF Aircraft Flying Hours and Inventory by MD. . 199

Appendix H: PACAF MSC Multiple Regression Model SAS Output . . 202

Appendix I: PACAF MSC Independent Variable Correlation Matrix 214

Appendix J: USAFE Aircraft Flying Hours and Inventory by MD. . 217

Appendix K: USAFE MSC Multiple Regression Model SAS Output . . 220

Appendix L: USAFE MSC Independent Variable Correlation Matrix 232

Appendix M: PACAF Transformed Network Data 235

Appendix N: USAFE Transformed Network Data 237

Appendix O: PACAF and USAFE MSC Multivariable Network Output 239

Appendix P: Trend and Seasonal Analysis (PACAF and USAFE MAC Data). 243

Appendix Q: Time Series Analysis (PACAF and USAFE MAC Data). 247

Appendix R: DSXR USAFE MAC Model SAS Regression Output 253

Appendix S: PACAF MAC Multiple Regression Model SAS Output . . 259

Appendix T: PACAF MAC Independent Variable Correlation Matrix 267

Appendix U: USAFE MAC Multiple Regression Model SAS Output . . 270

Appendix V: USAFE MAC Independent Variable Correlation Matrix 278

Page

Appendix W: PACAF and USAFE MAC Multivariable Network Output 281

Appendix X: PACAF and USAFE MAC Time Series Network Data . . . 285

Appendix Y: PACAF and USAFE MAC Time Series Network Output . . 286

Appendix Z: PACAF and USAFE MSC, MAC, and Total Flying Hour Data 288

Bibliography 290

Vita 293

List of Figures

Figure		Page
1.	A Neural Network Processing Element	29
2.	A Neural Network Processing Element Activation (Example 1)	30
3.	A Neural Network Processing Element Activation (Example 2)	31
4.	Sigmoid Function	32
5.	Hard Limiter and Threshold Logic Function	32
6.	Single Layer Perceptron and Decision Boundary	34
7.	The Exclusive-OR Decision Region	36
8.	A Multi-Layer Perceptron	38
9.	Multi-Layer Perceptron Exclusive-OR Decision Region	39
10.	PACAF MSC Tonnage versus PACAF Flying Hours	71
11.	Quarterly PACAF MSC Tonnage	72
12.	USAFE MSC Tonnage versus USAFE Flying Hours	73
13.	Quarterly USAFE MSC Tonnage	73
14.	Quarterly PACAF MSC Tonnage and Flying Hours	74
15.	Quarterly USAFE MSC Tonnage and Flying Hours	75
16.	Aircraft Procurement Dollars	75
17.	PACAF Military Population	76
18.	USAFE Military Population	77
19.	PACAF and USAFE 1/4 Pressures	80
20.	PACAF and USAFE 4/4 Pressures	80
21.	PACAF and USAFE Sealift/Flying Hour Ratios	81
22.	PACAF Flying Hours by Aircraft Type	83
23.	USAFE Flying Hours by Aircraft Type	83
24.	PACAF MSC Independent Variable Selection	85

Page

25. DSXR and Multiple Regression PACAF MSC Forecasts 90

26. USAFE MSC Independent Variable Selection 93

27. DSXR and Multiple Regression USAFE MSC Forecasts 98

28. MSC Full Multivariable Network Model. 101

29. MSC Reduced Multivariable Network Model 102

30. PACAF MSC Full Multivariable Network Model Training Iterations 103

31. PACAF MSC Reduced Multivariable Network Model Training Iterations 104

32. PACAF MSC Full and Reduced Multivariable Network Forecasts 106

33. USAFE MSC Full Multivariable Network Model Training Iterations 107

34. USAFE MSC Reduced Multivariable Network Model Training Iterations 108

35. USAFE MSC Full and Reduced Multivariable Network Forecasts 110

36. Quarterly PACAF MAC Tonnage 113

37. Quarterly USAFE MAC Tonnage 113

38. PACAF MAC Tonnage versus PACAF Flying Hours 114

39. USAFE MAC Tonnage versus USAFE Flying Hours 114

40. DSXR and Multiple Regression PACAF MAC Forecasts 125

41. DSXR and Multiple Regression USAFE MAC Forecasts 131

42. PACAF MAC Full Multivariable Network Training Iterations 134

43. PACAF MAC Reduced Multivariable Network Training Iterations 134

44. PACAF MAC Full and Reduced Multivariable Network Forecasts 136

45. USAFE MAC Full Multivariable Network Training Iterations 138

Page

46. USAFE MAC Reduced Multivariable Network Training Iterations 139

47. USAFE MAC Full and Reduced Multivariable Network Forecasts 140

48. USAFE and PACAF MAC Time Series Network Configuration . . 142

49. PACAF MAC Time Series Network Pattern Recognition 144

50. USAFE MAC Time Series Network Pattern Recognition 144

51. PACAF MAC Time Series Network Forecasts 146

52. USAFE MAC Time Series Network Forecasts 147

53. PACAF and USAFE MSC Model Rank Order 150

54. PACAF and USAFE MAC Model Rank Order 151

55. PACAF Aircraft Flying Hours by Percent 155

56. USAFE Aircraft Flying Hours by Percent 155

57. DSXR PACAF MSC Model Residuals versus Predicted Values. . 169

58. DSXR PACAF MSC Model Residuals versus Flying Hours. . . . 170

59. DSXR PACAF MSC Model Studentized Residuals versus Flying Hours 171

60. DSXR PACAF MSC Model Residuals versus Time (N) 172

61. DSXR USAFE MSC Model Residuals versus Predicted Values. . 177

62. DSXR USAFE MSC Model Residuals versus Flying Hours. . . . 178

63. DSXR USAFE MSC Model Studentized Residuals versus Flying Hours 179

64. DSXR USAFE MSC Model Residuals versus Time (N) 180

65. PACAF Multiple Regression MSC Model Residuals versus Predicted Values 204

66. PACAF Multiple Regression MSC Model Residuals versus A-10 Flying Hours. 205

67. PACAF Multiple Regression MSC Model Residuals versus F-16 Flying Hours. 206

Page

68. PACAF Multiple Regression MSC Model Residuals versus Officer Population (OFF) 207

69. PACAF Multiple Regression MSC Model Residuals versus Time (N) 208

70. PACAF Multiple Regression MSC Model (Second Order) Residuals versus Officer Population (OFF) 212

71. USAFE Multiple Regression MSC Model Residuals versus Predicted Values 222

72. USAFE Multiple Regression MSC Model Residuals versus C-130 Flying Hours 223

73. USAFE Multiple Regression MSC Model Residuals versus F-4 Flying Hours 224

74. USAFE Multiple Regression MSC Model Residuals versus Officer Population (OFF) 225

75. USAFE Multiple Regression MSC Model Residuals versus Time (N) 226

76. USAFE Multiple Regression MSC Model (Second Order) Residuals versus Officer Population (OFF) 230

77. DSXR USAFE MAC Model Residuals versus Predicted Values. . 255

78. DSXR USAFE MAC Model Residuals versus Flying Hours. . . . 256

79. DSXR USAFE MAC Model Residuals versus Time (N) 257

80. PACAF Multiple Regression MAC Model Residuals versus Predicted Values (LMAC). 261

81. PACAF Multiple Regression MAC Model Residuals versus B-52 Flying Hours. 262

82. PACAF Multiple Regression MAC Model Residuals versus F-15 Flying Hours. 263

83. PACAF Multiple Regression MAC Model Residuals versus Airman Population (AMN). 264

84. PACAF Multiple Regression MAC Model Residuals versus Time (N) 265

85. USAFE Multiple Regression MAC Model Residuals versus Predicted Values (LMAC). 272

Page

86. USAFE Multiple Regression MAC Model Residuals versus A-10 Flying Hours. 273

87. USAFE Multiple Regression MAC Model Residuals versus Officer Population (OFF) 274

88. USAFE Multiple Regression MAC Model Residuals versus Airman Population (AMN). 275

89. USAFE Multiple Regression MAC Model Residuals versus Time (N) 276

List of Tables

Table		Page
1.	AFLC SDT Major Categories	4
2.	Major Program and Requirement Source	4
3.	FY 1989 SDT Costs	6
4.	Exclusive OR Problem	36
5.	Types of Neural Networks	41
6.	DSXR PACAF MSC Model Analysis of Variance	62
7.	DSXR PACAF MSC Model Forecasting Accuracy	65
8.	DSXR USAFE MSC Model Analysis of Variance	67
9.	DSXR USAFE MSC Model Forecasting Accuracy	70
10.	MSC Trend and Seasonal Analysis	78
11.	PACAF MSC Multiple Regression Model Analysis of Variance	86
12.	PACAF MSC Multiple Regression Model Forecasting Accuracy	89
13.	PACAF MSC Multiple Regression Model Independent Variable Correlations	91
14.	USAFE MSC Multiple Regression Model Analysis of Variance	94
15.	USAFE MSC Multiple Regression Model Forecasting Accuracy	97
16.	USAFE MSC Multiple Regression Model Independent Variable Correlations	99
17.	PACAF MSC Network Independent Variables	101
18.	PACAF MSC Multivariable Network Forecasting Accuracy	105
19.	USAFE MSC Network Independent Variables	106
20.	USAFE MSC Multivariable Network Forecasting Accuracy	109
21.	DSXR PACAF MAC Model Forecasting Accuracy	116

Page

22. DSXR USAFE MAC Model Analysis of Variance 118

23. DSXR USAFE MAC Model Forecasting Accuracy 119

24. PACAF MAC Multiple Regression Model Analysis of Variance. 122

25. PACAF MAC Multiple Regression Model Forecasting Accuracy. 125

26. PACAF MAC Multiple Regression Model Independent Variable Correlations 126

27. USAFE MAC Multiple Regression Model Analysis of Variance. 128

28. USAFE MAC Multiple Regression Model Forecasting Accuracy. 130

29. USAFE MAC Multiple Regression Model Independent Variable Correlations 132

30. PACAF MAC Network Independent Variables 133

31. PACAF MAC Multivariable Network Forecasting Accuracy. 135

32. USAFE MAC Network Independent Variables 137

33. USAFE MAC Multivariable Network Forecasting Accuracy. 139

34. PACAF and USAFE MAC Time Series Network Independent Variables 142

35. Time Series Network Training and Forecasting Iteration Methodology 143

36. PACAF MAC Time Series Network Forecasting Accuracy. . . . 145

37. USAFE MAC Time Series Network Forecasting Accuracy. . . . 147

38. Multiple Regression Aircraft Variables. 154

39. Multiple Regression Military Population Variables 156

40. Overseas SDT Costs (Dollars). 159

41. Financial Implications (Dollars). 159

Page

42. Total Difference in Underestimations and Overestimations (Dollars) 160

43. Programmed and Actual Flying Hours. 161

44. Programmed and Actual Flying Hours Financial Implications (Dollars). 161

Abstract

The Air Force Logistics Command ~~(AFLC/DSXR)~~ currently uses a simple linear regression model to forecast overseas Second Destination Transportation (SDT) general cargo tonnage requirements for specific geographical areas. The independent variable for the model is the total flying hours for each geographical area while the dependent variable is the general cargo SDT tonnage requirement. This ~~research~~ explored the use of a multivariable approach for developing multiple regression and neural network models which was based on the breakout of the total flying hour variable into separate aircraft flying hours and the addition of military population variables.

The purpose of this research was to develop multiple regression and neural network models for predicting Pacific ~~(PACAF)~~ and European ~~(USAFE)~~ Military Airlift Command ~~(MAC)~~ and Military Sealift Command ~~(MSC)~~ general cargo tonnage requirements that were more accurate forecasting models than the simple regression forecasting models presently used by AFLC/DSXR. Once the models were developed, the multiple regression and neural network models were compared to determine which type of model was statistically more accurate.

Neural networks are an adaptive information processing system loosely based on the information processing capability of the human brain that mathematically develops associations between particular independent and dependent variables. Recent research indicates neural network models are an alternative to conventional mathematical techniques for solving problems that do not have a well defined model or theory. Unlike

regression models, neural networks determine the equation and parameters for independent variables, thereby eliminating the difficulty of prespecifying the model.

Overall, the use multivariable model development approach significantly increases SDT forecasting accuracy. The neural network models were the most accurate forecasting models. In three out of the four data sets used, the multiple regression models produced more accurate forecasts than the AFLC/DSXR simple regression model. The application of either model would significantly reduce the financial implications of overestimating and underestimating SDT tonnage requirements.

FORECASTING AIR FORCE LOGISTICS COMMAND SECOND DESTINATION TRANSPORTATION: AN APPLICATION OF MULTIPLE REGRESSION ANALYSIS AND NEURAL NETWORKS

I. Introduction

Overview

The Air Force Logistics Command (AFLC) is responsible for providing logistical support for the Air Force. A key element to AFLC's effectiveness in providing logistical support is the transportation system which adds the time and place utility for all logistical support. The increasingly severe Air Force funding limitations have placed a greater emphasis on the transportation system to provide timely distribution of logistical support to the end users at a minimum cost.

Recent budget reductions in Second Destination Transportation (SDT) have reduced the funding level to 80 percent for fiscal year 1989 and 1990. The impact has resulted in the following:

1. Decreased logistical support to the end user.
2. Increased cannibalization of parts.
3. Increased down time for parts.
4. Increased transit time for parts.
5. Decreased stockage effectiveness.
6. Incomplete War Readiness Materiel (WRM) spares kits.
7. Accruing secondary costs (temporary storage, handling charges, expiring warranties, corrosion).
8. Decreasing theater sustainability (spares, munitions).

9. U.S. manufactured vehicles destined for overseas bases being held at ports.

10. The retention of munitions in CONUS [Continental United States] while shortages exist at overseas locations.

11. Malpositioned munitions in operational theaters. (19)

Another result of the recent budget reductions for SDT and the topic of this research was the requirement for a statistically valid and accurate SDT forecasting method. Underestimations of the SDT requirement compound the problem of providing logistical support during budget reductions. Overestimations divert scarce funds from other Air Force programs.

In the Department of Defense (DOD), SDT budget requirements are a separate line item under Major Force Program (MFP) VII, Central Supply and Maintenance (20:3-4). Each branch of the military services develops SDT budget requirements which are included in the operations and maintenance budget for each respective service. Within the Air Force, the Air Force Logistics Command controls approximately 75-80 percent of the SDT budget (20:32). The remainder of the budget is divided between the other major commands and the Directorate of Administration, Office of the Chief of Staff. The Plans and Programs Division, Directorate of Transportation, Headquarters United States Air Force (HQ USAF/LETX), consolidates the forecasted SDT budget requirements from each of the major commands and the Directorate of Administration into the Air Force SDT budget requirement.

The Budget and Requirements Division, Directorate of Programs and Resources, Chief of Staff Distribution, Headquarters Air Force Logistics Command (HQ AFLC/DSXR) is responsible for forecasting AFLC Second Destination Transportation (SDT) funding requirements.

Definitions. In order to understand SDT, First Destination Transportation (FDT) must be defined. According to Department of Defense Instruction (DODI) 5000.8, First Destination Transportation is defined as:

> The movement of property from f.o.b. [free on board] point of origin to the point at which the materiel, in the form required for use, is first received for use or storage for subsequent distribution in the military supply system. The costs of such movement. (6:124)

In other words, First Destination Transportation is the inland movement of newly acquired material from the contractor or vendor to the first point of use or storage or to a CONUS port of embarkation for onward moves to an overseas location (19).

Second Destination Transportation is defined as:

> The subsequent movement of property for intradepartment or interdepartment distribution from the point of storage at which originally received from f.o.b. point of origin. The costs of such movement. (6:124)

Second Destination Transportation can be described as all subsequent movement of materials, including port handling. It covers all overocean transportation of logistics cargo, all shipments from the ALC's (Air Logistics Centers or depots), shipments between bases, and shipments from bases to repair depots (19).

AFLC SDT budgeting and funding responsibilities are divided into six major categories identified by Air Force Element of Expense/Investment Account Codes (EEIC) (Table 1). The six categories represent five modes of transport and the port handling requirement.

In addition to the six categories of shipments, there are fourteen major programs. Table 2 displays these programs along with the requirement sources. The requirement source is used to forecast future

tonnage requirements for each major program. This research was focused on the general cargo program for overseas SDT tonnage.

Table 1

AFLC SDT Major Categories (20:34)

Category	EEIC
Military Airlift Command (MAC)	454
Military Sealift Command (MSC)	461
Commercial Air	462
Commercial Surface	463
Logair	464
Port Handling	465

Table 2

Major Program and Requirement Source (19)

Major Program	Requirement Source
General Cargo	Flying hour program
Logair	Majcom (Major Command) Program Unit Priority Document
Fuel	Miles/Fuel burn rate
Subsistence	Overseas manpower
Air Munitions	Majcom WRM (War Readiness Material)/PTO (Peacetime Operations)
CENTAF SWA-WRM (Central Air Force Southwest Asia-War Readiness Material)	CENTAF
New Vehicles	Inventory Manager
Missiles	Program Manager
Special Weapons	Program Manager
PADS (Program Action Directives)	Majcom/Air Logistics Centers (ALC)
DOD Schools	History
PCS (Permanent Change of Station) Civilians	History
Commercial Paper	Flying Hour Program
CCP (Container Collection Point)	Flying Hour Program

Overseas general cargo is forecasted within the general cargo program and the requirement source is the overseas flying hour program. The overseas general cargo tonnage requirement is forecasted for two modes of transport, airlift (MAC) and sealift (MSC) for each geographical area by using a simple linear regression model. The flying hour program is used as an independent variable for predicting future tonnage requirements (dependent variable) for each mode of transport. The general cargo tonnage requirements are further divided into five geographical areas: Pacific (PACAF), Europe (USAFE), Northern (Greenland and Iceland), Southern (USAFSO), and Alaska. The general cargo tonnage requirements are divided into geographical areas for the following reasons:

1. The cost of shipment depends on distance moved and the weight of the shipment.

2. Each geographical area has different general cargo tonnage requirements.

3. The flying hours are programmed by geographical area. (19)

Table 3 depicts the costs of AFLC SDT for fiscal year 1989 by mode and theater. The airlift and surface (sealift) shipment of cargo to overseas bases represents 72 percent of the total SDT costs for FY 89. The remaining 28 percent of the total costs is airlift and surface SDT shipments within CONUS. By theater, PACAF and USAFE SDT shipments represent 62 percent of the total SDT costs while CONUS and shipments to other locations represent the remaining 38 percent.

Current Methodology. AFLC DSXR uses quarterly flying hours as an independent variable in a simple linear regression model to predict quarterly SDT general cargo tonnage requirements (dependent variable) for specific geographical areas as shown in the following equation;

$$Y = \beta_0 + \beta_1 X$$

where: Y = historical (predicted) quarterly SDT tonnage
X = historical (programmed) quarterly flying hours
β_0 = the y axis intercept
β_1 = flying hour parameter.

Table 3

FY 1989 SDT Costs (19)

Airlift	Percent of Cost	Total Cost ($ Million)
Overseas (MAC)	40	154
CONUS	19	75
Surface		
Overseas (MSC)	32	125
CONUS	9	33
	Total:	387
Theater		
PACAF	25	97
USAFE	37	143
CONUS	29	112
Other	9	35
	Total:	387

Historical flying hours and MAC and MSC tonnage requirements for each geographical area are used to develop the simple linear regression equations. Flying hours that are programmed six quarters into the future are used to predict future tonnage requirements using the simple linear regression equations that were developed from the historical data. The flying hour independent variable is the total programmed flying hours for

all aircraft assigned to each geographical area. This includes all transient aircraft in the geographical area for over 60 days (19).

There are two reasons for using the flying hour program as the independent variable in the simple regression model for predicting overseas general cargo. First, the use of the flying hour program as the independent variable is based on past experience (1:1). Variations in airlift and sealift tonnage were directly related to variations in flying hours (1:1). Second, the general cargo category is comprised of aircraft spare parts and general base supply items (19). As long as the relationship held constant, the use of the flying hour program as the only independent variable was justified. Recent research conducted by Strom indicated the relationship was not always constant and the forecasting technique used by DSXR was not always valid (25:62).

AFLC/DSXR utilizes an iterative approach to achieve the best simple regression model by beginning with the last 40 quarters of data and eliminating the oldest quarter of data until the last 8 recent quarters remain (19). After each iteration, the coefficient of correlation (r) and the standard deviation of each variable is calculated for each model. The model with the highest coefficient of correlation is used to predict the future six quarters of airlift and sealift tonnage requirements for each geographical area based on the quarterly programmed flying hours for each geographical area. The coefficient of correlation is defined in the following equation;

$$r = \pm [1 - (SSE / SSY)]^{1/2}$$

where: SSE = the unexplained sample variation
SSY = the total sample variation.

In addition to the iterations, two types of data are used for the simple linear regression model, smoothed and non-smoothed data. A smoothing technique is used to reduce the quarterly variation in the MAC and MSC tonnage requirements (19). This technique is based on a three quarter moving average and is started by setting the first moving average value equal to the earliest quarter value of the data set. The next moving average value is equal to two times the earliest quarter added to the second earliest quarter and the sum is then divided by three. The next to the last moving average is equal to two times the last quarter added to the next to the last quarter and the sum is then divided by three. All the quarters in between are computed as a regular moving average (19).

Once all the simple regression equations for MAC and MSC tonnage have been computed for each geographical area through the iterative technique, the simple regression equations (smoothed or nonsmoothed developed equations) that produce the highest positive coefficient of correlation for the regression model is used to forecast future SDT tonnage requirements.

Specific Problem

Research conducted by Captain Stephen L. Strom in 1989 analyzed the simple linear regression model used by DSXR to forecast SDT general cargo requirements for PACAF and USAFE. Strom's research statistically invalidated the iterative linear regression method used to predict MSC general cargo SDT tonnage for PACAF and USAFE by showing how the flying hour parameter (β_1 in the simple regression equation) statistically changed after iterations were conducted. This instability in the flying

hour parameter implied the model was invalid for predicting future tonnage values (25:80-81). Strom's research did not statistically invalidate the models used to predict MAC SDT tonnage for PACAF and USAFE. Strom developed a Box-Jenkins model that was statistically more accurate than DSXR's PACAF and USAFE MAC forecasting models, but failed to develop a Box-Jenkins model that was statistically more accurate than the DSXR PACAF and USAFE MSC forecasting models. Strom concluded that further research is needed to develop a valid forecasting model that is significantly more accurate than the one presently used by DSXR and suggested the development and testing of a multiple regression model (25:88).

Research Objectives

The usefulness of any forecasting model depends on how accurate it can forecast. In order to develop forecasting models with improved forecasting accuracy compared to the DSXR models, the data base of the independent variable was increased by separating the total flying hour variable into flying hours by aircraft type. Military population variables were also added to the data base and used in developing the models. The objectives of this research were to:

1. Develop multiple regression and neural network models using flying hours by aircraft type and military population variables that were statistically more accurate than the DSXR simple regression models.

2. Determine whether the neural network or multiple regression models were more accurate forecasting models.

A neural network (neurocomputing) is an adaptive information processing system that is 'trained' to develop associations between

particular inputs and a desired output. Neurocomputing is a rapidly emerging technology that is being applied in areas such as complex pattern recognition problems, identifying handwritten characters, understanding speech, and economic forecasting. Neurocomputing has been successfully used to solve problems that can not be solved using conventional algorithmic methods (13:36-37).

Scope of the Research

Since there was a need for a statistically correct and accurate forecasting model for the overseas general cargo program, this research was limited to developing multiple regression and neural network models for PACAF and USAFE MAC and MSC SDT requirements. In fiscal year 1989, the PACAF and USAFE MAC SDT represented approximately 70% of the total overseas MAC general cargo and the PACAF and USAFE MSC represented approximately 92% of the total overseas MSC general cargo.

Plan of Analysis

Chapter II is a review of previous SDT research studies and findings and also presents an introduction to neural networks and neural network forecasting applications. Chapter III outlines the methodology that was used to evaluate the DSXR simple regression models and to develop and validate the multiple regression and neural network models. The results and analysis of the PACAF and USAFE MSC data sets are presented in Chapter IV. Chapter V presents the results and analysis of the PACAF and USAFE MAC data sets. Finally, Chapter VI presents the research findings, implications, and future research recommendations.

II. Literature Review

This chapter is divided into four parts. The first part is a presentation of previous SDT reports and research that have had an impact on forecasting future SDT requirements. Part two examines the Navy and Army methodology for forecasting SDT requirements. The third part presents a background on neurocomputing and the backpropagating neural networks. The fourth part presents previous research on the use of neural networks in forecasting applications.

Previous SDT Reports and Research

The purpose of this section is to present research findings concerning SDT forecasting by examining six previous reports:

1. LMI Task 75-4 (1976)
2. Foster Report (1977)
3. Grayson Report (1977)
4. Lamb and Sarnacki Research (1978)
5. Abell Report (1982)
6. Strom Research (1989).

LMI Task 75-4 (1976). This report was prepared by E. A. Narragon and J. M. Neil at the request of the Assistant Secretary of Defense (Installations and Logistics) (ASD(I&L)). The purpose of the report was to identify opportunities for more effective and efficient utilization of transportation resources. The report evaluated the control and cost effectiveness of SDT funds by each military service and identified areas requiring increased participation by ASD(I&L) (20:1).

The report identified many problems with the management of SDT by the military services, but one finding that was related to forecasting

indicated there were three principle causes for changes in the SDT funding requirements:

1. Rate Changes: Changes in rates occur because of numerous economic pressures upon commercial carriers and Single Manager Operating Agencies:

2. Workload Changes: Changes in workload occur because distribution patterns are modified through force level changes, repositioning of stocks, and the like; and

3. Policy Decisions: Service and OSD policy decisions can have a direct effect upon the total Service SDT program. These decisions may result in changes in transportation modes or workload. (20:61)

These three causes continually contribute to the difficulty of accurately forecasting future SDT requirements. For example, a recent (1988) policy decision made transportation priority 2 (TP-2) cargo ineligible for airlift and has resulted in a significant decrease in overseas MAC tonnage requirements. Now, all TP-2 cargo can no longer be airlifted and must be transported by an alternate mode. In the case of overseas TP-2 cargo, the alternate mode is sealift (19).

Foster Report (1977). This unpublished report titled, A Working Paper on Second Destination Transportation (SDT) Forecasting, was prepared by Newton W. Foster, Directorate of Management Sciences, Deputy Chief of Staff Plans and Programs, Headquarters Air Force Logistics Command (HQ AFLC, XRS), Wright-Patterson AFB, Ohio. This report is no longer available, but was presented in Strom's research. The study was accomplished in response to a review conducted by the Office of the Secretary of Defense (OSD) that raised concern about the validity of forecasting SDT using the flying hour program (25:29). The study had two objectives:

1. To support the use of flying hours as a predictor of SDT.

2. To develop a better method of predicting SDT if the flying hour related computation could not be supported. (25:30)

Sixteen quarters of data (FY 73/1 through FY 76/4) were collected under 21 different categories (i.e. manpower, requisitions, overseas flying hours, worldwide flying hours) for six major geographical regions: PACAF, USAFE, AAC, USAFSO, Northeastern, and Worldwide. The transportation tonnage for MAC, MSC, and GBL (Worldwide) were also collected. Regression analysis was used to find the relationships between the 21 different categories of data and the transportation tonnage through simple and multiple linear regression equations (25:30-31).

The conclusions of the study were:

1. The forecast method [DSXR's simple linear regression models], although not a totally valid predictor of SDT tonnage, was the most logical predictor based on the data provided and examined.

2. A better forecasting method for predicting SDT tonnage for a particular geographical region by a specific transportation mode was not evident based on the data provided and examined. (25:31)

Grayson Report (1977). In this report, Major John Grayson investigated the procedures used by AFLC to determine SDT budget requirements. The author believed there was an over reliance on the flying hour program to forecast future SDT requirements. The author presented two reasons for this opinion:

1. Programmed flying hours were consistently overestimated by an average of 25 percent compared to the actual flying hours. This disparity between programmed and actual flying hours questioned the effectiveness of the flying hour program as an indicator of SDT requirements.

2. The generalized nature of the flying hour data did not account for different types of aircraft. Different aircraft require different logistical support, and any aircraft transitions

(F-4 and F-16 transition) would have a significant impact on SDT requirements. (11:16-19)

Grayson recommended two specific actions to eliminate the perceived over reliance on the flying hour program. First, AFLC and Air Staff should continue to identify programs that are not dependent on the flying hour program and separate them from the general cargo category (examples of programs that have been separated from the general cargo category are munitions and new vehicles). Second, AFLC should identify additional independent variables and develop a multivariate formula for determining SDT requirements (11:24-27).

Lamb and Sarnacki Research (1978). In this research, Captain Christopher J. Lamb and Captain Joseph B. Sarnacki developed a computerized discontinuous linear regression model to forecast SDT requirements utilizing flying hours and manpower as the independent variables. This research showed that the discontinuous linear regression model was statistically more accurate than the delta factor model used by AFLC during this time period. The research also showed that the flying hour and manpower variables were reliable predictors of SDT tonnage requirements (16:37-40).

The delta factor model consisted of manually computing a simple ton/flying hour ratio for each geographical area based on historical data. Programmed flying hours were used to determine future SDT requirements and were directly related to tonnage requirements. In other words, an increase in flying hours meant an increase in SDT tonnage; however, budget overestimations resulted because this was not always the case. The research indicated this method was not an "acceptable and understandable decision-making tool for budget estimations" (16:7) by

higher echelons (Office of the Secretary of Defense) and that a requirement existed for a validated method (16:1-7).

Lamb and Sarnacki used the discontinuous linear regression model to account for shifts or changes in slopes that were evident in scattergrams of actual tonnage versus total programmed flying hours and actual tonnage versus total manpower authorizations (16:14-18). This research was limited to forecasting MAC SDT requirements. The general form of the equation used is displayed in the following equation;

$$
\begin{aligned}
Z = {} & \beta_0 + \beta_1 X + \{\beta_2(X-X_1)X_{p1} + \beta_3 X_{D1}\}^* \\
& + \{\beta_4(X-X_2)X_{p2} + \beta_5 X_{D2}\}^* + \{\beta_6(X-X_3)X_{p3} + \beta_7 X_{D3}\}^* \\
& + \beta_8 Y + \{\beta_9(Y-Y_1)Y_{p1} + \beta_{10} Y_{D1}\}^* \\
& + \{\beta_{11}(Y-Y_2)Y_{p2} + \beta_{12} Y_{D2}\}^* \\
& + \{\beta_{13}(Y-Y_3)Y_{p3} + \beta_{14} Y_{D3}\}^* + \epsilon
\end{aligned}
$$

* Discontinuous adjustments

where: Z = Tonnage transported by MAC

X = Flying hours (either programmed or actual)

Y = Manpower (either programmed or actual)

X_1, X_2, X_3, Y_1, Y_2, Y_3 = Discontinuous data points

X_{p1}, X_{p2}, X_{p3}, Y_{p1}, Y_{p2}, Y_{p3},

X_{D1}, X_{D2}, X_{D3}, Y_{D1}, Y_{D2}, Y_{D3} = Dummy variables defined as:

$X_{p1} = X_{D1} = 1$, if $X > X_1$; otherwise 0

$X_{p2} = X_{D2} = 1$, if $X > X_2$; otherwise 0

$X_{p3} = X_{D3} = 1$, if $X > X_3$; otherwise 0

$Y_{p1} = Y_{D1} = 1$, if $Y > Y_1$; otherwise 0

$Y_{p2} = Y_{D2} = 1$, if $Y > Y_2$; otherwise 0

$Y_{p3} = Y_{D3} = 1$, if $Y > Y_3$; otherwise 0

$\beta_0, \ldots, \beta_{14}$ = Coefficients of regression

ϵ = Random error component. (16:18-19)

This research had two major developments. First, a statistically valid linear regression model was used to forecast future SDT requirements that was more accurate than the delta model used by AFLC. Second, the model showed that flying hours and manpower were valid predictors of SDT tonnage and could be used in the same model.

Abell Report (1982). This report, prepared by Joseph A. Abell, evaluated the use of the linear regression model used by AFLC to forecast future tonnage requirements. The model evaluated in this report is the same one presently used by AFLC/DSXR and is presented in the following equation;

$$Y = \beta_0 + \beta_1 X$$

where: Y = historical (predicted) quarterly SDT tonnage
X = historical (programmed) quarterly flying hours
β_0 = the y axis intercept
β_1 = flying hour parameter

The objective of this research was to evaluate the linear regression model and the data smoothing technique and determine if the model met four requirements that Abell believed were necessary for any forecasting model. These requirements were:

1. be verifiable,
2. be able to incorporate indicators of future trends in operations,
3. be relatively straight-forward in its application and interpretation,
4. be able to produce the most accurate results possible with the information available. (1:2)

In order to properly evaluate the model, Abell determined it was necessary to reproduce previous results achieved by the model. Abell

used the Statistical Package for the Social Sciences (SPSS) within the CREATE system to produce scattergrams of the data and generate the linear regression equations and the statistics associated with the regression variables. Abell was not able to duplicate any of the previous results achieved by the model because the computer program used by AFLC was flawed and produced inaccurate results. Abell pointed out that the nonduplication of the results did not invalidate the model, but did suggest that the computer program should be corrected (1:5-6).

Abell continued his research in order to determine the validity of the data smoothing technique. The report indicated that the smoothing technique removed the randomness in the data sets for tonnage and flying hours and isolated the underlying trends between tonnage and flying hours. It also indicated that the coefficient of correlation (r) and the coefficient of determination (r^2) for the regression equations using smoothed data significantly increased compared to r and r^2 values computed for regression equations using nonsmoothed data. For example, the report showed an r value for a regression equation using nonsmoothed data as r = .04163, while the r value for the regression equation using smoothed data increased to r = .79056. The sum of squares error (SSE) for the regression equations using smoothed data decreased compared to the SSE computed for regression equations using nonsmoothed data (1:11-13).

Abell believed that if the smoothing technique was in fact isolating the trends in the flying hour and tonnage data sets, then the increased values of the coefficient of correlation and the coefficient of determination computed for the regression equations utilizing smoothed data sets were true indicators of the strength of the relationship

between tonnage and flying hours. The stronger relationship between tonnage and flying hours would improve the accuracy and reliability of the forecasts (1:13). Abell proved this by comparing nonsmoothed data forecasts with smoothed data forecasts using the mean absolute deviation (MAD) formula and determined the smoothed data forecasts were more accurate than the nonsmoothed forecasts (1:16-18).

Abell recommended that the regression model should be continued based on the following reasons:

1. it is dependable and defendable,
2. it is able to incorporate the effects of past trends into the estimate,
3. it is able to incorporate indicators of increases/decreases of future operations into the estimate,
4. provides a measure of the probable error in the estimate,
5. provides a measure of the strength of the relationship between tonnage movements and flying hours, the correlation coefficient. (1:22-23)

Abell believed the smoothing technique was justifiable, but required further evaluation. The report also recommended an investigation into the use of other regression models such as logarithmic or multiple regression models and the evaluation of alternative data smoothing techniques (1:23-25).

Strom Research (1989). This research, conducted by Captain Stephen L. Strom, was initiated as a result of concern expressed by Headquarters Unites States Air Force, Plans and Programs Division, Directorate of Transportation (HQ USAF/LETX) with respect to the effectiveness and accuracy of the overseas general cargo forecasting model used by DSXR. With the increase in budget reductions throughout the Air Force, and especially in the transportation system, more emphasis was placed on DSXR to make the most accurate SDT forecasts possible (25:10).

Strom's research was limited to MAC and MSC general cargo tonnage forecasting models for USAFE and PACAF. The research had two objectives:

1. Validate the current forecasting method used for computing tonnage estimates to derive future SDT budget requests.

2. If the current method's validity was not supported, develop a new forecasting model, using the same input data, that would produce more accurate and reliable tonnage estimates. (25:10)

In order to validate the current forecasting method, Strom tested DSXR's iterative linear regression forecasting technique by examining β_1 in the simple regression equation to determine whether it changed during the iterative process. If β_1 did not change, then the data was stationary and the iterative linear regression technique could be used. If it changed, then the iterative approach was invalid and could not be used (25:41). DSXR's simple regression model is;

$$Y = \beta_0 + \beta_1 X$$

where: Y = historical quarterly SDT tonnage
X = historical quarterly flying hours
β_0 = the y axis intercept
β_1 = flying hour parameter

A three step process was used to determine the stability of the flying hour parameter:

1. The flying hour parameter, β_1, for each iteration was computed.

2. The standard error ($s_{\beta 1}$) for each flying hour parameter was computed.

3. Using these standard errors, 95 percent confidence intervals for β_1 were established for each iteration using the following equation:

$$\beta_1 \pm t_{\alpha/2} s_{\beta 1}$$

where: $t_{\alpha/2}$ = the value of the two-tailed test-statistic for α = .05 and n - 2 degrees of freedom. (25:41-42)

Strom concluded that if any of the confidence intervals did not overlap, then the flying hour parameter changed during the iterations and

consequently invalidated DSXR's linear regression iterative technique. The following hypothesis test was conducted to statistically determine whether the flying hour parameter changed during the iteration process:

H_o: $\beta_{1,8} = \beta_{1,9} = \ldots = \beta_{1,40}$

H_a: $\beta_{1,i} \neq \beta_{1,8} = \ldots = \beta_{1,n}$

where: i = any one iteration conducted with 8 to 40 periods of data
n = total number of iterations conducted excluding i

Test Statistic: $\beta_1 \pm t_{\alpha/2} s_{\beta 1}$

where: $t_{\alpha/2}$ = the value of the test-statistic for $\alpha = .05$ and n - 2 degrees of freedom.
Rejection Region: Reject H_o if any two of the regression iteration confidence intervals did not overlap. (25:58-59)

Strom's research statistically invalidated the iterative linear regression method used to predict MSC SDT tonnage for PACAF and USAFE by showing how the flying hour parameter (β_1) had statistically changed after iterations were conducted. The null hypothesis for each of the two models was rejected based on a 95 percent confidence interval. This instability in the flying hour parameter implied that the models were invalid for predicting future tonnage requirements. Strom's research did not statistically invalidate the models used to predict MAC SDT tonnage for PACAF and USAFE, but did cast suspicion on the validity of the linear regression iterative technique. Strom continued the research in order to develop Box-Jenkins time series forecasting models (25:80-81).

The Box-Jenkins time series model was used because it can identify patterns in historical time series data and use the patterns to make forecasts. A computer software package (TIMES) was used to accomplish the time series analysis, the model building and evaluation, and the

forecasting requirements. Four steps were used in the development of the Box-Jenkins (BJ) time series model:

1. Identification of any patterns in the time series.
2. Model specification based on these identified patterns.
3. Diagnostic tests to ensure the appropriate model is specified.
4. Hypothesis testing and forecasting. (25:63-64)

Based on the raw data, autocorrelation function (ACF), and partial autocorrelation function (PACF) patterns, four ARIMA (autoregressive integrated moving average) models were selected:

ARIMA (1,2,2) for MAC tonnage destined for USAFE,

$$Y_t = Y_{t-2} + \sigma_2(Y_{t-2} - Y_{t-4}) - \alpha_1 e_{t-1} - \alpha_2 e_{t-2} + e_t$$

ARIMA (1,1,1) for MSC tonnage destined for USAFE,

$$Y_t = Y_{t-1} + \sigma_1(Y_{t-1} - Y_{t-2}) - \alpha_1 e_{t-1} + e_t$$

ARIMA (1,1,2) for MSC tonnage destined for PACAF,

$$Y_t = Y_{t-1} + \sigma_1(Y_{t-1} - Y_{t-2}) - \alpha_3 e_{t-3} - \alpha_4 e_{t-4} + e_t$$

ARIMA (0,0,1) for MAC tonnage destined for PACAF,

$$Y_t = e_t - \alpha_1 e_{t-1} + \mu$$

where: Y_t = the forecasted time series values

Y_{t-i} = historical value of the time series

σ_i = the AR (autoregressive) parameter at period i

α_i = the MA (moving average) parameter at period i

e_{t-i} = the error associated with period t - 1.

μ = the mean of the time series. (25:68-70)

Eight diagnostic tests were performed on the models:

1. The residuals were plotted and evaluated to determine changes in the variance.

2. The residual autocorrelation function was inspected for any significant values or spikes.

3. The Portmanteau Lack of Fit Test (Q-statistic) was evaluated to determine how well the model fits the data.

4. The cumulative periodogram of the residuals was evaluated to determine linearity.

5. A histogram of the residuals was reviewed to determine normality.

6. The sum of the squared error (SSE) and the mean squared error (MSE) were evaluated to determine whether the model was a good predictor of the time series.

7. The Fourier Transform of the autocorrelations (Power Spectrum) was reviewed to determine whether problems existed in the models.

8. The Schwartz statistic (Bayesian Information Criterion (BIC)) was reviewed to assess the goodness of fit as well as penalizing for complexity within the model. (25:45-47)

The accuracy of the BJ model forecasts was evaluated using the mean absolute percent error (MAPE) equation and compared to the forecasts and MAPE for the DSXR model. The BJ models for PACAF and USAFE MAC tonnage were statistically more accurate than the DSXR model forecasts. The BJ PACAF and USAFE MSC models failed to forecast more accurately than the DSXR models for PACAF and USAFE MSC tonnage (25:83-86).

Strom's research conclusion recommended further research to develop a valid forecasting model that is significantly more accurate than the one presently used by DSXR. Strom recommended an econometric model with multiple independent variables (25:88).

Navy and Army SDT Program

Navy SDT. A telephone interview was conducted with Mr. Bill Wall (Budget Analyst) of the Navy Material Transportation Office (Service Wide

Transportation Branch (SWT)), to determine how the Navy forecasts their SDT requirements. The Navy's MAC and MSC requirements are manually forecasted using a subjective method which consists of looking for trends and seasonal fluctuations in the historical tonnage data for each channel of traffic (51 channels for MAC, 450 - 500 channels for MSC) and predicting a short range forecast (100 days) and a long range forecast (2 years) based on the characteristics of the historical data (a channel is a particular cargo route with a specific point of embarkation (POE) and debarkation (POD), i.e. Atlantic Coast (U.S.) to Europe). Once the tonnage forecasts are established, the SDT budget requirement is determined by using cost figures from applicable MAC and MSC regulations (27).

Although no standard statistical procedures are used to develop the forecasts, the data was reported to be consistent which indicates the past tonnage requirements are a good predictor of future tonnage requirements. Simple averages of past tonnage requirements have yielded good results for the Navy, but there are some problems the Navy encounters with forecasting SDT. First, it is difficult to determine in advance the geographical area (i.e. Greenland, Mediterranean) where the carrier groups and squadrons are going to be conducting exercises. Without knowing the geographical area, it's impossible to forecast the channel tonnage requirement to support the exercises. Second, the historical data for some of the larger cargo volume channels are disrupted (large peaks and valleys) from SDT policy changes and budget reductions. This makes it difficult to forecast new requirements based on past historical data which have been disrupted from these changes.

The smaller cargo volume channels do not seem to be affected by these changes (27).

Army SDT. A telephone interview was conducted with Mr. Robert Saylor (Traffic Management Specialist) of the Army Materiel Command (AMC) (Logistics Control Authority (LCA)), to determine how the Army forecasts their SDT requirements. The Army's MAC and MSC requirements are manually forecasted by each command using a subjective method which consists of looking for trends and seasonal fluctuations in the historical tonnage data for each channel of traffic. The AMC then combines the forecasts and develops the Army's CONUS outbound, CONUS inbound, and intratheater SDT requirements. Cost figures are used from the applicable MAC and MSC regulations to develop the overall SDT budget (24).

The Army appears to have some problems with forecasting SDT. First, the forecasts submitted by each of the commands are usually inaccurate. Although no standard statistical procedure is used to develop the forecasts, the Army is in the process of developing an automated forecasting system using historical data and a Winter's time series model. Second, MAC and MSC tonnages are reported by channel, but the Army develops their forecasts by command. Tonnages reported by command are more useful for developing forecasts (24).

The Army reported the same problems with policy changes and budget cuts as the Navy. They also reported that simple statistical methods such as a simple average can yield good forecasting results, but they believe the Winter's time series model will be capable of extracting the trend and seasonal characteristics of the past historical tonnage and will ultimately improve their forecasting accuracy (24).

Summary of the SDT Reports and the Army/Navy SDT

The past literature indicates the forecasting methodology used by DSXR has been extensively evaluated by other researchers. Most of the research evaluated the use of the flying hour program to forecast SDT and attempted to justify its use in forecasting models or make recommendations on how to improve the models.

The Foster Report indicated the methodology was not totally valid, but no other method was explored for possible use. The LMI report cited rate, workload, and policy changes as the most important causes of SDT requirement changes. This means SDT forecasting models should be capable of modeling these changes. Grayson's report identified the problem of overestimated programmed flying hours and the possible negative impact it would have on SDT forecasting accuracy. Grayson believed the use of the total flying hour variable was too much of a generalization and did not account for particular types of aircraft. Abell's study indicated DSXR's simple regression model should be continued, but other models should be evaluated such as multiple regression models. Strom's research addressed DSXR's iterative simple regression development technique and discovered it was not always a statistically accurate technique. Strom developed Box Jenkins time series models, but only two models out of the four that were developed were statistically more accurate than the DSXR models.

DSXR is presently in the process of separating additional programs from the general cargo category such as aircraft engines and communication equipment. DSXR believes this will allow them to track the shipment of these items separately from the general cargo category and provide more accurate forecasts for SDT. Three of the previous researchers (Lamb, Sarnaki, and Strom) developed different forecasting

models in order to improve forecasting accuracy, but none of the models have been implemented by DSXR.

DSXR's forecasting methodology is different in some aspects from the Army and Navy forecasting methodology. For budgeting purposes, DSXR forecasts the total MAC and total MSC tonnage requirement by geographical area based on flying hours using simple regression models. The Army and the Navy subjectively forecast their MAC and MSC requirements by channel and add them together to arrive at a total MAC and MSC tonnage requirement for funding. DSXR does subjectively alter their forecasts produced by the regression models if the forecasted tonnage values appear too low or high with respect to current historical levels of tonnage. Once the tonnage forecasts are established, the tonnage forecasts are converted into dollars for funding.

This research addressed the recommendations for improving the forecasting models made in the prior research. Specifically, this research examined the use of multiple regression models with other independent variables such as military population and flying hours by the type of aircraft. Also, previous research indicated conventional mathematical techniques for forecasting SDT requirements have yielded marginal results. Since other neurocomputing research has proven neural networks can be successfully employed in applications that have yielded poor results through conventional mathematical techniques, neural network models were developed and evaluated for SDT forecasting accuracy. The next section presents an introduction to the back-propagating neural network that was used in this research.

Neurocomputing

The artificial intelligence community has long been interested in developing machines that mimic human characteristics. Expert systems and robotics are two results of this quest. Expert systems attempt to capture the knowledge of one or more human domain experts in a computer program, while robotics deals with controlling the visual and tactile aspects of robotic activity.

Neural networks are an adaptive information processing system that mathematically develops associations between particular inputs and a desired output. The network is given sets of example inputs (independent variables) and desired outputs (dependent variables) and extracts a relationship between the inputs and outputs by analyzing the input/output pairs (13:36).

Neurocomputing is loosely based on the information processing capability of the human brain which is composed of thousands of biological neurons and neural connections (13:37). A biological neuron is composed of three parts: the cell body which contains the nucleus; the dendrites which receive the input signals and transmit them to the cell body; and the axon which transmits the signals from the cell body and dendrites (5:635). The neuron is the basic cellular unit of the central nervous system (brain, spinal cord) and the peripheral nervous system (neurons outside the central nervous system). The nervous system's major function is internal communication and this function is accomplished by an electrical and chemical transmission from one neuron to other neurons (5:635-640).

A single neuron in the brain may receive signals from thousands of other neurons by synaptic connections between its dendrites and the axons

of other neurons, but depending on the summation of the signals, the neuron may or may not be excited (fire) and initiate an impulse (5:640). The neurons in the brain react to inputs based on how it is genetically organized and what it has previously learned. The neural network functions in a similar fashion. Input and output training examples are presented to the network which, in turn, forms mappings or associations between the inputs and outputs. The strength of a neural network comes from its ability to associate new inputs with a particular output based on the input/output associations it has developed through training (13:37).

Neural Networks. Neurocomputing attempts to 'train' computer simulated neurons to fire based on particular inputs they receive. The computer simulated neurons are known as processing elements (Figure 1) and consist of a small local memory and processing power (contained in the processing node). The local memory is used to store previous computations and a value known as an adaptive coefficient or connection weight. Each weight determines the connection strength between processing elements and is used to increase or decrease the input signal from other processing elements. Weights can be negative (inhibitory, decreases output signal from the processing element) or positive (excitatory, increases output signal from the processing element). Each processing element receives many different input signals, but only transmits one output signal. This output signal is transmitted to other processing elements and acts as the input signal to each of them. A processing element computes output signals or values by multiplying the input values received from other processing elements with the corresponding connection weights and summing them together.

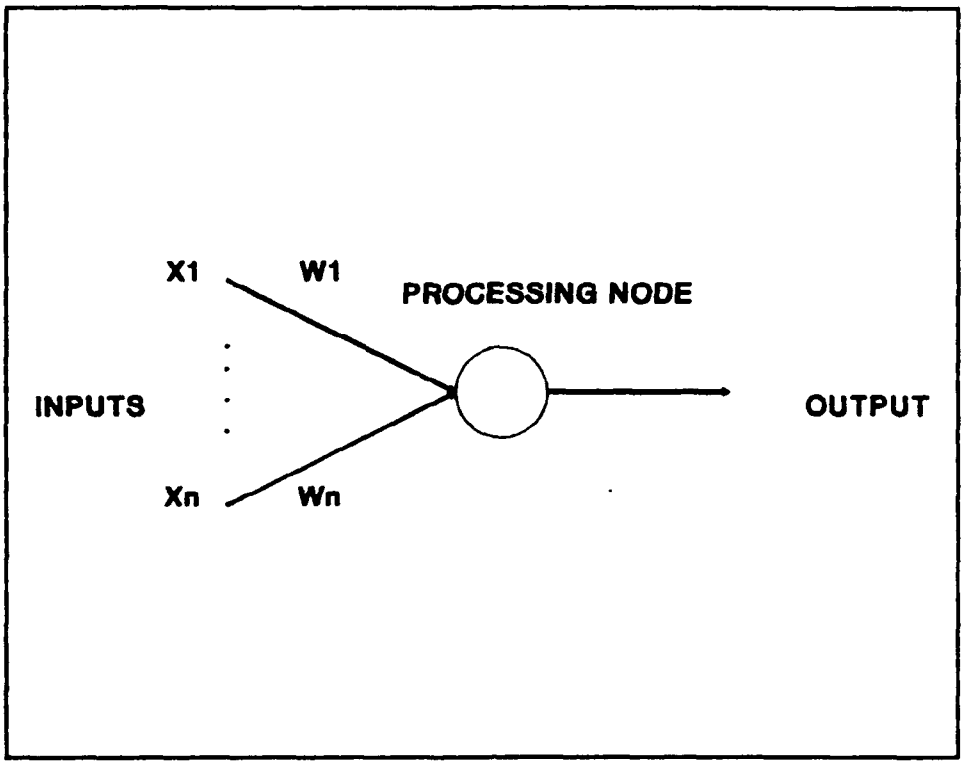

Figure 1. A Neural Network Processing Element

Figure 2 shows an example of a processing element with three inputs values and one output value. The three input values (x_1, x_2, and x_3) are multiplied by their corresponding weight values (w_1, w_2, and w_3) and the products are summed with a neuron bias value (w_0 = +1) (the neuron bias is added to mathematically improve the network's performance). The processing element then modifies the weighted summation with a transfer function which is typically the sigmoid function (13:36-38). The equations that are used are presented below;

$$z = x_1w_1 + x_2w_2 + \ldots x_nw_n + w_0$$

$$f(z) = 1 / (1 + e^{-z})$$

where: z = weighted summation value with bias value
$f(z)$ = sigmoid function
x_i = input value
w_i = input weight value
w_0 = neuron bias value (+1)

n = the number of inputs
e = natural exponential function. (17:13-17; 28:44-45)

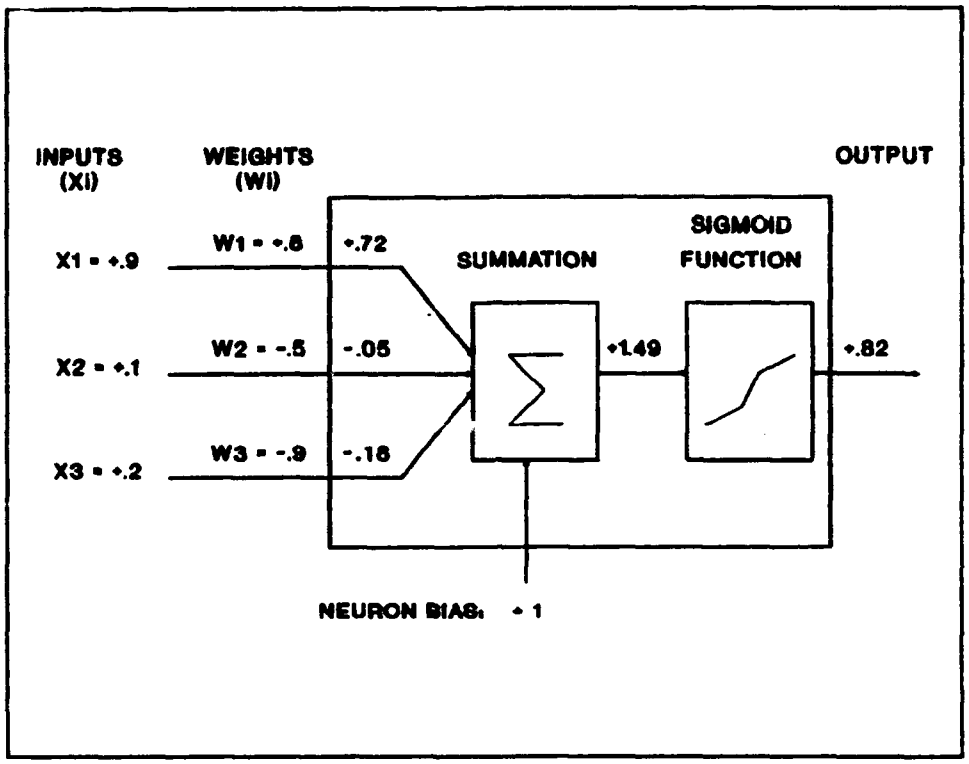

Figure 2. A Neural Network Processing Element Activation (Example 1)

Figure 3 is another example of the same processing element, but with three different input values. Figure 4 shows how the summation value is transformed by the sigmoid function. The sigmoid 'squashes' the summation value so that it ranges between 0 and 1 (28:45).

The transfer function represents the processing power of the processing element and is usually a nonlinear positively increasing function. Figure 5 shows two other common nonlinear functions, the hard limiter function and the threshold logic function (17:4).

The network is presented, in an iterative fashion, pairwise sets of input (independent variables) and desired output values (dependent variables) during a training process. The network produces a predicted

output each time the input/output set is presented to the network. The network compares the predicted output with the desired output value from the training set and determines the error between the two values. The network adjusts or changes the connection weights between processing elements in a way that minimizes the error. Connection weights that contribute the most to the error are changed first. Through this iteration process, the network continues to reduce the error between the predicted outputs it produces and the desired outputs until it reaches a global minimum error (28:47-53).

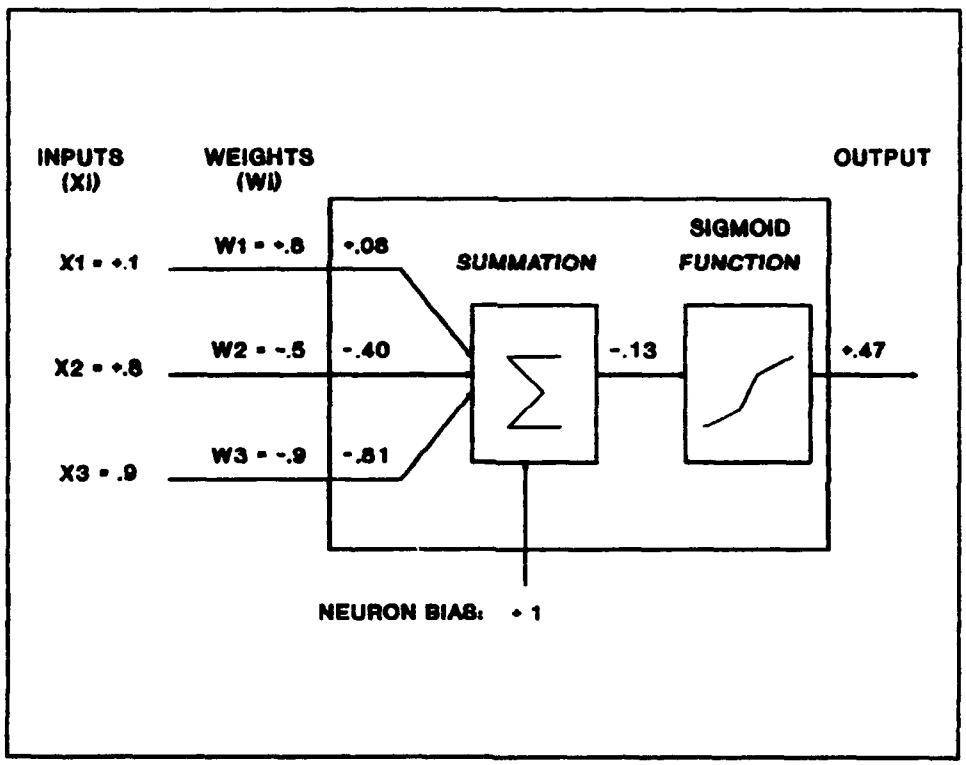

Figure 3. A Neural Network Processing Element Activation (Example 2)

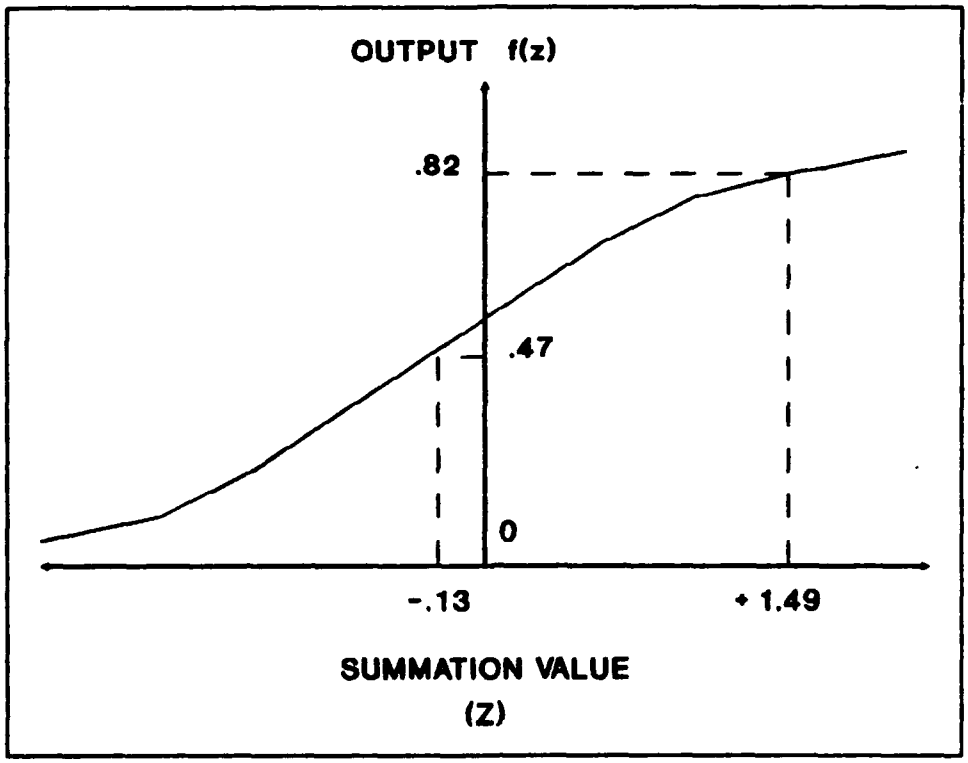

Figure 4. Sigmoid Function

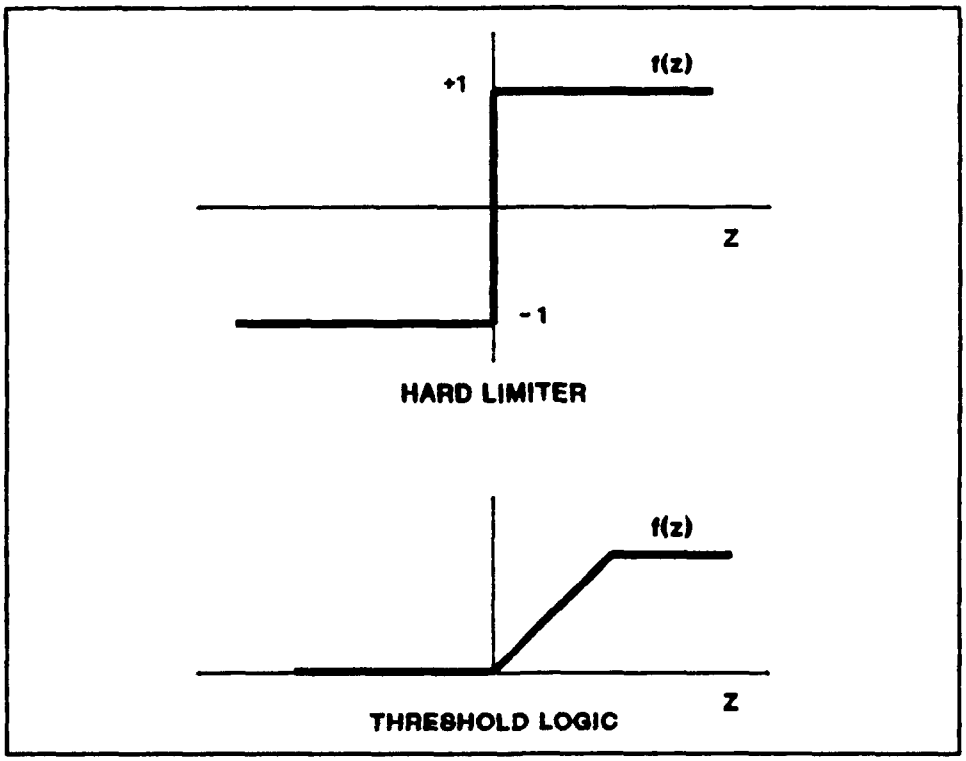

Figure 5. Hard Limiter and Threshold Logic Function (17:5)

Single Layer Perceptron. The earliest implementation of neurocomputing was the development of a network in the late 1950's that had the ability to recognize simple patterns that could be separated by a single plane in a two dimensional space. This single layer network was called a perceptron and subsequent work by researchers continued the experimentation and understanding of the perceptron's capabilities (22:20-21).

The single layer perceptron is limited to deciding whether input patterns belong to one of two classes. This is accomplished by the mathematical computations that the processing element performs. The processing element in the single layer receives weighted input signals and sums them together. A threshold factor (β) is added to the sum and is passed through a hard limiting nonlinearity to produce an output of -1 or +1. An output of -1 would correspond to one class, while an output of +1 would correspond to the other. A decision boundary is formed by a hyperplane in a two dimensional space that separates the two classes into separate half planes (17:13). Figure 6 displays the input/output relationship of the processing element and the decision boundary.

The perceptron is initialized by setting the connection weights and threshold values to small random numbers, usually between -1 and +1. The perceptron is trained by presenting examples of inputs and desired output. During this training process, the perceptron compares the output it computed (actual output) to the desired output and adapts or corrects the connection weights if an error exists. The rate of adaptation is controlled by a momentum or gain term (n) that ranges between 0 and 1. Low gain terms result in stable weight changes and an averaging of past

input values while high gain terms cause large weight modifications for changes in the input values.

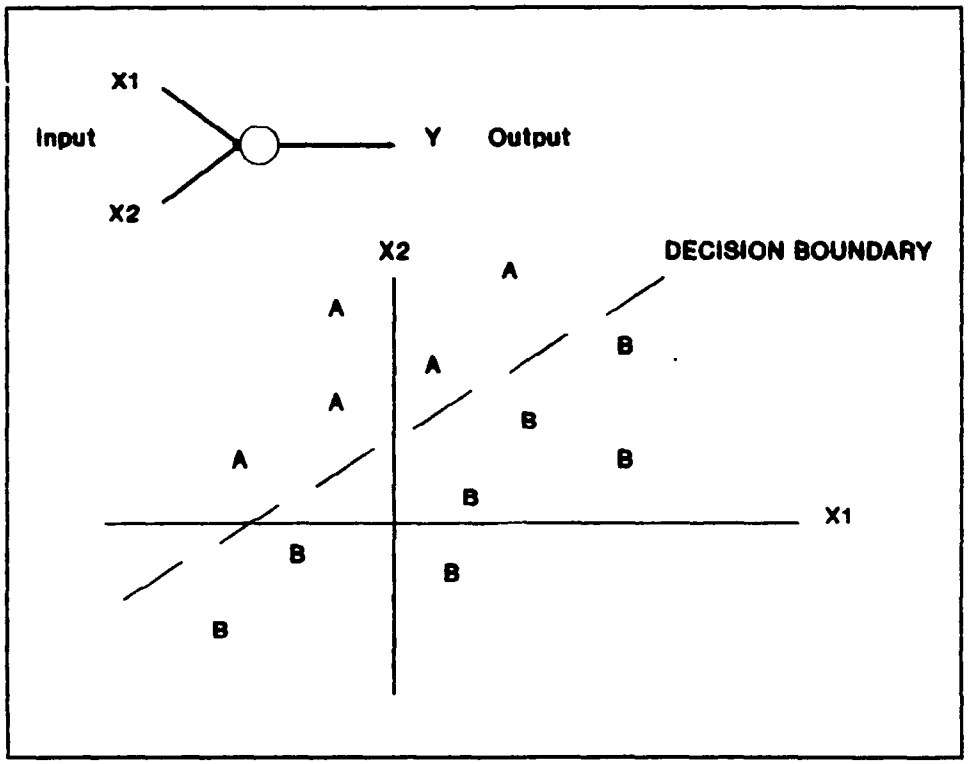

Figure 6. Single Layer Perceptron and Decision Boundary (17:13)

Through the training process, the perceptron forms a decision boundary between the two classes. This process is called perceptron convergence. The original mathematical equations were modified to eliminate the possibility of oscillating decision boundaries. The modified method is called the least mean square (LMS) algorithm which minimizes the error between the desired output and the perceptron's actual output. The LMS algorithm follows a gradient descent heuristic by changing connection weights that maximize the change in error. The connection weights from inputs that contribute to the greatest error undergo the largest adaptation or correction. Other modifications

included replacing the hard limiting nonlinearity function with a linear or threshold logic nonlinearity and updating the connection weights during every trial depending on the error between the desired output and the actual output (17:14).

The first equation of the following two equations shows the mathematical formula used to calculate the actual output of the processing element. The second equation displays how the weights are adapted using the LMS method.

$$y(t) = f[\Sigma w_i(t)x_i(t) + \beta)]$$

$$w_i(t+1) = w_i(t) + n[d(t) - y(t)]x_i(t)$$

where: $y(t)$ = output

$x_i(t)$ = input values where $i = 1...n$ inputs

$w_i(t)$ = connection weights for each input

β = threshold value

f = a linear or threshold logic nonlinearity function

$d(t)$ = desired output (0 or +1)

n = gain term. (17:13)

In 1969, Minsky and Papert published a discussion of the perceptron that proved the single layer perceptron was incapable of performing complex pattern recognition problems. Soon after its publication, interest in perceptron research declined until 1986, when Rummelhart and others developed a training algorithm for the multi-layered perceptron.

Papert and Minsky exposed the limitations of the perceptron by demonstrating a number of problems the perceptron could not solve. The perceptron convergence procedure is not an appropriate method to use when two classes cannot be separated by a hyperplane. The problem used by

Papert and Minsky to demonstrate this was the exclusive-OR problem (Table 4). In this problem, the two classes are disjoint and cannot be separated by a single hyperplane. Figure 7 shows how the two classes cannot be separated by a single hyperplane.

Table 4

Exclusive OR Problem (17:14)

Input values		Output values	
X1	X2	Y	Class
0	0	0	B
1	0	1	A
0	1	1	A
1	1	0	B

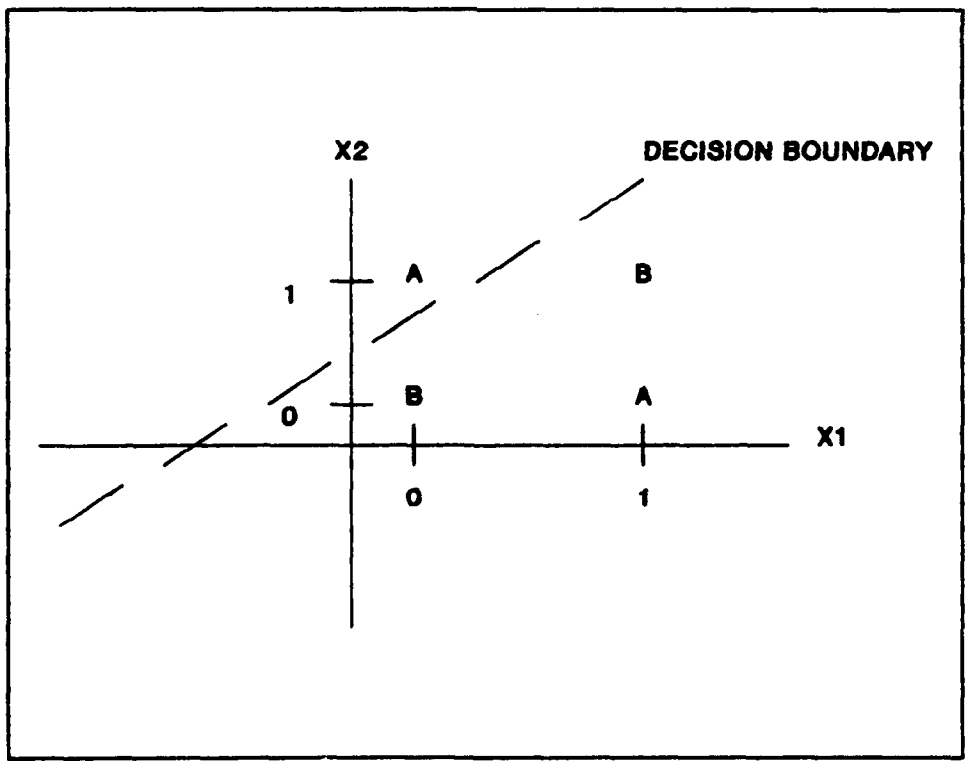

Figure 7. The Exclusive-OR Decision Region (17:14)

Training algorithms were later developed for multi-layered networks. Multi-layered networks have much more processing power than the single-layer networks and can learn complicated multi-dimensional associations between input patterns and output values. The development of the backpropagating training algorithm enabled the multi-layer perceptron to overcome the limitations of the single-layer perceptron (17:15-17).

Multi-Layer Perceptron. The multi-layer perceptron contains at least one additional layer of processing elements between the input and output processing elements. The additional layers are called hidden layers because they are not directly connected to inputs or outputs. Each hidden layer processing element adds capability for the network to recognize associations between the inputs and outputs. Usually, a network will not contain more than two hidden layers since most problems can be solved with one or two layers. The output signal from one processing element is the input to other processing elements in subsequent layers and each element affects the performance of the entire network. The connections between processing elements can be fully connected or randomly connected (13:38). Figure 8 displays a multi-layer perceptron with 8 inputs and one output with two hidden layers consisting of 12 processing elements in the first layer and 5 processing elements in the second layer (17:15). Unlike the single-layer perceptron, the multi-layer perceptron can form bounded or unbounded convex decision regions. A bounded convex region means a particular class is contained in a particular finite region while an unbounded convex region contains a class in an infinite region. The convex regions are constructed by the intersections of half plane regions which are

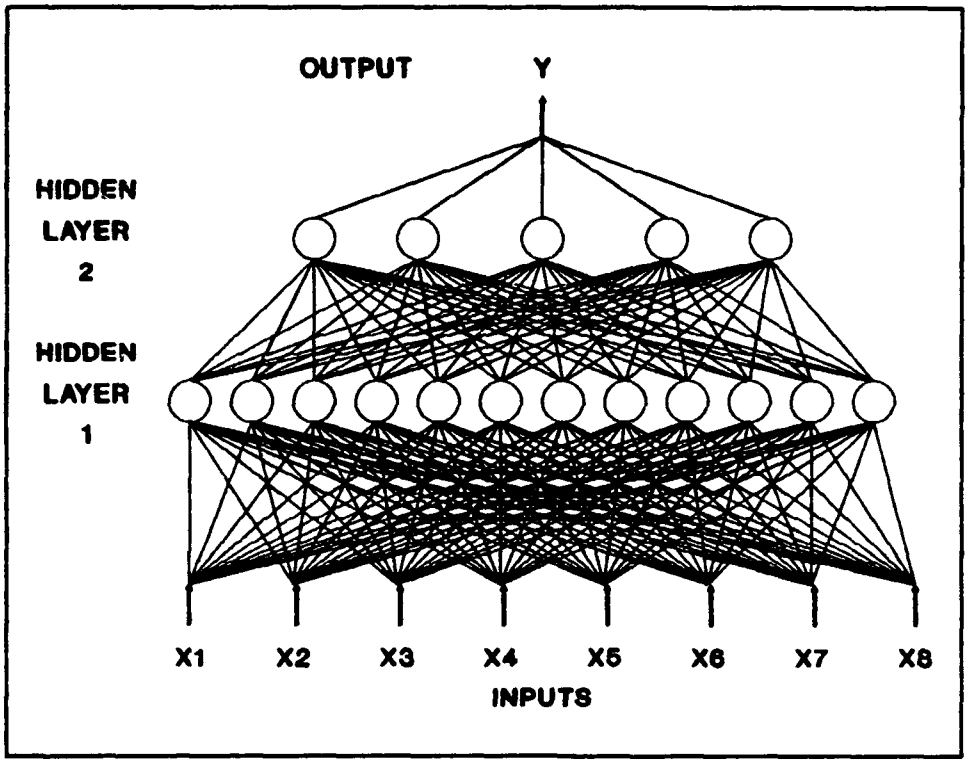

Figure 8. A Multi-Layer Perceptron

formed by the processing elements in the first hidden layer. Each of the processing elements in the first layer acts like a single perceptron and forms a half plane region bounded by a hyperplane. The decision region for a particular class becomes the intersection of all the half planes that are formed by each of the processing elements. The complexity of the decision regions increases as the number of additional layers and processing elements increase. The number of sides for a convex region are limited to the number of procession elements in the first hidden layer. Figure 9 displays the decision region formed by a multi-layer perceptron to solve the exclusive-OR problem (17:15-16).

The shapes of the decision regions can change depending on the types of transfer functions used by the processing elements. When sigmoidal nonlinearities are used instead of hardlimiting nonlinearities,

the decision boundaries are curved instead of straight line segments. Networks form these decision regions by using the backpropagation training algorithm (17:16).

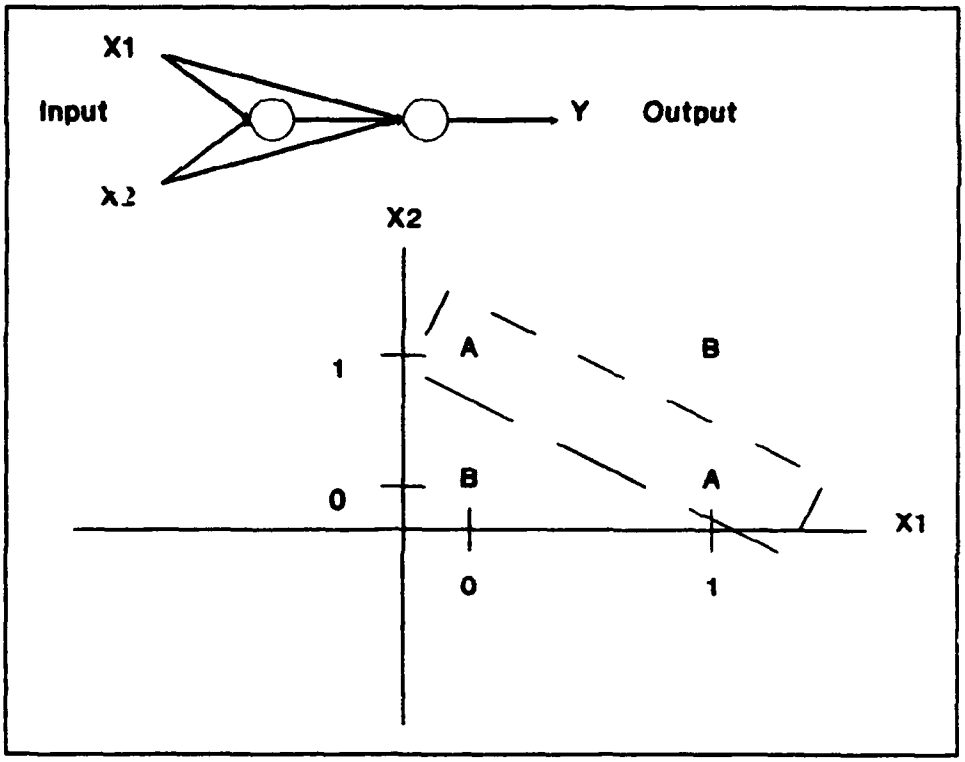

Figure 9. Multi-Layer Perceptron Exclusive-OR Decision Region (17:14)

Backpropagation. The backpropagation training algorithm incrementally reduces the global error between the desired output and the actual output. It is a generalization of the LMS algorithm and uses a iterative gradient technique to minimize the global error of the network. The error is minimized by adapting or correcting the connection weights in the hidden layers. This is accomplished by 'backpropagating' the error from the output processing elements back through the processing elements in the hidden layers. The gradient aspect of the technique

ensures outputs from hidden layer processing elements that contributed to the greatest amount of error are modified the most (12:1-20; 17:17).

The initiation of the backpropagation training algorithm is similar to the initiation of the LMS training algorithm for the single-layer perceptron. The initial connection weights and threshold values are set to small random numbers between -1 and +1. The network is presented input and desired output values during the development process and the connection weights between the processing elements in the hidden layers are modified to minimize the error between the actual output and the desired output. Once an acceptable error is achieved, the development process is terminated.

Starting at the output layer and working back to the previous layers, the weights are adjusted as follows;

$$w_{ij(t+1)} = w_{ij(t)} + n\delta_j x'_i$$

where: $w_{ij(t)}$ = connection weight from an input node i node j at time t
n = gain term
δ_j = error term for node j
x'_i = output from node i. (17:17)

The error terms are calculated differently depending on where the node is in the network. For output nodes, the error term is determined by the following equation;

$$\delta_j = y_j(1-y_j)(d_j - y_j)$$

where: δ_j = error term for output node j
y_j = actual output for output node j
d_j = desired output for output node j. (17:17)

For hidden nodes, the error term is;

$$\delta_j = x'_i(1-x'_i)\Sigma\delta_k w_{jk}$$

where: δ_j = error term for hidden node j
x'_i = output from node j
δ_k = error term for a forward node k
w_{jk} = connection weight from node j to forward node k. (17:17)

Types of Networks. Neural network models are specified in three ways: network topology, processing element characteristics, and network training rules. The network topology determines the mapping of connections between processing elements and influences the information processing capability of the network by controlling what data each processing element receives. The processing element characteristics are the type of transfer function and mathematical formula used to combine the input values and connection weights. Network training rules determine how and when the connection weights are adapted or changed in order to improve the performance of the network. Network topology, processing element characteristics, and network training rules determine the type of network (17:4).

There are six important types of networks (shown in Table 5) that can be classified by binary or continuous valued inputs and supervised or unsupervised training. Binary inputs are zero and one while continuous valued inputs are usually a number between zero and one.

Table 5

Types of Neural Networks (17:6)

Binary input	Continuous Valued Input
Supervised:	Supervised:
Hopfield Net	Perceptron
Hamming Net	Multi-layer Perceptron
Unsupervised:	Unsupervised:
Carpenter/Grossberg Classifier	Kohonen Self-Organizing Feature Maps

Under supervised training, the network is presented training data with input and desired output values. The network compares its actual output with the desired output and corrects itself by making changes to the connection weights for each of the processing elements. The process of presenting the training data is iterative and the network gradually corrects itself according to its training rules (13:38; 17:7).

Unsupervised training allows the network to make changes by itself. The network is presented training data with only the input values. The network organizes itself internally by responding to input values with different processing elements. In other words, the processing elements group themselves to respond to a particular set of input values or a closely related set of input values (13:38; 17:7).

Neural Network Applications. Neural networks applications are grouped into two categories, pattern recognition and generalization. In pattern recognition applications, the networks are developed to extract patterns from distorted inputs or inputs with added background noise (28:2-3). This application is used for speech and handwritten character recognition because there are many slightly distorted ways to pronounce the same word or write the same letter.

The generalization category of neural network applications consists of two types of problems, classification and prediction. Networks used in these applications are presented new inputs which are distinct from the inputs used in network training. For classification applications, the network groups the new input with other similar types of inputs and generates a particular output for that group. The output response signifies the group into which the new input was classified.

Networks for prediction applications are trained to recognize underlying patterns in a particular data set. When a new input is presented, the network generates an output based on the underlying data patterns that were previously learned (7:443).

The remainder of this section presents two neural network forecasting applications. The first application compares the forecasting accuracy of several neural network models to regression models in predicting ratings for corporate bonds. The second application compares the forecasting accuracy of neural network models to several standard time series forecasting techniques.

Neural Network Forecasting Applications

Neural Network vs. Regression (Bond Rating Problem). In this research, Soumitra Dutta used a neural network to predict the ratings for corporate bonds. In the past, the use of conventional mathematical modeling techniques to solve this problem have produced poor results. This was a generalization classification application of a neural network in a domain lacking a well defined theory or model. The underlying functional form or mathematical model that determines a corporate bond rating (dependent variable) from various independent variables is not well defined or known. Statistical techniques require an assumption or correlation concerning the functional form relating independent variables and dependent variables. The reason for utilizing a neural network for this type of problem is that the network does not need to know the apriori functional form. The network extracts the underlying data patterns from input-output training pairs. This research indicated that

the neural networks performed better than the multiple regression models (7:443-446).

The researchers selected ten financial variables (independent variables) to predict bond rating (dependent variables). Different neural network and regression models were developed using ten variables and six variables as independent variables. Forty seven sets of bond ratings and financial values were randomly selected. Thirty sets were used to train the neural networks and develop the regression models, while the other seventeen were used to test the performance of both models (7:446-448).

The Berkeley Interactive Statistical Package was used to develop the regression coefficients and t-statistics. A neural network simulator was used to develop 2 layered (input and output layer) and 3 layered (1 hidden layer with a varying number of nodes) network configurations. The output responses to the test data for the neural network and regression models were compared to determine which model was more accurate (7:448).

The results of the tests indicated the neural networks significantly outpreformed the regression models in predicting bond rating. The success rate for predicting bond rating was 88.3% (2 layered, 10 variable neural network) compared to 64.7% for the ten variable regression model. The 3 layered neural networks had a smaller total squared error for the learning data than the 2 layered network, but there was no significant difference in the predictive ability using the test data. When the neural network models were in error, the magnitude of the error was one rating while the regression models were often in error by several ratings (7:448-450).

Application of Neural Networks in Time Series Forecasting. The researchers in this study (Brian Huffman and Thomas Hoffmann) explored the use of neural networks in time series forecasting. The researchers compared the forecasting accuracy of neural networks with other conventional techniques such as moving averages, simple exponential smoothing, Winter's exponential smoothing, and naive or random walk approach. Two types of time series were examined, generating functions with and without noise and real-world data. Forecasting accuracy was evaluated on five criteria:

1. Average algebraic error
2. Standard deviation of algebraic error
3. Minimum error
4. Maximum error
5. Average Absolute error. (14:162)

The time series that were used in the research are listed below.

1. Sine wave function without slope (Sin00, Sin10, Sin30). The average value of the sine wave was 100 and was represented by the equation:

$$F(t) = 100 + 50(\sin(t*30/360)) + \epsilon$$

where,

$F(t)$ = function's value during period t

ϵ = uniformly distributed random error. (14:162)

2. Sine wave function with slope. A noiseless slope time series was added to sine wave function Sin00 to form the following equation:

$$F(t) = 200 + .8(t) + 50(\sin(t*30/360)) + \epsilon$$

where,

$F(t)$ = function's value during period t

ϵ = uniformly distributed random error. (14:162)

3. The actual data represented the number of international airline passengers during a high growth decade. The researchers believed this empirical data were more complex since it had an upward trend and seasonal characteristics with no known generating parameters (14:163).

The researchers wanted to put the simple moving average and simple exponential smoothing technique on equal footing with the neural networks as much as possible for comparison purposes. The simple moving average used N periods where N represented the number of inputs to the neural network.

The simple exponential formula used an alpha value of 2/(N+1) where N was the number of months used for inputs to the neural network model. Winter's model (alpha, beta, and gamma value of .15) was the most data intensive model and required a 12 month period to develop the seasonal constants. Four different neural networks were used with 3, 4, 5, and 9 input nodes and one output node for each model. Each of the forecasting models made 192 predictions and the results were compared using the above listed criteria (14:163).

The results of the sine wave and empirical data experiments indicated the neural networks were generally superior to the other methods, but under certain conditions other methods did outperform the networks on some of the criteria. The Winter's model also appeared to do very well, but the researchers believed precomputing the seasonal factors for the model may have biased the results in its favor. The researchers noted that more than three inputs and more than a year's worth of data may actually decrease the performance of the network. They also stated

that the networks seemed to forecast with greater accuracy, but that the predictive ability of the networks tended to be biased either positively or negatively (14:164).

Chapter Summary

This chapter presented previous SDT reports and research that have had an impact on forecasting future SDT requirements. This research was based on several of the further research suggestions from previous reports and research. The second part of this chapter examined the Navy and Army methodologies for forecasting SDT requirements. Both the Army and Navy relied on subjective evaluations, but noted that simple techniques such as the simple average of historical requirements produced adequate results. Part three of this chapter was a background on neurocomputing and the backpropagating neural network which was used in this research. The fourth part of this chapter presented previous research on the use of neural networks in two different types of forecasting applications. One report compared the forecasting accuracy of neural networks to the forecasting accuracy of multiple regression models. The second report compared the forecasting accuracy of neural networks to the forecasting accuracy of several conventional time series forecasting models. Overall, the neural networks produced more accurate forecasts compared to the other models.

III. Methodology

This chapter presents the methodology used to accomplish the research objectives and is divided into five parts. The methodology begins with the collection of the flying hour data by aircraft type and the military population data. The second part is a data analysis methodology section which consists of plot evaluations, trend and seasonal analysis, business cycle analysis, and a time series analysis. DSXR's simple regression model validation and forecasting evaluation methodology is presented in part three. Part four presents the multiple regression development, validation and forecasting evaluation methodology. Finally, part five is the neural network development and forecasting evaluation methodology.

Collection of the Data

One objective of this research was to develop multiple regression and neural network models using aircraft flying hours by type of aircraft and military population variables in order to increase SDT forecasting accuracy. Presently, DSXR receives a product of the G033B system (Aerospace Vehicle Inventory Status and Utilization Reporting System (AVISURS)) for updating their historical data base of quarterly total flying hours by geographical area. DSXR also receives the programmed flying hours by quarter and uses this information to determine future tonnage requirements.

The G033B system can produce reports showing flying hours categorized by mission design (MD) or mission design series (MDS) for each geographical area (4). For this research, PACAF and USAFE flying hours by MD from FY 1985/1 to 1988/2 (14 quarters) were used to develop

the PACAF and USAFE MSC multiple regression and neural network models. Six quarters (FY 1988/3 to 1989/4) were used to test forecasting accuracy. Flying hours by MD from FY 1985/3 to 1988/4 (14 quarters) were used to develop the PACAF and USAFE MAC multiple regression and neural network models. Five quarters (FY 1989/1 to 1990/1) were used to test forecasting accuracy.

PACAF and USAFE military population data were obtained from AFLC Director of Military Personnel, Systems Division (DPMSD), and are categorized by yearly officer and airman manpower strength for each geographical area. Manpower strength is programmed for future outyears and DSXR currently receives this information for forecasting SDT for the subsistence program.

Data Analysis Methodology

Plots. Various plots of PACAF and USAFE MSC tonnage, MAC tonnage, flying hours, and military population variables were developed so that relationships could be graphically identified. Many of the plots were developed from data in the last appendix in this report (Appendix Z) which presents the PACAF and USAFE MAC, MSC and total flying hour data from FY 1978, quarter 1 to FY 1990, quarter 1.

Trend and Seasonal Analysis. PACAF and USAFE MSC and MAC quarterly tonnage and total quarterly flying hours were analyzed using Gardner's trend and seasonal analysis methodology. Gardner's methodology is based on comparing the variance of the actual data set with the variance of the difference between same quarters for each year, the first difference between each quarter, and the second difference between each quarter.

The data set with the lowest variance determines the strength of the trend/seasonal characteristics in the actual data set (9:44-50).

Business Cycle Analysis. The PACAF and USAFE MSC quarterly tonnage were analyzed using Gardner's business cycle pressure analysis methodology (MAC tonnage data were not analyzed). The pressure analysis methodology was used in this research to determine the peaks, troughs, turning points and other changes that have occurred in each of the MSC tonnage data sets (10:40-43). Pressure values are determined by comparing data from a particular time period with data from a previous time period (time periods that are compared have equal lengths). The pressure values are ratios of the data for time period (t) to the data for a previous time period (t-x) (where x is a pre-determined number of previous time periods). A pressure value above 100% indicates the present MSC tonnage is greater than the past year's tonnage for the same period (a value below 100% means the opposite is true).

For this research, 1/4 and 4/4 pressure plots were developed. A 1/4 pressure plot shows the comparison of the quarterly tonnage with the quarterly tonnage for the previous year. A 4/4 pressure plot shows the comparison of the sum of four quarters of quarterly tonnage with the same sum of tonnage for the previous year.

Pattern Identification. PACAF and USAFE MSC and MAC tonnage autocorrelations were computed using a statistics computer program (SAS, proc ARIMA) to determine whether there were any patterns (autoregressive (AR), moving average (MA)) in the data sets. The autocorrelation function was used to identify MA aspects and data stationarity while the partial autocorrelation function was used to identify AR aspects. The Q-

statistic was computed to test whether a data set was white noise (random).

DSXR Simple Regression Model Validation and Forecasting Evaluation

The DSXR simple regression models were replicated in this research (SAS, proc REG) in order to identify model specification problems. The models were statistically validated and evaluated (forecasting accuracy) with the following tests:

1. The Two Tailed T-test.

H_o: $\beta_1 = 0$

H_a: β_1 does not equal 0

Test Statistic: t

Rejection Region: $t < -t_{.025}$ or $t > t_{.025}$

where, $t_{.025}$ is based on n - 2 df.

The two tailed t-test proves whether the model is significant at the .05 level of confidence. The rejection of the null hypothesis indicates the independent variable contributes information for the prediction of the tonnage (dependent) variable.

2. R^2 Value Evaluation. The R^2 value (coefficient of determination) can be used to determine the predictive power of the model. The R^2 value indicates the model's fit to the data. This test measures the proportionate reduction of total variation or error associated with the use of the independent variables (21:89-90). A model with a coefficient of determination of approximately .70 or greater is normally considered an effective model for predicting the dependent variable, but R^2 values can be artificially forced to take a high value by adding more independent variables to the model even though the model

contributes no useful information for predicting the dependent variable (18:581). The R^2 value can be determined from the following equation;

$$R^2 = 1 - (SSE / SSY)$$

where: SSE = the unexplained sample variation

SSY = the total sample variation.

3. Residual Analysis. Plots of the residuals versus the predicted values were used to determine whether the residuals were randomly distributed (no heteroscedasticity problems). Plots of the residuals versus the independent variables were examined for random distribution (no problem with the assumption of linearity between tonnage and the independent variables). The Wilk Shapiro Test for Normality was used to determine whether the residuals were normally distributed. A properly specified model will have normally distributed residuals.

H_o: The residual distribution function is a normal distribution function.

H_a: The residual distribution function is not a normal distribution function.

Test Statistic: W

Rejection Region: $W < W_{.05}$

where, $\alpha = .05$, n = number of observations.

The rejection of the null hypothesis proves the residuals are not normally distributed and the model is not properly specified.

4. Outlier Detection. The studentized residuals were computed so that residuals falling beyond 3 standard deviations could be identified.

5. Durbin Watson (DW) Test. This test was used to test for the existence of first order autocorrelation (23:158-161).

H_o: The residuals are not autocorrelated.

H_a: The residuals are autocorrelated.

Test Statistic: DW

Rejection Region:

$$4 - d_l < DW < 4 \quad \text{(negative autocorrelation)}$$

$$0 < DW < d_l \quad \text{(positive autocorrelation)}$$

Acceptance (null hypothesis) Region:

$$2 < DW < 4 - d_u \quad \text{(no autocorrelation, negative test)}$$

$$d_u < DW < 2 \quad \text{(no autocorrelation, positive test)}$$

Indeterminate Results:

$$4 - d_u < DW < 4 - d_l$$

$$d_l < DW < d_u$$

where d_l and d_u are based on k independent parameters and n observations.

The acceptance or rejection of the null hypothesis indicates whether a positive or negative autocorrelation problem is evident in the residuals.

6. Forecasting Accuracy. The forecasting accuracy of the models was evaluated in three ways, mean absolute error (MAE), minimum absolute error, and maximum absolute error. Out of the three, the MAE value was the most critical value for evaluating forecasting accuracy. The following equation was used to calculate the MAE value;

$$MAE = (|E_1| + |E_2| + \ldots + |E_n|) / n$$

where:

$|E_i|$ = absolute error between the predicted and actual value at time period i

n = number of periods in the forecast series.

The MAE values for the PACAF and USAFE MSC models were calculated with six quarters (FY 1988/3 to 1989/4, n = 6) while the MAE values for the PACAF and USAFE MAC models were calculated with five quarters (FY 1989/1 to 1990/1, n = 5).

Multiple Regression Model Development, Validation, and Forecasting Evaluation

The multiple variable regression model is an extension of the simple linear regression model. There are two reasons for using it:

1. To reduce the random error denoted by ϵ along with its variance denoted by s^2. This reduces the prediction and confidence intervals and increases the precision of the intervals.

2. To eliminate bias by including independent variables that contribute to the prediction of the dependent variable. (29:71)

Before using the multiple variable regression models, four assumptions must be met concerning the random error ϵ:

1. The mean of the probability distribution for ϵ must be equal to zero.

2. The variance of the probability distribution of ϵ is constant for all given sets of independent variables.

3. The probability distribution of ϵ is normal.

4. The error associated with the a y value (dependent variable) is independent of any other y values. (18:501,558)

It is very easy to graphically determine how well a simple regression model fits the data by plotting a two dimensional representation of the predicted and actual dependent values against the independent variable (15:133). When more than one independent variable is used, the data is represented as a hypersurface in a k + 1 dimensional space where k is the number of independent variables (15:133). In this

case, the problem of graphically determining how well the model fits the data becomes more difficult.

In this research, the PACAF and USAFE data sets each had approximately 18 different types of aircraft flying hours and two military population variables. In order to reduce the number of variables into a manageable data set, an initial selection of aircraft flying hour variables was made based on whether the aircraft was a major weapon system (i.e. F-16, A-10, F-4, F-15 etc.) and/or the aircraft flew a significant percentage of the total flying hours (i.e. C-130, C-135, B-52, etc.). Once this initial selection was made, the computer program SAS (proc reg) was used to develop statistically significant first order multiple regression models at the 95% confidence level. If the residual analysis revealed any nonlinearities between the dependent and independent variables, higher order terms were added to improve the fit of the model. The first order general representation of the model is shown in the following equation;

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_i x_i + \epsilon$$

where: y = tonnage (airlift or sealift)

x_1 through x_i = quarterly flying hours variables by type of aircraft or officer/airman military population variables.

The models were statistically validated and evaluated (forecasting accuracy) using the same six steps used to validate and evaluate the DSXR simple regression models, but the F-test was substituted for the two tailed t-test and a multicollinearity test was added.

1. The analysis of variance F-test.

H_0: $\beta_1 = \beta_2 = \beta_3 = \ldots \beta_k = 0$

H_a: At least one β_i does not equal 0

Test Statistic: F

Rejection Region: $F > F_\alpha$

where,

$\alpha = .05$

k = the number of independent variables

n = the number of observations in the data set

$v_1 = k$

$v_2 = n - (k + 1)$

The global usefulness of the model was tested using the analysis of variance F test. The value of F must be greater than the value of F_α in order to reject H_o and accept H_a and conclude that the model is useful for predicting the dependent variable (18:578).

2. Multicollinearity. Before developing the models, the Pearson's coefficient of correlation (r) was computed for the independent variables for each data set so that highly correlated ($r > +.5$, $r < -.5$) could be identified (3:206-207). Once the models were developed, the variance inflation factors (VIF) were computed to determine whether a problem of multicollinearity existed between the variables.

H_o: Variables x_1, x_2, x_3 ... x_k are more closely related to the dependent variable than each other.

H_a: Variables x_1, x_2, x_3, ... x_k are more closely related to each other than the dependent variable.

Test Statistic:

VIF (x_1), VIF (x_2), VIF (x_3), . . . VIF (x_k)

Rejection Region: VIF $(x_i) > 1 / (1 - R^2)$

where,

i = one of the independent variables

k = the number of independent variables.

The VIF for each independent variable is subjected to the test and VIF's that exceed $1 / (1 - R^2)$ indicate the model has a problem with multicollinearity (8:80).

Neural Network Model Development and Forecasting Evaluation

The objective of the neural network models was the same as the multiple regression models, to forecast future PACAF and USAFE SDT requirements (measured in airlift and sealift tonnage). To solve this problem, two different techniques were used to develop the neural networks. One technique was based on developing the networks with data (independent and dependent variables) similar to the data (14 quarters) used by the multiple regression models (in order for the networks to process the data, the data were transformed so that it ranged between 0 and 1). An iterative trial and error approach was used to find the best network architecture (training algorithm and the number of inputs, outputs, and hidden layers) that could achieve a statistically significant level of learning convergence (learning convergence is the capacity of a network to correctly associate input values to a desired output value and for this research, network learning convergence was measured in terms of R^2 values). The networks (labeled as multivariable networks) were developed by presenting the 14 quarters of transformed independent and dependent variable data as training examples (PACAF and USAFE MSC FY 1985/1 to 1988/2, PACAF and USAFE MAC FY 1985/3 to 1988/4) and tonnage forecasts were computed by presenting new independent

variable data (6 quarters for PACAF and USAFE MSC (1988/3 to 1989/4), 5 quarters for PACAF and USAFE MAC (FY 1989/1 to 1990/1)).

In addition to the multivariable network development technique, a time series development technique was used for the PACAF and USAFE MAC data sets. This technique consisted of presenting twenty sets of four quarters (five years) of MAC tonnage at times t-3, t-2, t-1, and t as network input values and one forecast quarter at time t+1 as the network output value. Forecasting accuracy was evaluated by forecasting for the same five quarter period (FY 1989/1 to 1990/1).

The type of network used in this research was a multi-layer back-propagation network using the Neural Ware Professional II computer program (neural network simulation program). All the networks in this research used the generalized delta rule algorithm and the input and output layers used a linear transfer function while the two hidden layers used the sigmoid transfer function. Six steps were followed in the development of the networks:

1. Collect the data and accomplish data transformations.
2. Determine the input and output variables and construct the network.
3. Train the network with sample data.
4. Analyze the network for training inefficiencies and pathological conditions.
5. Validate the model.
6. Predict future values by presenting new data.

A large part of the research was devoted to experimenting with different arrangements of nodes, layers, interconnections, inputs, outputs, training algorithms, and weights for each network. After the networks were constructed, training began by presenting sample data to

the networks. Steps 2, 3, and 4 were iterative steps in order to find the best possible networks.

In step 4, the networks were analyzed to determine the effectiveness of the nodes and to identify inefficient training methods. Networks with nodes that increased without bound were terminated since no additional training could achieve learning convergence. Networks with other pathological conditions such as nodes with weights of zero, nodes that had the same weight or opposite weight of another node and nodes that continued to fire regardless of the input were also terminated (26:59-61).

In step 5, R^2 values were used to validate the network's pattern recognition capability. A network with a low R^2 value (below .5) indicated the network was not properly configured for the problem or the network did not receive enough training iterations. The networks in this research made rapid progress (large increases in the R^2 value) with 1000 to 2000 training iterations, but progress usually slowed down with increased training iterations. To prevent undertraining or overtraining, network training was terminated once the R^2 value became relatively stable. The Durbin-Watson statistic was also calculated for each network output.

Finally, in step 6, the networks were presented new input data to make forecasts. Like the DSXR models and the multiple regression models, forecasting accuracy was evaluated with the MAE, minimum absolute error, and maximum absolute error and was compared to the multiple regression and DSXR model forecasts.

Methodology Summary

This chapter presented the steps used to conduct this research. In order to develop new forecasting models, the research began with the collection of additional independent variable data. The total flying hour parameter used by DSXR for PACAF and USAFE general cargo tonnage was restructured into separate flying hour parameters for each type of aircraft. Other independent variables (officer and airman population) were also added to the model.

The DSXR simple regression models were replicated using the SAS computer program and validated with five statistical diagnostic tests. Forecasting accuracy was evaluated by measuring the MAE and minimum and maximum absolute errors. Multiple regression models were developed using the SAS computer program and were subjected to similar diagnostic tests and forecasting accuracy evaluation as the DSXR models.

Neural network models were developed using the Neural Ware Professional II computer program. Multivariable and time series networks were constructed and evaluated for pattern recognition capability and forecasting accuracy.

Finally, the DSXR model, the multiple regression model, and the neural network model were compared to determine the best forecasting model.

IV. MSC SDT Forecasting Results and Analysis

This chapter is divided into four parts. The first part examines the current DSXR regression models used to forecast sealift tonnage requirements to PACAF and USAFE. The second part is an analysis of the PACAF and USAFE data sets. Part three presents the development and results of the multiple regression models and part four is the development and results of the multivariable neural network models.

DSXR Simple Regression Model Validation and Forecasting Evaluation

DSXR PACAF MSC Model. Appendix A is the complete SAS output of the DSXR simple regression model used to forecast for the six quarter period from fiscal year 1988/3 to 1989/4. The dependent variable is PACAF sealift tonnage and the independent variable is the total PACAF aircraft flying hours. This model was developed by using 34 quarters of data (1980/1 to 1988/2). Table 6 displays a portion of the SAS output. The following diagnostic output showed the following:

1. Two Tailed Test.

H_o: $\beta_1 = 0$

H_a: β_1 does not equal 0

Test Statistic: $t = 3.084$

Rejection Region: $t_{.025} < -2.042$, $t_{.025} > 2.042$

where,

$\alpha = .05$

$df = 32$.

The two tailed test indicates the model is significant at the .05 level of confidence (the flying hour (independent) variable contributes information for the prediction of the tonnage (dependent) variable).

Table 6

DSXR PACAF MSC Model Analysis of Variance

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB>F
MODEL	1	339797123.86	339797123.86	9.511	0.0042
ERROR	32	1143223160	35725723.76		
C TOTAL	33	1483020284			

ROOT MSE	5977.1	R-SQUARE	0.2291
DEP MEAN	41325.85	ADJ R-SQ	0.2050
C.V.	14.46334		

PARAMETER ESTIMATES

VARIABLE	DF	PARAMETER ESTIMATE	STANDARD ERROR	T FOR HO: PARAMETER=0	PROB > |T|
INTERCEP	1	-3382.12	14532.78074	-0.233	0.8175
FH	1	1.13325080	0.36745715	3.084	0.0042

2. R^2 Value. Although the two tailed test indicates the model is useful, the R^2 value is low (.2291) which signifies a lack of fit of the model to the data.

3. Residual Analysis. A plot of the residuals versus the predicted values (Appendix A, Figure 57) appears to show a problem with heteroscedasticity. The plot is funnel shaped with increasing residual variance as the predicted values increase. The plot of the residuals versus flying hours (Appendix A, Figure 58) does not appear to show a problem with the assumption of linearity between tonnage and flying hours (the residuals seem to be randomly distributed). The Wilk Shapiro Test for Normality was used to determine whether the residuals were normally distributed (see Appendix A).

H_0: The residual distribution function is a normal distribution function.

H_a: The residual distribution function is not a normal distribution function.

Test Statistic: $W = .97273$

Rejection Region: $W < W_{.05} = .933$

where,

$\alpha = .05$, $n = 34$.

The Wilk Shapiro Test indicates there is insufficient evidence to reject the null hypothesis and accept the alternative hypothesis based on a 95% confidence interval.

4. Outlier Detection. The plot of the studentized residual values versus flying hours (Appendix A, Figure 59) shows all the residuals falling within 3 standard deviations from the mean of zero and all but one residual (observation 23, student residual = 2.73) falling within 2 standard deviations. Based on this finding, no outliers are present in the data set.

5. Durbin Watson (DW) Test. This test was used to test for the existence of first order autocorrelation.

H_0: The residuals are not positively autocorrelated.

H_a: The residuals are positively autocorrelated.

Test Statistic: $DW = .655$

Rejection Region: $0 < DW < d_l$ (positive autocorrelation)

Acceptance Region: $d_u < DW < 2$ (no autocorrelation, positive test)

where,

$d_l = 1.39$

$d_u = 1.51$

$n = 34$

$k = 1$.

The null hypothesis is rejected and the alternative hypothesis is accepted with a 95% level of confidence based on the fact that the DW statistic (.655) is lower than d_l (1.39). The residuals are positively autocorrelated and the plot of the residuals versus N (Appendix A, Figure 60) (N = automatic observation counter that creates a sequential period indicator) shows how the residuals start out negative, become positive, and then become negative again (cyclical residual effect).

6. Forecasting Accuracy. The DSXR model was used to forecast for the six quarter period from fiscal year (FY) 1988/3 to 1989/4. Since the R^2 value of the model was low, a comparison was made between the DSXR model and a simple 34 quarter (1980/1 to 1988/2) and 12 quarter (1985/3 to 1988/2) tonnage average. The averages were used as forecasts for the six quarter forecasting period. Table 7 shows the results of the model forecasts and the 34 and 12 quarter average tonnage forecasts.

According to the forecasting results, the 12 quarter average is a slightly more accurate forecasting model compared to the DSXR model. This is not surprising for three reasons. First, the DSXR model is a statistically useful model, but the low R^2 value and a high MSE value indicate the flying hour variable does not explain a large amount of the tonnage variance. Second, the DSXR model was developed from a large data base (34 quarters of data (1980/1 to 1988/2)) containing old data as well as recent data. Despite the numerous weapon system changes that have taken place since 1980, one has to make the unreasonable assumption that the relationship between tonnage and flying hours remained constant for the entire 8 year time period and will continue to remain constant in the future. Third, the 12 quarter average was developed from a small data

base (1985/3 to 1988/2) and represents current sealift tonnage requirements.

The 34 quarter average did not forecast as well the other models, but it did achieve the minimum error for one forecast. This model had the same problem as the regression model, too much old data was used in developing the forecasts.

Table 7

DSXR PACAF MSC Model Forecasting Accuracy

FY/Qtr	Actual Tonnage	DSXR Model Forecasts	12 Qtr Average Forecasts	34 Qtr Average Forecasts
1988/3	55113	43902	43526	41326
1988/4	42250	40698	43526	41326
1989/1	43079	42714	43526	41326
1989/2	45086	44169	43526	41326
1989/3	44009	44614	43526	41326
1989/4	41224	38077	43526	41326
	MAE:	2966	2943	3835
	Minimum Error:	365	447	102
	Maximum Error:	11211	11587	13787

7. Summary of Analysis. Overall, the model was useful based on the results of the two tailed t-test with a 95% confidence level, but the low R^2 value indicated a need for model improvement. The residual analysis revealed a problem with heteroscedasticity and the Durbin-Watson test proved the tonnage data was positively autocorrelated (two violations of the probability distribution assumptions of ϵ) (18:500-501).

The simple 12 quarter average had comparable forecasting accuracy to the DSXR regression model. The forecasting capability of the DSXR regression model was degraded because the model was developed from a 34 quarter data base which contained old, irrelevant data. The relationship

between tonnage and flying hours does not remain constant during the 34 quarter period. Strom's research supports this finding. His research proved DSXR's iterative procedure for finding a regression model was invalid because the β coefficients statistically changed as the number of quarters were reduced in developing the model. This means the relationship between tonnage and flying hours changes with respect to time.

DSXR USAFE MSC Model. Appendix B contains the complete SAS output of the DSXR simple regression model used to forecast for the six quarter period from fiscal year 1988/3 to 1989/4. The dependent variable is USAFE sealift tonnage and the independent variable is the total USAFE aircraft flying hours. This model was also developed by using 34 quarters of data (1980/1 to 1988/2). Table 8 displays a portion of the SAS output. The following diagnostic output showed the following:

1. Two Tailed Test.

H_o: $\beta_1 = 0$

H_a: β_1 does not equal 0

Test Statistic: $t = 2.434$

Rejection Region: $t_{.025} < -2.042$, $t_{.025} > 2.042$

where,

$\alpha = .05$

$df = 32$.

The two tailed test indicates the model is significant at the .05 level of confidence (the flying hour (independent) variable contributes information for the prediction of the tonnage (dependent) variable).

Table 8

DSXR USAFE MSC Model Analysis of Variance

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB>F
MODEL	1	1200722483	1200722483	5.926	0.0207
ERROR	32	6484081098	202627534.30		
C TOTAL	33	7684803580			

ROOT MSE	14234.73	R-SQUARE	0.1562
DEP MEAN	70816.56	ADJ R-SQ	0.1299
C.V.	20.10085		

PARAMETER ESTIMATES

VARIABLE	DF	PARAMETER ESTIMATE	STANDARD ERROR	T FOR HO: PARAMETER=0	PROB > |T|
INTERCEP	1	5192.00148	27068.71949	0.192	0.8491
FH	1	0.89342350	0.36701624	2.434	0.0207

2. R^2 Value. Although the two tailed test indicates the model is useful, the R^2 value is low (.1562) which signifies a lack of fit of the model to the data.

3. Residual Analysis. The plot of the residuals versus the predicted values (Appendix B, Figure 61) does not appear to be randomly distributed (increasing variance) which indicates a heteroscedasticity problem. The plot of studentized residuals versus flying hours shows how most of the residuals are negative and fall within 0 and -1.3 standard deviations while the positive residuals fall within 0 and 2.1 standard deviations with a grouping of residuals near 2 standard deviations. The plot of the residuals versus flying hours (Appendix B, Figure 62) does not appear to be randomly distributed and looks similar to the plot of the residuals versus the predicted values. This means the assumption of linearity between tonnage and flying hours may not be correct (fitting a

straight line through curvilinear data). The Wilk Shapiro Test for Normality was used to determine whether the residuals were normally distributed (Appendix B).

H_o: The residual distribution function is a normal distribution function.

H_a: The residual distribution function is not a normal distribution function.

Test Statistic: $W = .907887$

Rejection Region: $W < W_{.05} = .933$

where,

$\alpha = .05$

$n = 34$.

The Wilk Shapiro Test proves the residual distribution function is not normal (reject the null hypothesis and accept the alternative hypothesis based on a 95% confidence interval). This confirms the interpretation of the residual plots.

4. Outlier Detection. The plot of the studentized values versus the flying hours (Appendix B, Figure 63) shows all the residuals falling within 3 standard deviations from the mean of zero and all but one residual (observation 21 (1985/1), student residual = 2.1062) falling within 2 standard deviations. This is similar to the DSXR PACAF MSC model which had one residual falling outside 2 standard deviations at the same time period (observation 23, (1985/3)). Based on this finding, no outliers are present in the data set.

5. Durbin Watson (DW) Test. This test was used to test for the existence of first order autocorrelation. The statistics are displayed in Appendix B.

H_o: The residuals are not positively autocorrelated.

H_a: The residuals are positively autocorrelated.

Test Statistic: DW = .644

Rejection Region: $0 < DW < d_l$ (positive autocorrelation)

Acceptance Region: $d_u < DW < 2$ (no autocorrelation, positive test)

where,

$d_l = 1.39$

$d_u = 1.51$

$n = 34$

$k = 1$.

The null hypothesis is rejected and the alternative hypothesis is accepted with a 95% level of confidence based on the fact that the DW statistic (.644) is lower than d_l (1.39). The residuals are positively autocorrelated and the plot of the residuals versus N (Appendix B, Figure 64) (N = automatic observation counter that creates a sequential period indicator) shows how the residuals start out negative, become positive, and then become negative again (cyclical effect).

6. Forecasting Accuracy. The model was used to forecast for the six quarter period from fiscal year (FY) 1988/3 to 1989/4. Since the R^2 value of the model was low, a comparison was made between the DSXR model and a simple 34 quarter (1980/1 to 1988/2) and 12 quarter (1985/3 to 1988/2) tonnage average. The averages were used as forecasts for the six quarter forecasting period. Table 9 shows the results of the DSXR USAFE MSC regression model forecasts and the 34 and 12 quarter average tonnage forecasts.

Like the DSXR PACAF MSC model, the 12 quarter average is a slightly more accurate forecasting model than the DSXR model. The forecasting

accuracy explanation for the DSXR PACAF MSC model also applies to this model.

Table 9

DSXR USAFE MSC Model Forecasting Accuracy

FY/Qtr	Actual Tonnage	DSXR Model Forecasts	12 Qtr Average Forecasts	34 Qtr Average Forecasts
1988/3	81479	71601	71714	70817
1988/4	76619	71977	71714	70817
1989/1	73847	66066	71714	70817
1989/2	63853	71427	71714	70817
1989/3	74806	81697	71714	70817
1989/4	88613	79357	71714	70817
	MAE:	7670	7443	8041
Minimum Error:		4641	2133	3030
Maximum Error:		9877	16899	17796

7. Summary of Analysis. The results of this model were very similar to the analysis results of the DSXR PACAF MSC model. Both models were useful based on the results of the two tailed t-test with a 95% confidence level, but the low R^2 value indicated a need for model improvement. The residual analysis for both models revealed a problem with heteroscedasticity and the Durbin-Watson test indicated the tonnage data was positively autocorrelated (two violations of the probability distribution assumptions of ϵ (18:500-501). Unlike the PACAF MSC model results, the Wilk-Shapiro test for normality statistically proved the residuals were not normally distributed. The results of the forecasting evaluation between the DSXR USAFE MSC model and the 12 and 34 quarter averages were similar to the results of the DSXR PACAF MSC model.

Data Analysis

Plots. Figure 10 is a plot of PACAF MSC tonnage versus PACAF flying hours. The relationship between tonnage and flying hours appears to be fairly linear when flying hours are below 40,000 hours. Above 40,000 flying hours, the data points are widely dispersed and the linear relationship is no longer present. Out of the 42 flying hour data points that were plotted, 26 of them are below 40,000 hours. Twenty-four of the twenty-six sub-40,000 flying hour data points occurred from 1978/1 to 1983/4. After 1983, the linear relationship between flying hours and tonnage is no longer apparent.

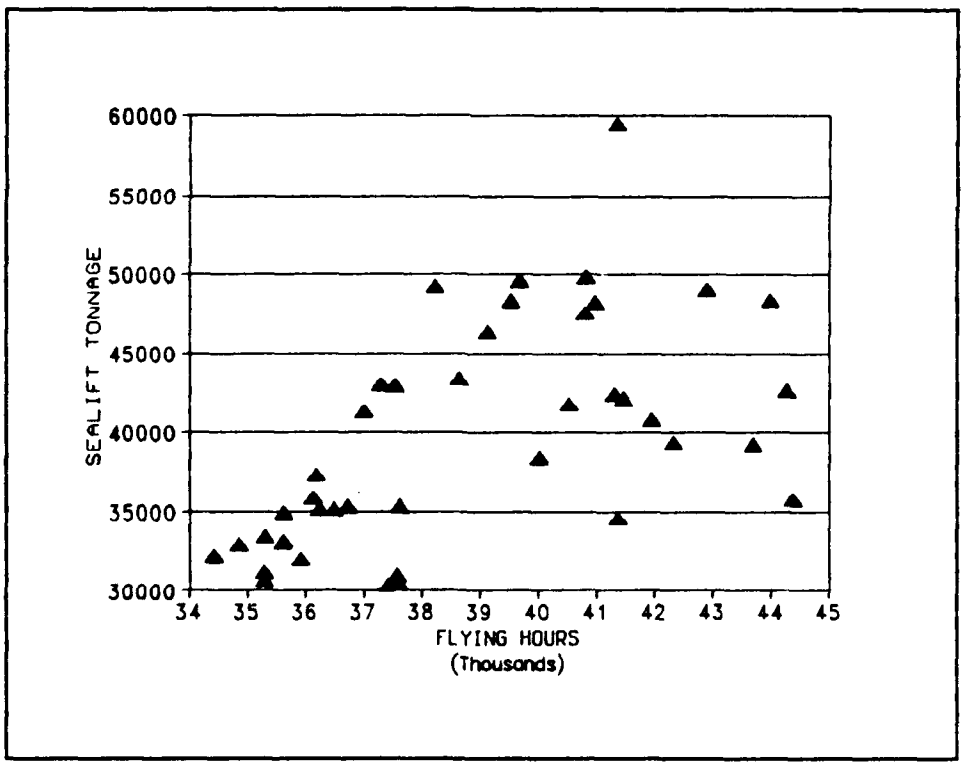

Figure 10. PACAF MSC Tonnage versus PACAF Flying Hours

The plot of the PACAF quarterly sealift tonnage (Figure 11) reveals how sealift requirements have changed with time. Two large peaks are evident in fiscal years 1983 and 1985.

Figure 12 is a plot of USAFE MSC tonnage versus USAFE flying hours. The relationship between USAFE tonnage and flying hours is similar to the relationship between PACAF tonnage and flying hours. In this case, the relationship is fairly linear when flying hours are below 65,000 hours. Above 65,000 flying hours, the data points are widely dispersed and the linear relationship is no longer present.

The quarterly USAFE MSC tonnage plot (Figure 13) shows similar peaks to the quarterly PACAF MSC tonnage plot (Figure 11). Unlike the PACAF tonnage, the USAFE tonnage has one continuous peak from 1983 to 1985 instead of two peaks.

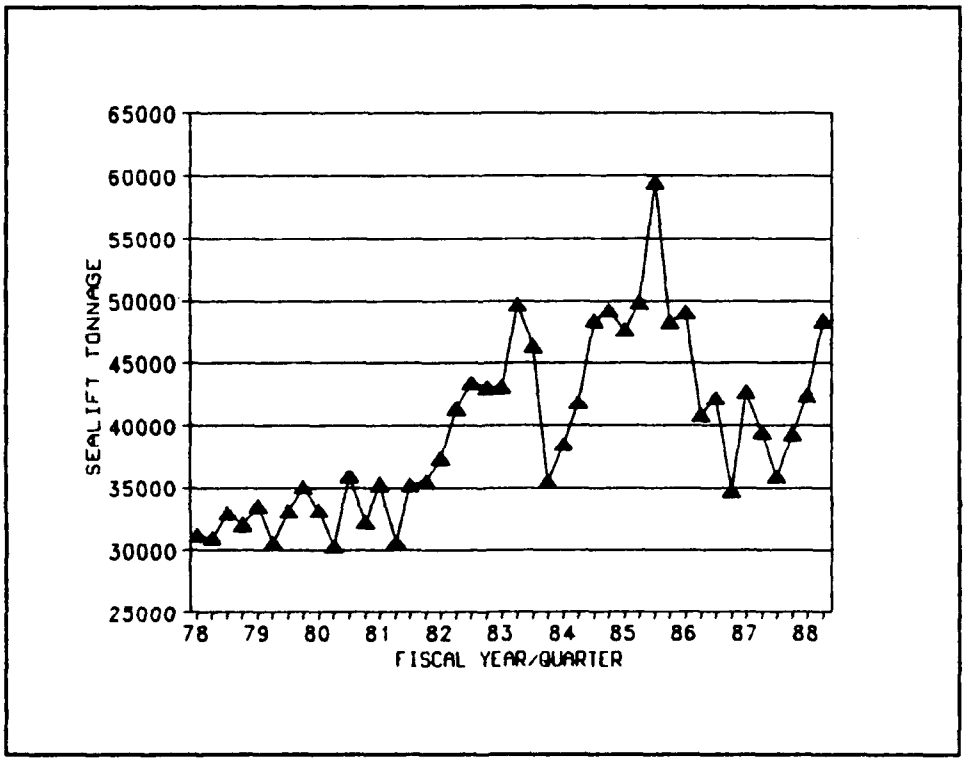

Figure 11. Quarterly PACAF MSC Tonnage

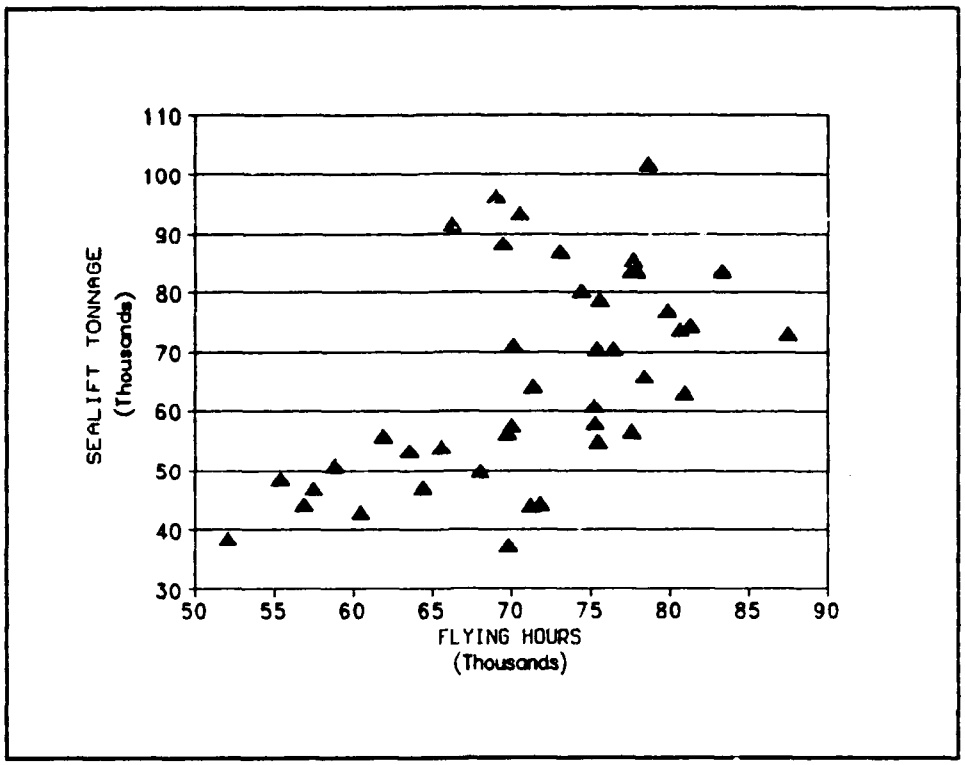

Figure 12. USAFE MSC Tonnage versus USAFE Flying Hours

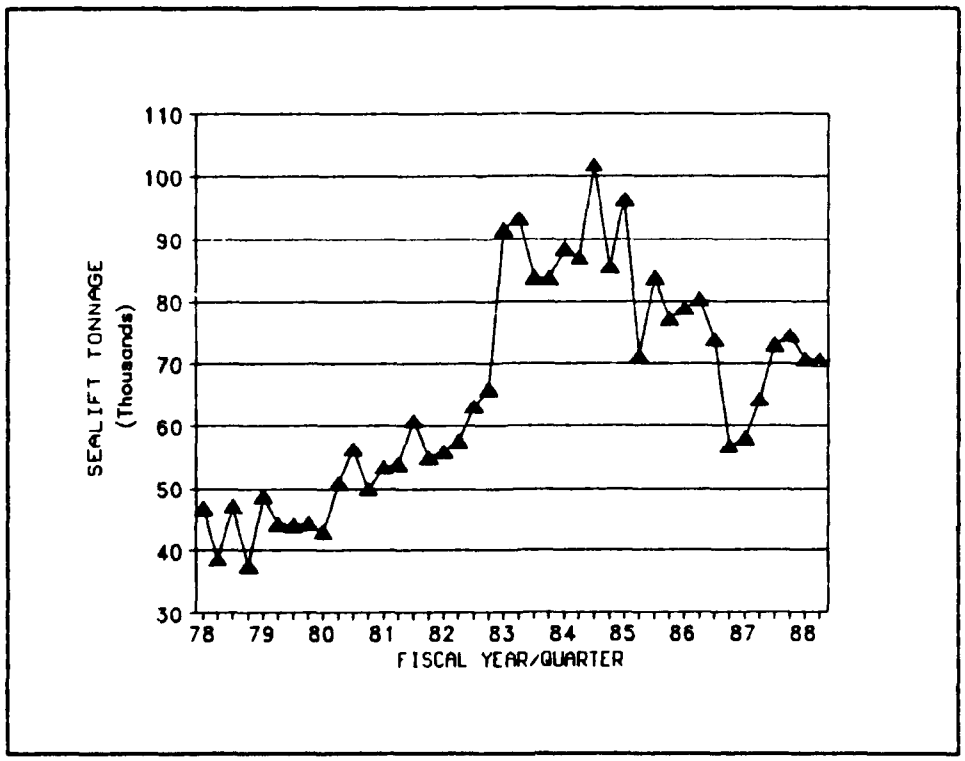

Figure 13. Quarterly USAFE MSC Tonnage

Interviews conducted with the DSXR personnel indicated the tonnage peaks for PACAF and USAFE could be the result of high growth and spending periods for the Air Force (19). Figures 14 and 15 are plots of tonnage and flying hours for PACAF and USAFE. Both plots show an increasing trend in flying hours and tonnage requirements, but the flying hour data do not exhibit the same large peaks that are evident with the tonnage data. It appears that factors other than flying hours have caused the large historical increases in sealift tonnage requirements.

Two other factors that could have caused the PACAF and USAFE sealift tonnage peaks are the aircraft funding levels and the overseas military populations. Figure 16 is a bar chart of the 3010 procurement dollars for aircraft. The chart shows large procurement dollar increases occurring from 1982 to 1985.

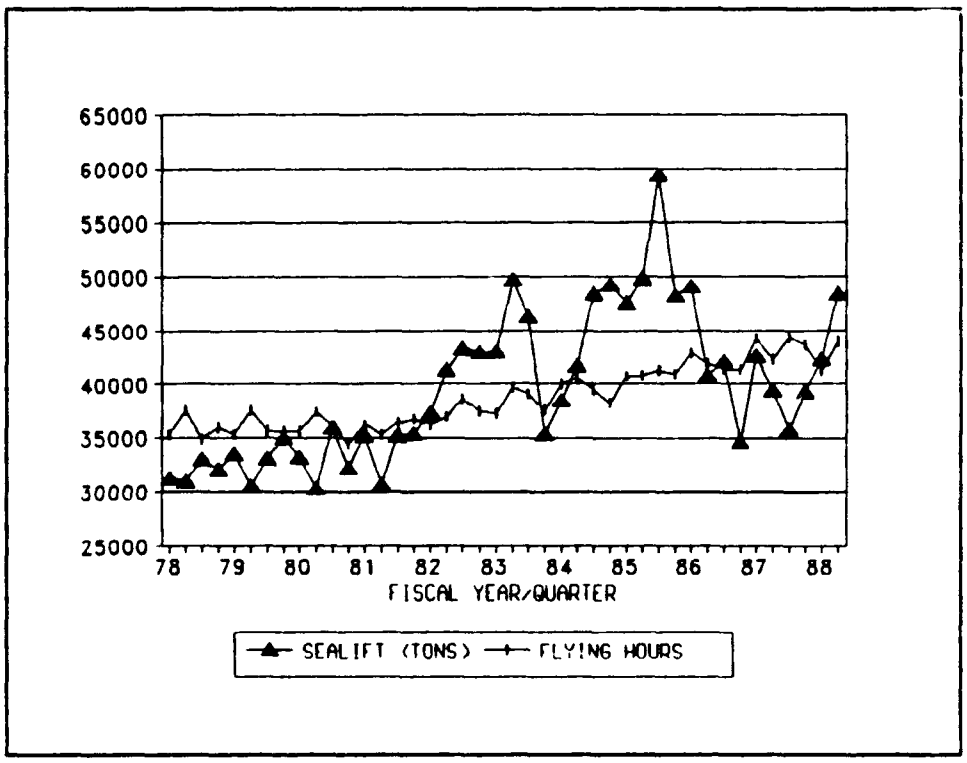

Figure 14. Quarterly PACAF MSC Tonnage and Flying Hours

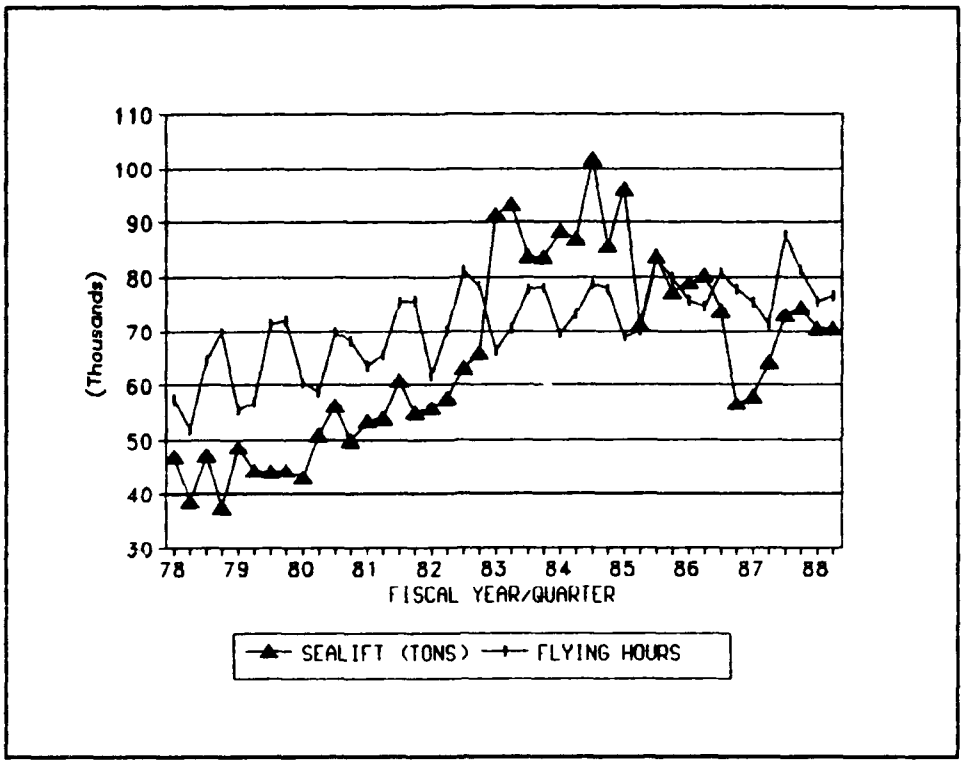

Figure 15. Quarterly USAFE MSC Tonnage and Flying Hours

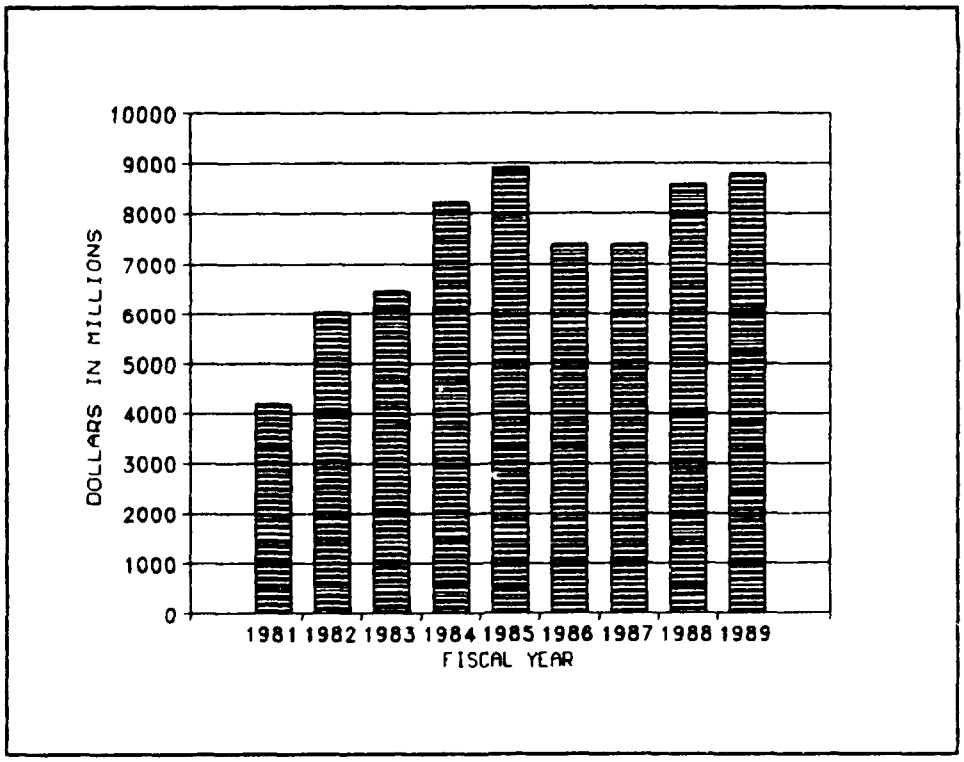

Figure 16. Aircraft Procurement Dollars

Figure 17 is a plot of PACAF officer and airman populations which shows manpower increases from 1982 to 1986. Figure 18 is a plot of USAFE officer and airman populations which shows officer manpower increases from 1982 to 1985 and an airman manpower increase in 1984.

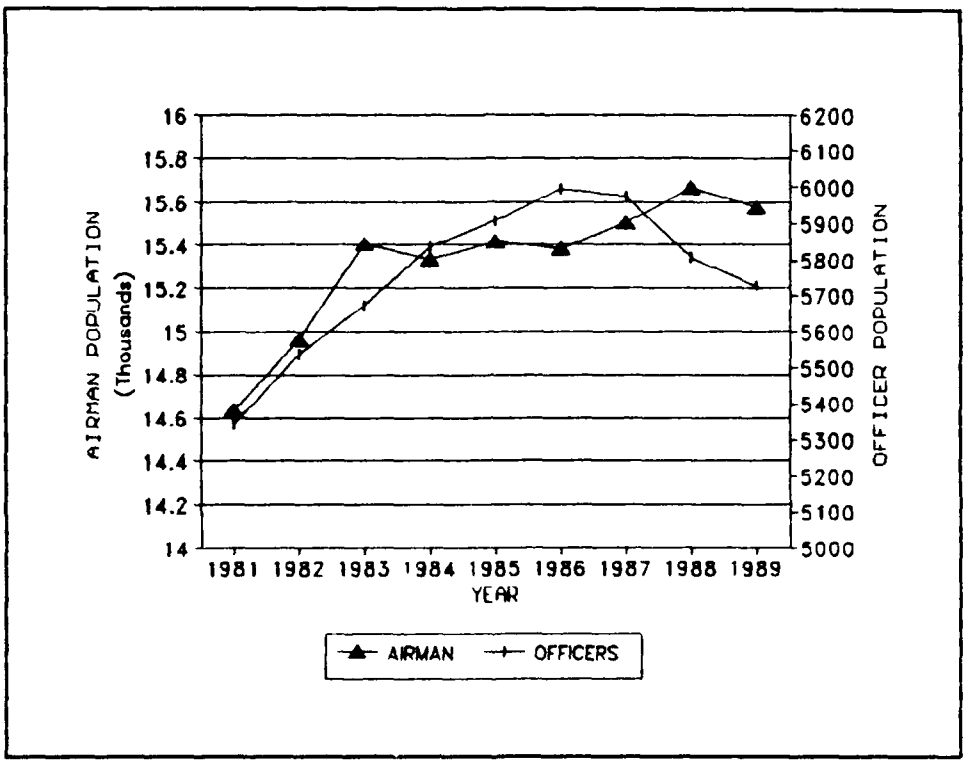

Figure 17. PACAF Military Population

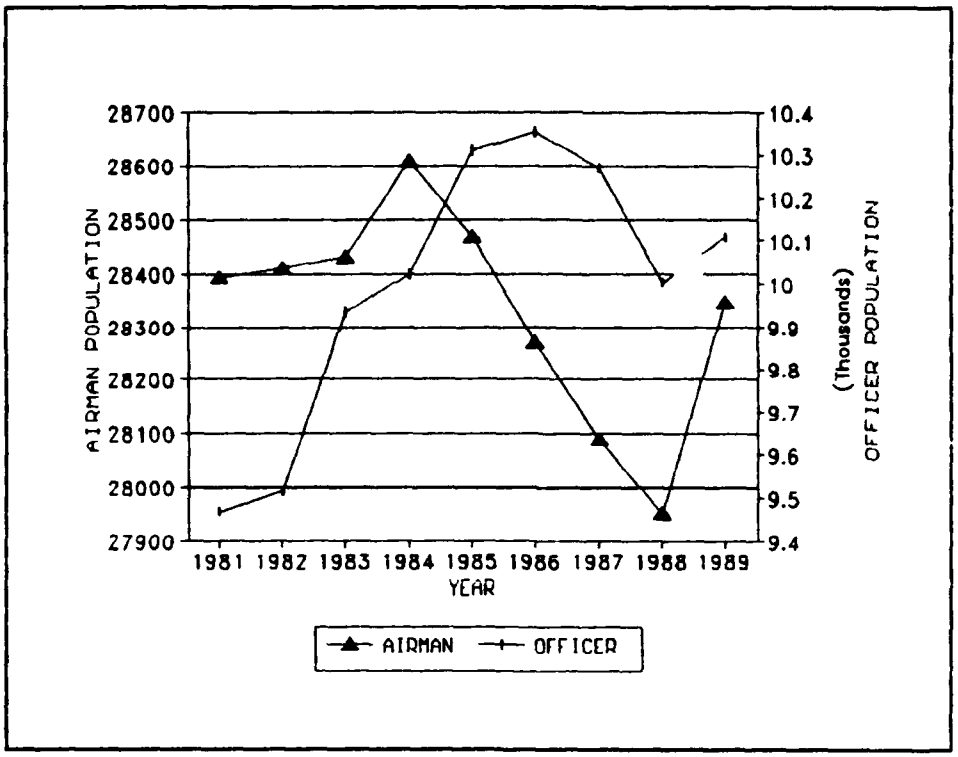

Figure 18. USAFE Military Population

Trend and Seasonal Analysis. PACAF and USAFE MSC tonnage and flying hours were analyzed using Gardner's trend and seasonal analysis methodology (Appendix C). According to Gardner's methodology, the variance of the actual data is compared to the variance of the difference between same quarters for each year (DBQ), the first difference between quarters (DBD-1), and the second difference between each quarter (DBD-2) (9:44-50). The data with the lowest variance indicates whether trend and/or seasonal characteristics exist. The MSC results are summarized in Table 10.

PACAF and USAFE MSC tonnage exhibited seasonality and a moderate trend since the variance of the first difference between quarters (DBD-1) was the lowest for each data set. PACAF and USAFE flying hours exhibited seasonality (the variance of the difference between same

quarters for each year (DBQ) was lowest), but the slight trend that was evident in the plots (Figures 14 and 15) was not detected.

Table 10

MSC Trend and Seasonal Analysis

PACAF MSC TONNAGE

	Actual	DBQ	DBQ-1	DBQ-2
Variance:	49027789	51731495	44276024	1.02E+08
Index:	100%	106%	90%	209%
Trend:			Moderate	
Seasonal:			Yes	

USAFE MSC TONNAGE

	Actual	DBQ	DBQ-1	DBQ-2
Variance:	2.84E+08	1.69E+08	1.45E+08	3.56E+08
Index:	100%	60%	51%	126%
Trend:			Moderate	
Seasonal:			Yes	

Pattern Identification. Appendix D is the SAS output of the autocorrelation analysis for the PACAF MSC data set (FY 1978/1 to 1988/2). The Q-statistic indicates this series is not white noise (the data set has autoregressive (AR) or moving average (MA) aspects) since the value (Q = 75.17) is greater than the chi square value (X^2 = 12.5916 with 6 df and 95% confidence level). The autocorrelations remain positive and significantly different from zero to r_{12} which means the series is not stationary. The autocorrelation function and partial autocorrelation function have large spikes at lag one (r_1 = .74251) which would indicate an AR(1) and/or MA(1) aspect. The autocorrelation function also shows a significant spike at lag two (r_2 = .62825) signifying a possible MA(2) aspect.

Appendix D also displays the SAS autocorrelation analysis output for the USAFE MSC data set (FY 1978/1 to 1988/2). The results of this

analysis are similar to the results for the PACAF MSC data set. The Q-statistic indicates this series is not white noise (Q = 129.46 is greater than the chi square value (X^2 = 12.5916 with 6 df and 95% confidence level). The series is not stationary since the autocorrelations remain positive and significantly different from zero to r_9. The autocorrelation function and partial autocorrelation function have large spikes at lag one (r_1 = .85238) and lag two (r_2 = .79221) which would indicate an AR(1 or 2) and/or MA(1 or 2) aspect.

Business Cycle Analysis. Gardner's business cycle pressure analysis methodology was used to determine the peaks, troughs, and turning points in the sealift tonnage (Appendix E). Pressures are values that show how the sealift tonnage for a quarter compares with the sealift tonnage for the same quarter a year earlier (10:40-42). The pressure values are ratios of the comparisons and are converted into percentages. A value above 100% means the present sealift tonnage is greater than the past year's sealift tonnage for the same period (a value below 100% means the present sealift tonnage is less than the past year's sealift tonnage for the same period).

The 1/4 pressure plot (Figure 19) shows the comparison of sealift tonnage with the same quarter for the previous year while the 4/4 pressure plot (Figure 20) shows the comparison of the sum of four quarters of sealift tonnage with the same sum for the previous year. Both plots show the peaks and troughs, but the turning points are more evident in the 4/4 pressure plot. Since 1987, sealift tonnage has been increasing and indicates the formation of another peak.

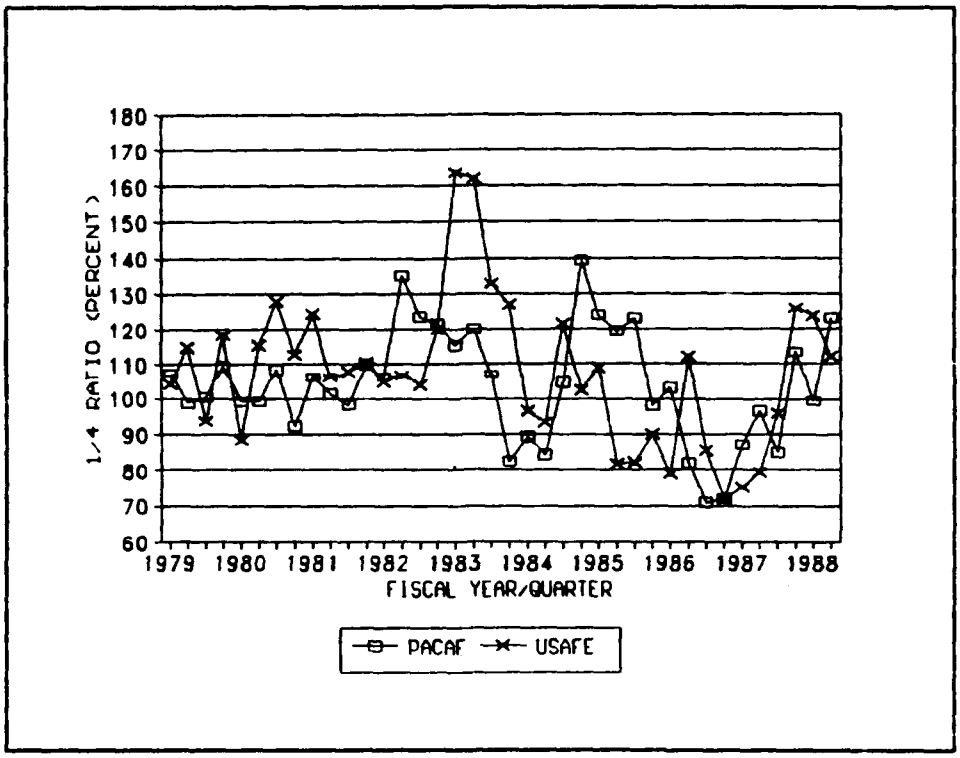

Figure 19. PACAF and USAFE MSC 1/4 Pressures

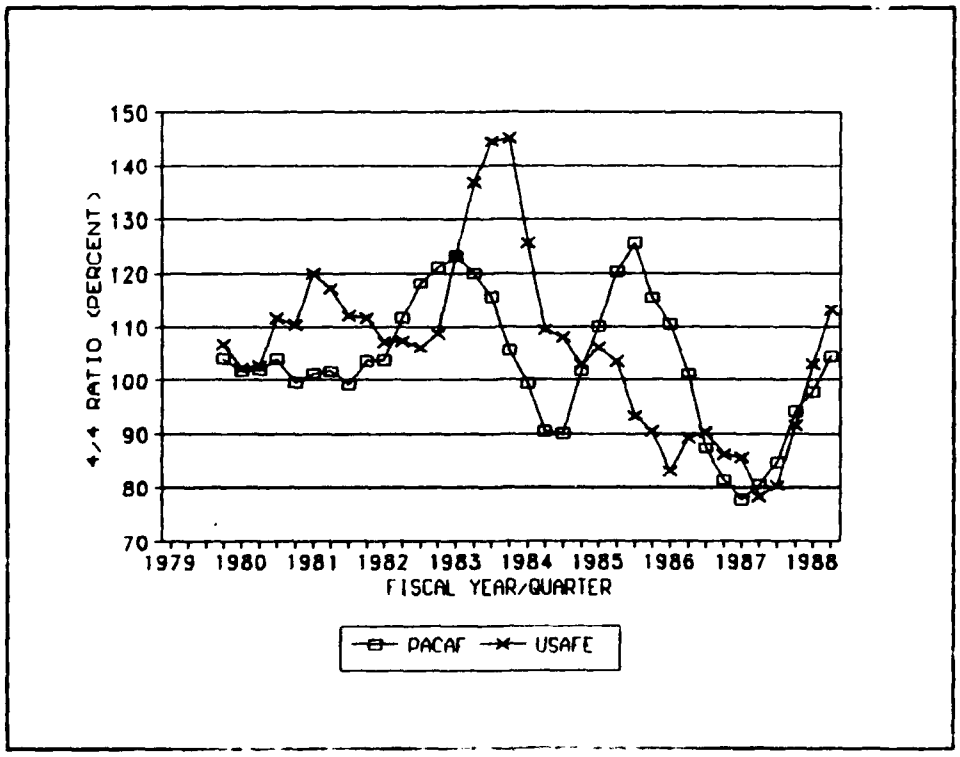

Figure 20. PACAF and USAFE MSC 4/4 Pressures

Figure 21 displays a plot of the ratio of sealift to flying hours for USAFE and PACAF. The plot shows how the PACAF and USAFE sealift tonnage per flying hour ratios have changed over an eleven year time period. The ratios are fairly similar and are consistent for particular time periods. From 1978 to 1982, the ratio is below 1.0 and averages approximately .9. During the peak sealift periods (1982 to 1985), the ratio increased above 1.0. It appears PACAF has historically required (on the average) more sealift tonnage per flying hour than USAFE.

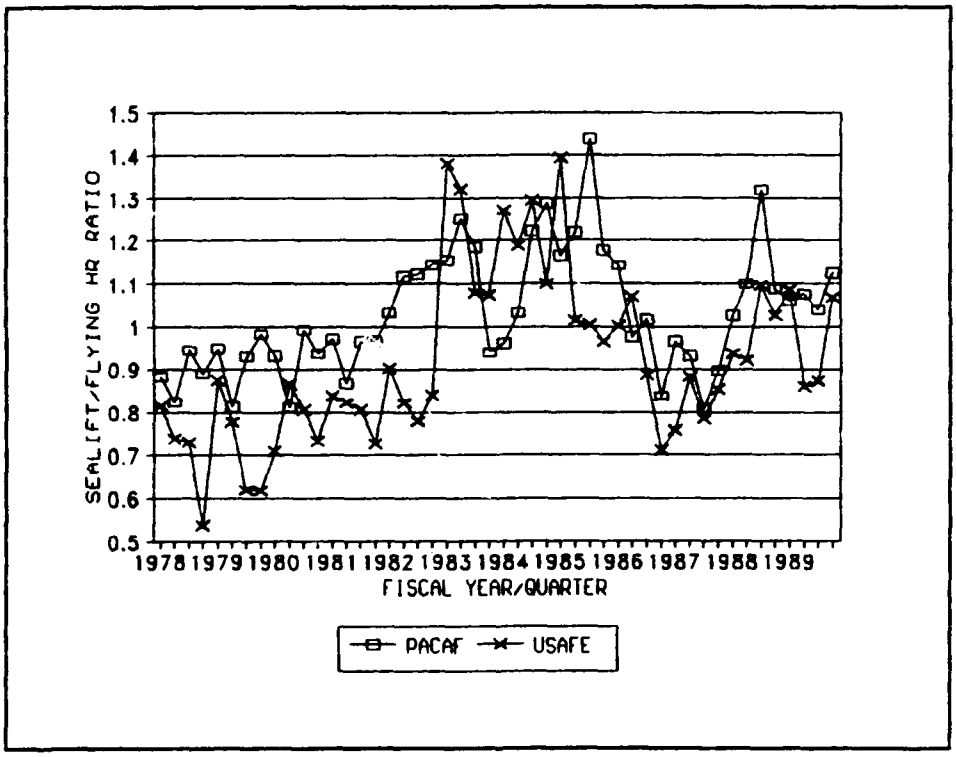

Figure 21. PACAF and USAFE Sealift/Flying Hour Ratios

Multiple Regression Models

The objective of the multivariable regression models was to determine whether the breakout of the total flying hour variable into specific aircraft types and the addition of military population variables

(Appendix F) contributed to increasing PACAF and USAFE sealift forecasting accuracy.

Figure 22 is a plot of PACAF flying hours versus fiscal year/quarter for four major types of aircraft (A-10, F-4, F-16, and F-15). This plot shows how the flying hours changed with respect to time. In FY 1982, the F-16 was a new weapon system and by FY 1985 the F-16 inventory was 59 aircraft which flew approximately 4500 total flying hours each quarter. By FY 1989, the F-16 inventory increased to 128 and the aircraft flew approximately 10,000 hours each quarter. Contrary to the F-16, the F-4 is a weapon system that is reaching the end of its life cycle. In FY 1981, the F-4 inventory was 112 aircraft and total flying hours were approximately 6,000 hours each quarter. By the end of FY 1989, the inventory decreased to 70 and the aircraft flew approximately 4,000 hours. Other aircraft such as the T-39, A-37, T-33, OV-10, and the E-3 have been introduced or phased out at different time periods from FY 1982 to FY 1989.

Figure 23 is a plot of USAFE flying hours versus fiscal year/quarter for the same four major types of aircraft (A-10, F-4, F-16, and F-15). Like the PACAF flying hours, this plot shows how the USAFE flying hours changed with respect to time. In the beginning of FY 1982, there were no F-16's, but by the end of FY 1989 the F-16 inventory was 242 and the aircraft flew over 16,000 total hours each quarter. In FY 1982, the F-4 inventory was over 200 and the aircraft flew over 12,000 hours each quarter. By the end of FY 1989, the inventory decreased to 52 and the aircraft flew approximately 4,000 hours. Like the PACAF flying hours, other aircraft such as the C-140, C-20, and the F-5 have flown at different time periods from FY 1982 to FY 1989.

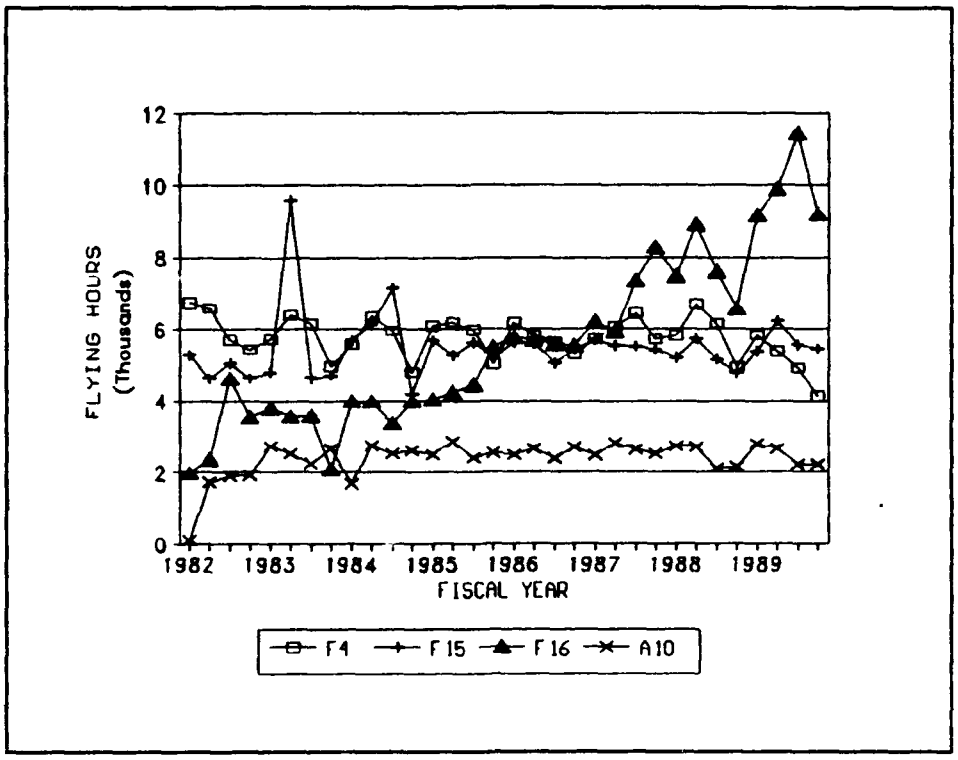

Figure 22. PACAF Flying Hours by Aircraft Type

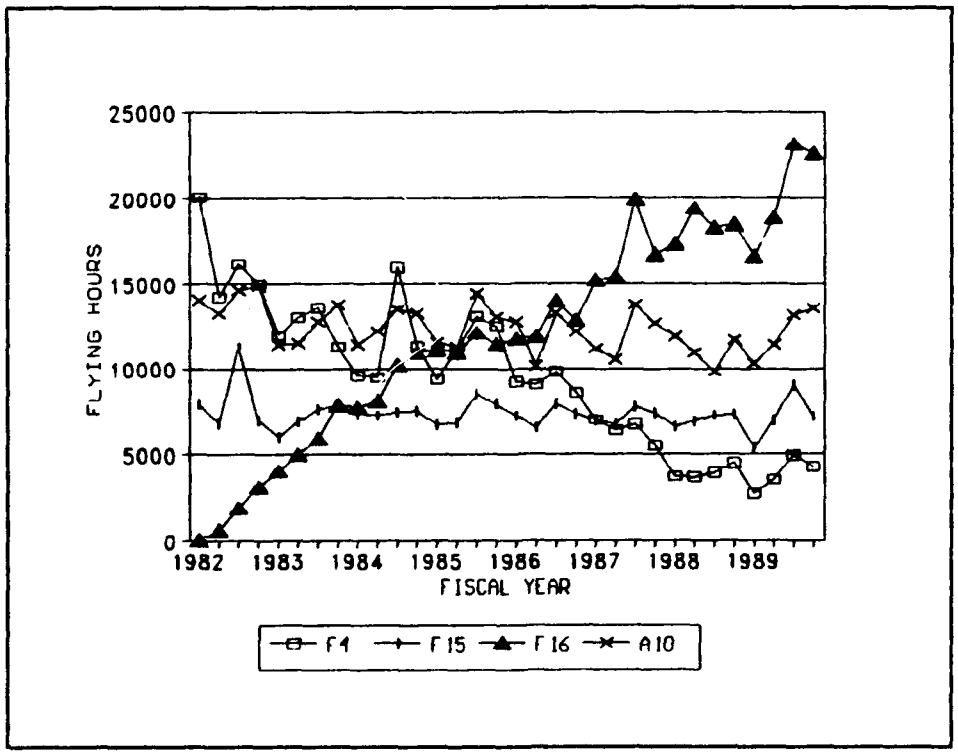

Figure 23. USAFE Flying Hours by Aircraft Type

Data. In order to model the current relationship between flying hours, military populations and sealift tonnage, fourteen quarters (three and a half years) of data (quarter 1, FY 1985 to quarter 2, FY 1988) were used to develop the regression models. Similar to the DSXR models, six quarters (quarter 3, FY 1988, quarter 4, FY 1989) were withheld to measure forecasting accuracy. The R^2, F-test, and Durbin Watson values were used to assess the fit of the models and the mean absolute error (MAE), minimum absolute error, and maximum absolute error were used to measure their forecasting accuracy.

PACAF MSC Multiple Regression Model Development. Eighteen different types of aircraft defined by MD (Mission Design) comprise the PACAF flying hour program (Figure 24). Appendix G shows the quarterly flying hours and inventory for each aircraft from FY 1985 to FY 1989. Seven aircraft types were selected as independent variables and used to develop the multiple regression model. The seven aircraft types account for approximately 85% of the total PACAF flying hours (Figure 24). Airman (AMN) and officer (OFF) military population independent variables were also selected and used to develop the multiple regression model (Figure 24).

Out of the nine initially selected independent variables (seven aircraft, two military population variables), six were eliminated because of multicollinearity problems or statistically non-significant t-values (Figure 24). The resulting three variable model was statistically significant at the 95% confidence level (Appendix H). The three variable model was;

$$y = 385572 + 13.82x_1 - 1.94x_2 - 61.52x_3$$

where: y = quarterly sealift tonnage

x_1 = quarterly A-10 flying hours

x_2 = quarterly F-16 flying hours

x_3 = quarterly officer population.

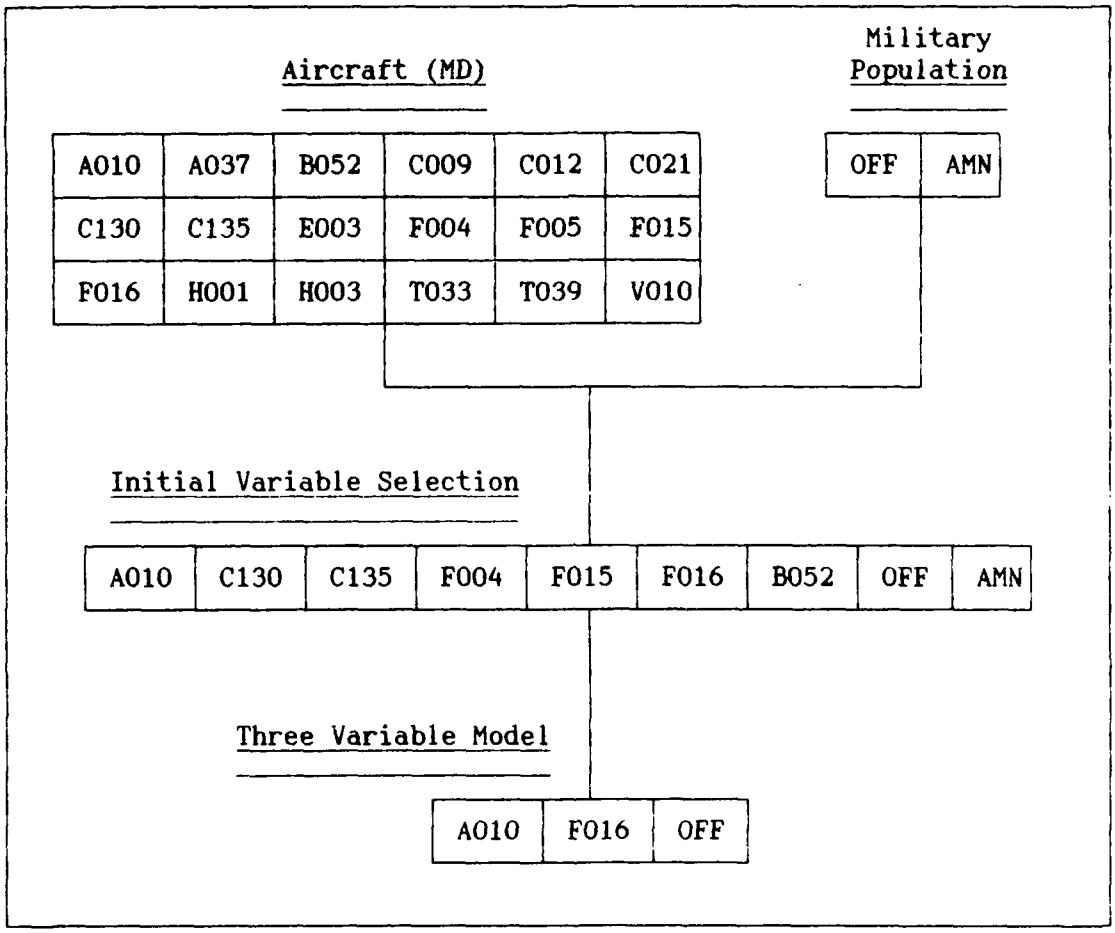

Figure 24. PACAF MSC Independent Variable Selection

Model Validation and Forecasting Evaluation. Appendix H is the complete SAS output of the multiple regression (three variable) model. Table 11 displays a portion of the SAS output. The following diagnostic output showed the following:

1. The analysis of variance F-test.

H_0: $\beta_1 = \beta_2 = \beta_3 = 0$

H_a: At least one β_i does not equal 0

Test Statistic: $F = 6.153$

Rejection Region: $F > F_{.05} = 3.71$

where,

$\alpha = .05$

$v_1 = 3$

$v_2 = 10$

The F test proves the model is significant at the .05 level of confidence (the independent variables contribute information for the prediction of the tonnage (dependent) variable).

2. R^2 Value. The model has a moderately good fit with a R^2 value higher than the DSXR model (.65 compared to .23 for the DSXR model).

Table 11

PACAF MSC Multiple Regression Model Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	3	364567520.74	121522506.91	6.153	0.0122
Error	10	197486140.19	19748614.019		
C Total	13	562053660.93			

Root MSE	4443.94127	R-square	0.6486
Dep Mean	44264.92857	Adj R-sq	0.5432
C.V.	10.03942		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	385572	125758.88080	3.066	0.0119
A10	1	13.819258	9.02533356	1.531	0.1567
OFF	1	-61.515186	20.19108244	-3.047	0.0123
F16	1	-1.944066	0.86970203	-2.235	0.0494

3. Residual Analysis. The plot of the residuals versus the predicted values (Appendix H, Figure 65) appears to be randomly

distributed and shows no heteroscedasticity problem. The plots of the residuals versus each of the independent variables (A-10 and F-16) (Appendix H, Figures 66 and 67) also appear to be randomly distributed and show no problem with the assumption of linearity between tonnage and each of the aircraft independent variables. The plot of residuals versus officer manpower independent variable (Appendix H, Figures 68) shows a slight curvature which indicates the possible need for a second order term. The Wilk Shapiro Test for Normality was used to determine whether the residuals were normally distributed.

H_0: The residual distribution function is a normal distribution function.

H_a: The residual distribution function is not a normal distribution function.

Test Statistic: $W = .940673$

Rejection Region: $W < W_{.05} = .874$

where,

$\alpha = .05$, $n = 14$.

The Wilk Shapiro Test indicates there is insufficient evidence to reject the null hypothesis and accept the alternative hypothesis based on a 95% confidence interval.

4. Multicollinearity. This model contains two variables (F-16 and OFF) which are negatively correlated ($r = -.61$) (Appendix I). The variance inflation factors were computed to determine whether a problem of multicollinearity existed between the variables.

H_0: Variables x_1, x_2, and x_3 are more closely related to the dependent variable than each other.

H_a: Variables x_1, x_2, and x_3 are more closely related to each other than the dependent variable.

Test Statistic: VIF = 1.09 (x_1)

VIF = 1.04 (x_2)

VIF = 1.07 (x_3)

Rejection Region: VIF > $1 / (1 - R^2) = 2.86$.

The results prove the independent variables are more closely related to the dependent variable than to each other.

5. Outlier Detection. The output of the studentized residuals (Appendix H) shows all the residuals falling within 2 standard deviations except for two residuals (observation 1 and 3) falling at approximately 2 standard deviations. Based on this finding, no outliers are present in the data set.

6. Durbin Watson (DW) Test. This test was used to test for the existence of first order autocorrelation.

H_0: The residuals are not positively autocorrelated.

H_a: The residuals are positively autocorrelated.

Test Statistic: DW = 1.386

Rejection Region: $0 < DW < d_l$ (positive autocorrelation)

Acceptance Region: $d_u < DW < 2$ (no autocorrelation, positive test)

where,

$d_l = .82$

$d_u = 1.75$

$k = 3$

$n = 14$.

There is insufficient evidence to reject the null hypothesis and accept the alternative hypothesis based on a 95% level of confidence.

The plot of the residuals versus N (Appendix B, Figure 69) (N = automatic observation counter that creates a sequential period indicator) appears to be randomly distributed.

7. Forecasting Accuracy. The multiple regression model was used to forecast for the six quarter period from fiscal year (FY) 1988/3 to 1989/4. Table 12 shows the results of the multiple regression model forecasts compared to the DSXR PACAF MSC model and the 12 quarter average tonnage forecasts.

The multiple regression model overestimated on four of the six forecasts, but achieved the lowest maximum absolute error compared to the other models. Figure 25 is a plot of the DSXR forecasts compared to the multiple regression forecasts which shows the overestimations by the multiple regression model in the last three quarters.

Table 12

PACAF MSC Multiple Regression Model Forecasting Accuracy

FY/Qtr	Actual Tonnage	DSXR Model Forecasts	12 Qtr Average Forecasts	Multiple Regression Forecasts
1988/3	55113	43902	43526	50977
1988/4	42250	40698	43526	44943
1989/1	43079	42714	43526	40265
1989/2	45086	44169	43526	52728
1989/3	44009	44614	43526	47772
1989/4	41224	38077	43526	46229
	MAE:	2966	2943	4342
	Minimum Error:	365	447	2693
	Maximum Error:	11211	11587	7642
	Rsquare:	.23	na	.65

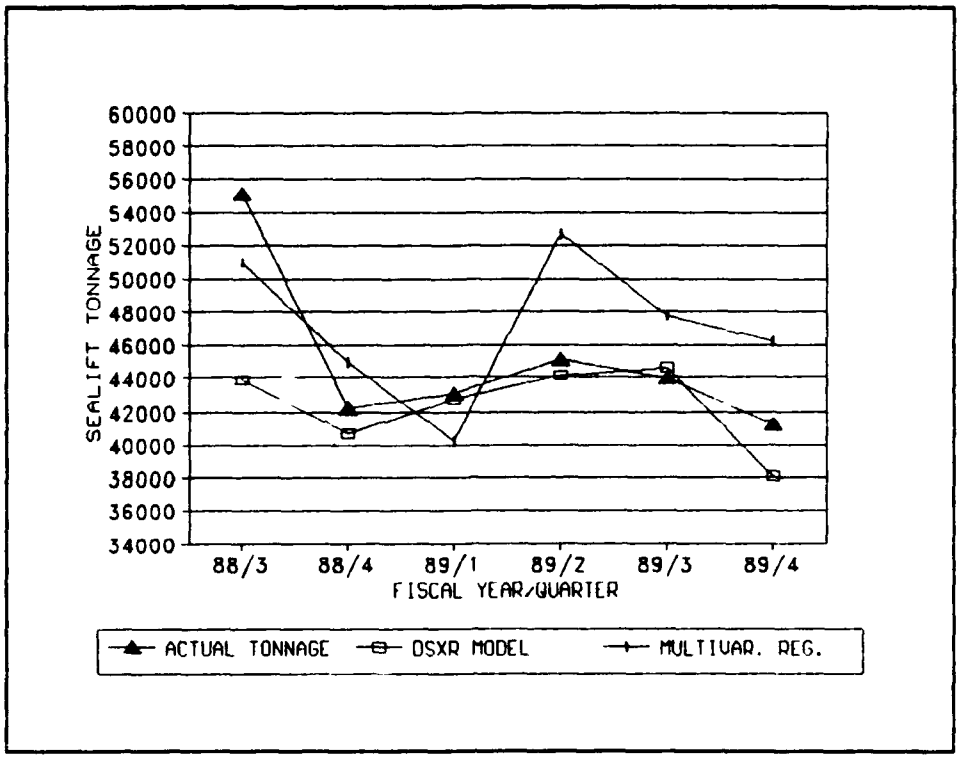

Figure 25. DSXR and Multiple Regression PACAF MSC Forecasts

8. Summary of Analysis. The multiple regression model was developed from nine independent variables, but the resulting model only had three independent variables. Although the model only had three variables, the three variables were highly positively correlated ($r > +.5$) and negatively correlated ($r < -.5$) with several of the other variables that were not included in the model (Table 13) (Appendix I). Overall, the model was a useful based on the results of the F-test (95% confidence level) and the residual analysis revealed no problems with heteroscedasticity, multicollinearity and autocorrelation.

The relationship between tonnage and officer manpower may not be linear based on the results of the residual analysis. A second order (squared) officer manpower variable (SOFF) was added to the model to

Table 13

PACAF MSC Multiple Regression Model Independent Variable Correlations

Independent Variable	Positively Correlated Variables	Negatively Correlated Variables
A-10	F-15, F-4	(none)
F-16	AMN	C-130, C-135
OFF	C-130	AMN

improve fit, but the F-16 variable was dropped because the variable was not statistically significant. The resulting model (Appendix H) had an improved fit (R^2 = .70) (Appendix H, Figure 70) compared to the first order model, but it did not have improved forecasting capability (MAE = 8761). The resulting second order model was;

$$y = -32409859 + 12.18210x_1 + 11035x_3 - .938747x^2_3$$

where: y = quarterly sealift tonnage

x_1 = quarterly A-10 flying hours

x_3 = quarterly officer population.

According to the forecasting results, the simple 12 quarter tonnage average and the DSXR regression model forecasted with greater accuracy than the multiple regression model. The multiple regression model suffered from an extrapolation problem because the officer manpower variable decreased to 5727 (79 personnel) in 1989 which represented the lowest level in 20 quarters. The forecasts for FY 1989/2 through 1989/4 were extrapolations since the officer manpower data for FY 1989/2 through 1989/4 was not contained in the sample data set that was used to develop the model. The F-16 variable data for FY 1989/2 was also not contained in the sample data set. This explains the relatively large

overestimations in FY 1989/2 through 1989/4. The extrapolation problem was amplified with the second order model because there were two officer variables contained in the model.

USAFE MSC Multiple Regression Model Development. Eighteen different types of aircraft defined by MD (Mission Design) comprise the USAFE flying hour program (Figure 26). Appendix J shows the quarterly flying hours and inventory for each aircraft from FY 1985 to FY 1989. Like the PACAF model, seven aircraft types and the airman (AMN) and officer (OFF) military population data (Appendix F) were initially selected as independent variables and used to develop the multiple regression model. The seven aircraft types account for approximately 87% of the total USAFE flying hours (Figure 26).

Out of the nine initially selected independent variables (seven aircraft, two military population variables), six were eliminated because of multicollinearity problems or statistically non-significant t-values (Figure 26). The resulting three variable model was statistically significant at the 95% confidence level (Appendix K). The three variable model was;

$$y = 630601 + 4.46x_1 - 59.98x_2 + 2.49x_3$$

where: y = quarterly sealift tonnage

x_1 = quarterly C-130 flying hours

x_2 = quarterly officer population

x_3 = quarterly F-4 flying hours.

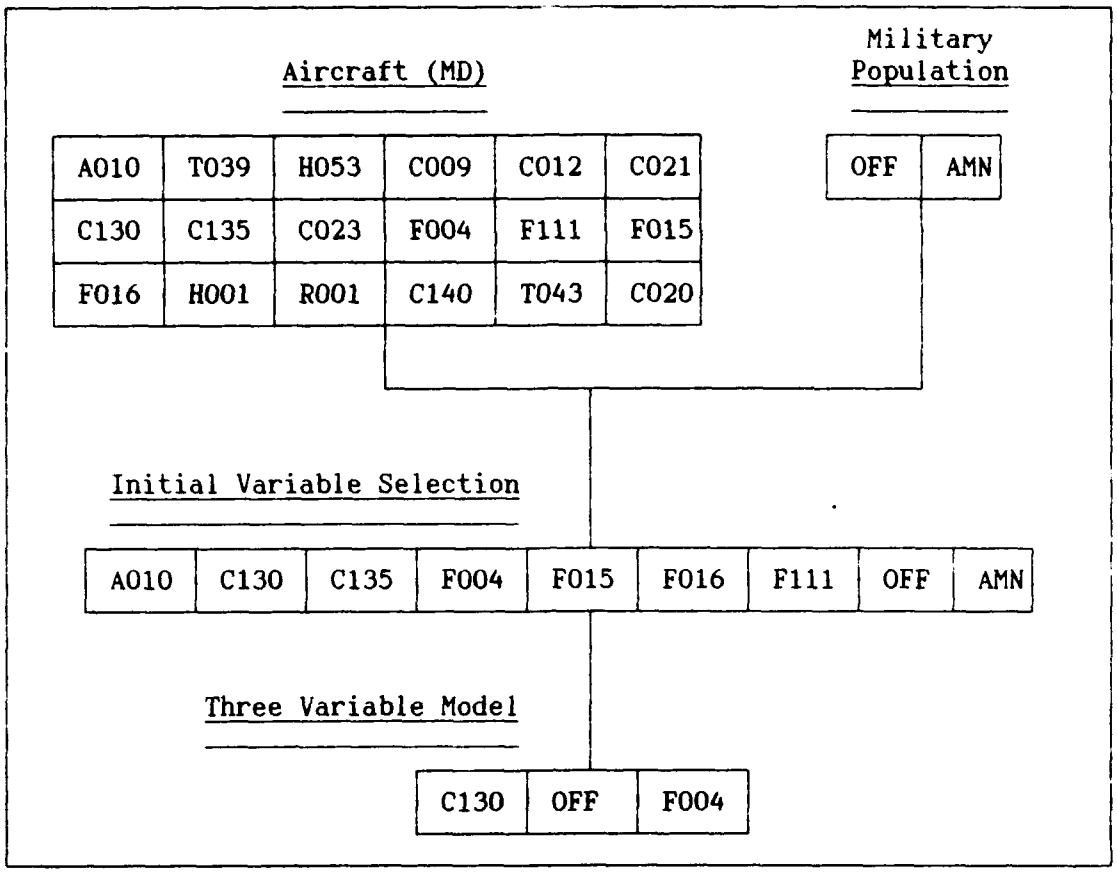

Figure 26. USAFE MSC Independent Variable Selection

Model Validation and Forecasting Evaluation. Appendix K is the complete SAS output of the multiple regression model. Table 14 displays a portion of the SAS output. The following diagnostic output showed the following:

1. The analysis of variance F-test.

H_o: $\beta_1 = \beta_2 = \beta_3 = 0$

H_a: At least one β_i does not equal 0

Test Statistic: $F = 7.136$

Rejection Region: $F > F_{.05} = 3.71$

where,

$\alpha = .05$

$v_1 = 3$

$v_2 = 10.$

The F test proves the model is significant at the .05 level of confidence and indicates the independent variables contribute information for the prediction of the tonnage (dependent) variable.

Table 14

USAFE MSC Multiple Regression Model Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	3	915845924.65	305281974.88	7.136	0.0076
Error	10	427834146.85	42783414.685		
C Total	13	1343680071.5			

Root MSE	6540.90320	R-square	0.6816
Dep Mean	73417.50000	Adj R-sq	0.5861
C.V.	8.90919		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	630601	175849.60077	3.586	0.0050
C130	1	4.464226	1.93861574	2.303	0.0440
F4	1	2.478257	0.68166009	3.636	0.0046
OFF	1	-59.980187	17.29779539	-3.468	0.0060

2. R^2 Value. The model has a moderately good fit with an R^2 value higher than the DSXR model (.68 compared to .16 for the DSXR model).

3. Residual Analysis. The plot of the residuals versus the predicted values (Appendix K, Figure 71) appears to be randomly distributed and shows no heteroscedasticity problem. The plots of the residuals versus each of the independent variables (C-130 and F-4) (Appendix K, Figures 72 and 73) also appear to be randomly distributed and show no problem with the assumption of linearity between tonnage and

each of the independent variables. Like the PACAF MSC model, the plot of residuals versus officer manpower (OFF) (Appendix K, Figure 74) appears to have a slight curvature. The Wilk Shapiro Test for Normality was used to determine whether the residuals were normally distributed (Appendix K).

H_o: The residual distribution function is a normal distribution function.

H_a: The residual distribution function is not a normal distribution function.

Test Statistic: $W = .951712$

Rejection Region: $W < W_{.05} = .874$

where,

$\alpha = .05$

$n = 14$.

The Wilk Shapiro Test indicates there is insufficient evidence to reject the null hypothesis and accept the alternative hypothesis based on a 95% confidence interval.

4. Multicollinearity. This model contains two variables which are correlated, F-4 flying hours and officer population ($r = .65$) (Appendix L). The variance inflation factors were computed to determine whether a problem of multicollinearity existed between the variables.

H_o: Variables x_1, x_2, and x_3 are more closely related to the dependent variable than each other.

H_a: Variables x_1, x_2, and x_3 are more closely related to each other than the dependent variable.

Test Statistic: $VIF = 1.04$ (x_1)

$VIF = 1.20$ (x_2)

$$VIF = 1.61 \ (x_3)$$

Rejection Region: $VIF > 1 / (1 - R^2) = 3.14$.

The results prove that the independent variables are more closely related to the dependent variable than to each other.

5. Outlier Detection. The output of the studentized residuals (Appendix K) shows all the residuals falling within 2 standard deviations. Based on this finding, no outliers are present in the data set.

6. Durbin Watson (DW) Test. This test was used to test for the existence of first order autocorrelation.

H_o: The residuals are not negatively autocorrelated.

H_a: The residuals are negatively autocorrelated.

Test Statistic: DW = 2.123

Rejection Region: $4 - d_l < DW < 4$ (negative autocorrelation)

Acceptance Region: $2 < DW < 4 - d_u$ (no autocorrelation, neg.test)

where,

$d_l = .82$

$d_u = 1.75$

$k = 3$

$n = 14$.

The test rejects the alternative hypothesis and accepts the null hypothesis based on a 95% level of confidence. The plot of the residuals versus N (Appendix K, Figure 75) (N = automatic observation counter that creates a sequential period indicator) is randomly distributed.

7. Forecasting Accuracy. The multiple regression model was used to forecast for the six quarter period from fiscal year (FY) 1988/3 to 1989/4. Table 15 shows the results of the multiple regression model

forecasts compared to the DSXR USAFE MSC model and the 12 quarter average tonnage forecasts.

Table 15

USAFE MSC Multiple Regression Model Forecasting Accuracy

FY/Qtr	Actual Tonnage	DSXR Model Forecasts	12 Qtr Average Forecasts	Multiple Regression Forecasts
1988/3	81479	71601	71714	80827
1988/4	76619	71977	71714	77635
1989/1	73847	66066	71714	76831
1989/2	63853	71427	71714	68757
1989/3	74806	81697	71714	75914
1989/4	88613	79357	71714	74597
	MAE:	7670	7443	4113
	Minimum Error:	4641	2133	652
	Maximum Error:	9877	16899	14016
	Rsquare:	.16	na	.68

The results of the forecasts show the multiple regression model with the lowest MAE and minimum absolute error compared to the other models. Figure 27 shows the multiple regression model producing more accurate forecasts for five of the six forecasts compared to the DSXR model.

8. Summary of Analysis. The USAFE MSC multiple regression model was developed in a similar manner to the PACAF MSC multivariable regression model. Both were developed from nine independent variables with the resulting models only having three statistically significant independent variables that were highly positively correlated ($r > .5$) and negatively correlated ($r < -.5$) with several of the other variables (Table 16) (Appendix L) that were not included in the model. Overall, the model was useful based on the results of the F-test (95% confidence level) and the residual analysis revealed no problems with heteroscedasticity and autocorrelation.

The relationship between tonnage and officer manpower may not be linear based on the results of the residual analysis. A second order (squared) officer manpower variable (SOFF) was added to the model to improve fit (Appendix K, Figure 76). The resulting model (Appendix K) had an improved fit (R^2 = .75) compared to the first order model, but it did not have improved forecasting capability (MAE = 5110). The resulting second order model was;

$$y = -30448299 + 4.58x_1 + 6050.50x_2 - .30x^2_2 + 2.71x_3$$

where: y = quarterly sealift tonnage

x_1 = quarterly C-130 flying hours

x_2 = quarterly officer population

x_3 = quarterly F-4 flying hours.

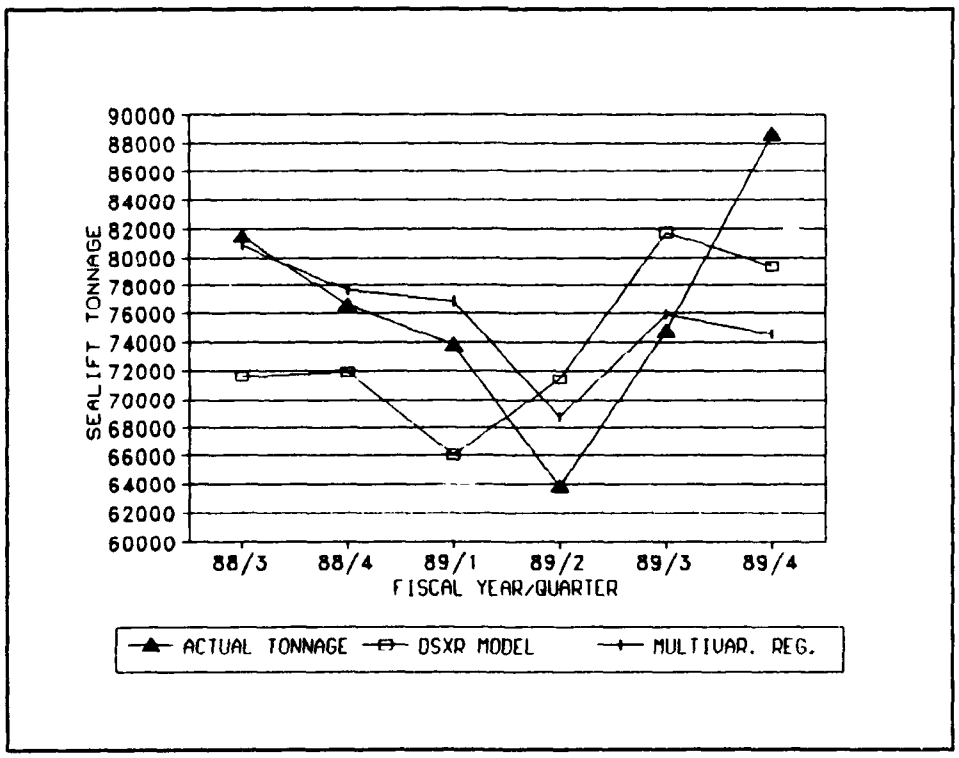

Figure 27. DSXR and Multiple Regression USAFE MSC Forecasts

Table 16

USAFE MSC Multiple Regression Model Independent Variable Correlations

Independent Variable	Positively Correlated Variables	Negatively Correlated Variables
C-130	(none)	(none)
F-4	C-135, AMN	F-16
OFF	(none)	F-16

According to the forecasting results, the multiple regression model achieved greater forecasting accuracy than the DSXR USAFE MSC model and the 12 quarter tonnage average. The multiple regression model contained a negative coefficient for the officer population variable which was similar to the PACAF MSC multiple regression model. The other two variables (C-130 and F-4 flying hours) had positive coefficients which were contrary to the negative coefficient for the F-16 variable in the PACAF MSC multiple regression model.

Neural Networks

The neural network forecasting models were developed and used in this research to determine whether they could produce more accurate forecasts. Neural networks are capable of recognizing and extracting patterns from data and are applicable to this type of problem.

Data. Similar to the regression models, fourteen quarters of data (1985/1 to 1988/2) were used to develop the networks, but the data (Appendix F, G, and J) were transformed so that the networks could process the data. The transformation involved converting the data so that it ranged between 0 and 1. Appendix M and N display the

transformation equations and the transformed data for the PACAF and USAFE flying hours and tonnages.

PACAF MSC Networks. Unlike the regression models, networks do not need to be specified a priori and multicollinearity is not a problem for the network. For this research, two network models were developed. One network (full multivariable network model) used all nine variables that were initially selected as independent variables (seven aircraft, two military population variables), and the other network (reduced multivariable network model) used the three variables from the multiple regression model (two aircraft variables, one population variable) (Figure 24).

The full multivariable network model (Figure 28) consisted of nine inputs (x_1 through x_9 in Table 17) and one output (y) with two hidden layers consisting of fourteen processing elements in the first layer and five processing elements in the second layer. The reduced multivariable network model (Figure 29) consisted of three inputs (x_1 through x_3 in Table 17) and one output (y) with two hidden layers consisting of eight processing elements in the first layer and four processing elements in the second layer. Some experimentation was required to find the optimal number of processing elements, but the first hidden layer typically contains more processing elements than inputs. Increasing the number of processing elements increases the processing capability of the network and increases the chances for the network to find the correct solution.

Table 17

PACAF MSC Network Independent Variables

	<u>Output</u>
y	= quarterly sealift tonnage
	<u>Inputs</u>
x_1	= quarterly A-10 flying hours
x_2	= quarterly F-16 flying hours
x_3	= quarterly officer population
x_4	= quarterly C-130 flying hours
x_5	= quarterly C-135 flying hours
x_6	= quarterly F-4 flying hours
x_7	= quarterly F-15 flying hours
x_8	= quarterly B-52 flying hours
x_9	= quarterly airman population.

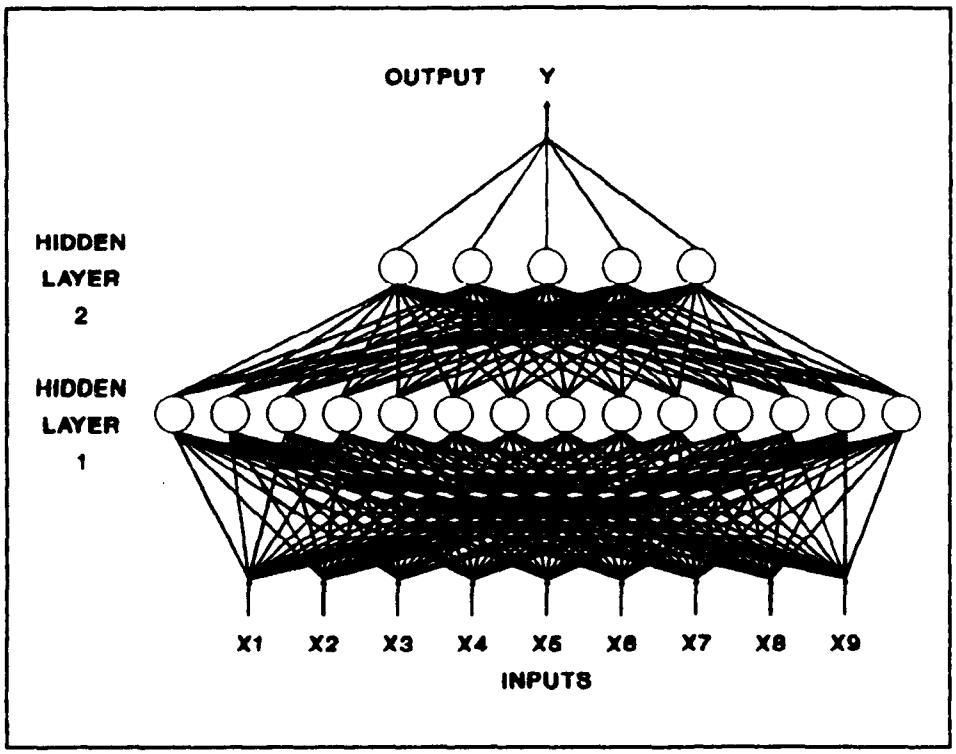

Figure 28. MSC Full Multivariable Network Model

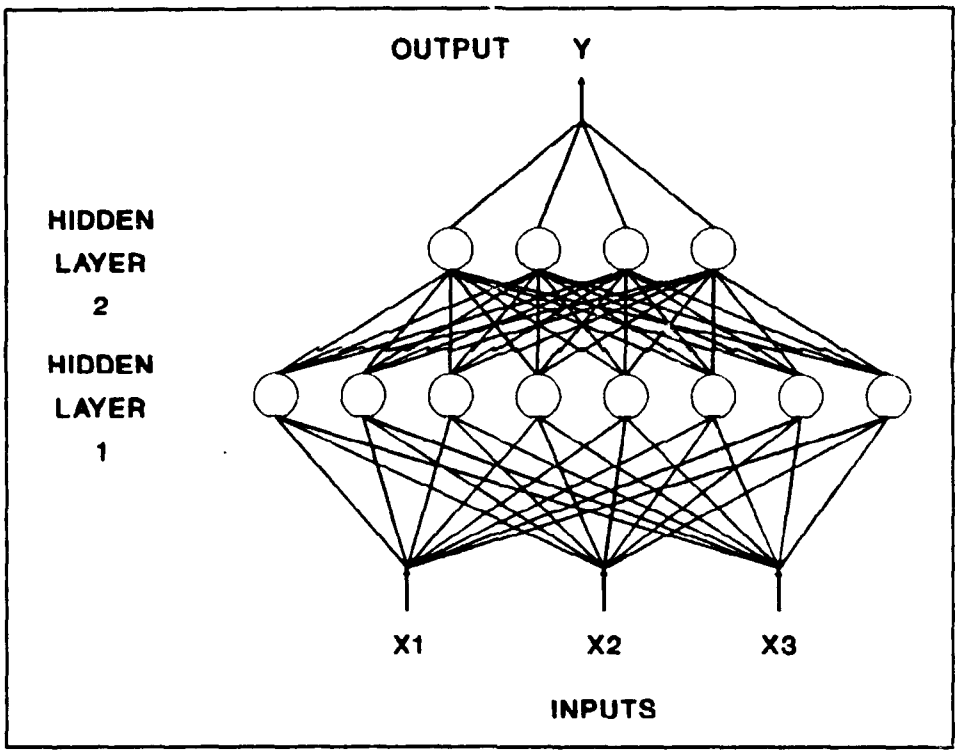

Figure 29. MSC Reduced Multivariable Network Model

Network Development. The number of training iterations used to develop a network affects the network's performance in terms of pattern recognition (R^2 value) and forecasting ability (MAE value). Figure 30 is a plot of the R^2 and MAE values with respect to the number of training iterations (full multivariable network). As the number of training iterations increase, the network continues to minimize the global error between the input and output values and as a consequence the R^2 value increases. The network makes rapid progress from 500 to 1500 training iterations, but after 1500 iterations the network's progress slows down (small increases in the R^2 value). Since the R^2 value began to plateau at approximately .65, the training was terminated at 4000 iterations and the network was evaluated.

The MAE values of the forecasts usually start out high, but in this case the MAE values are the lowest between 500 and 1000 training iterations. From 1 to 1000 training iterations, the network is forming mathematical relationships between the inputs and output and the network forecasts resemble a simple average of the actual tonnages. This explains the initial low MAE values because the forecast period (FY 1988/3 to FY 1989/4) could be predicted very accurately with a simple 12 quarter tonnage average. After 1000 training iterations, the MAE values begin to increase and fluctuate between 2300 and 2500.

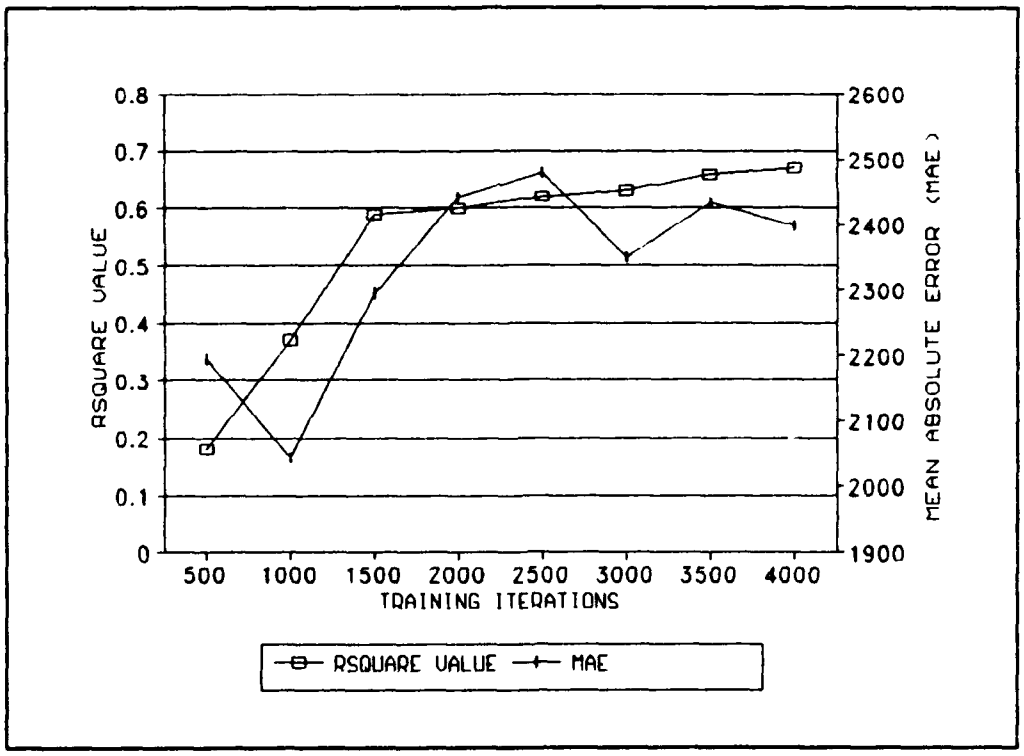

Figure 30. PACAF MSC Full Multivariable Network Model Training Iterations

Figure 31 is a plot of the R^2 and MAE values with respect to the number of training iterations for the reduced multivariable network.

After 2000 iterations, the network begins to plateau at an R^2 value between .6 and .65. Like the full multivariable network, the training was terminated at 4000 iterations and the network was then evaluated.

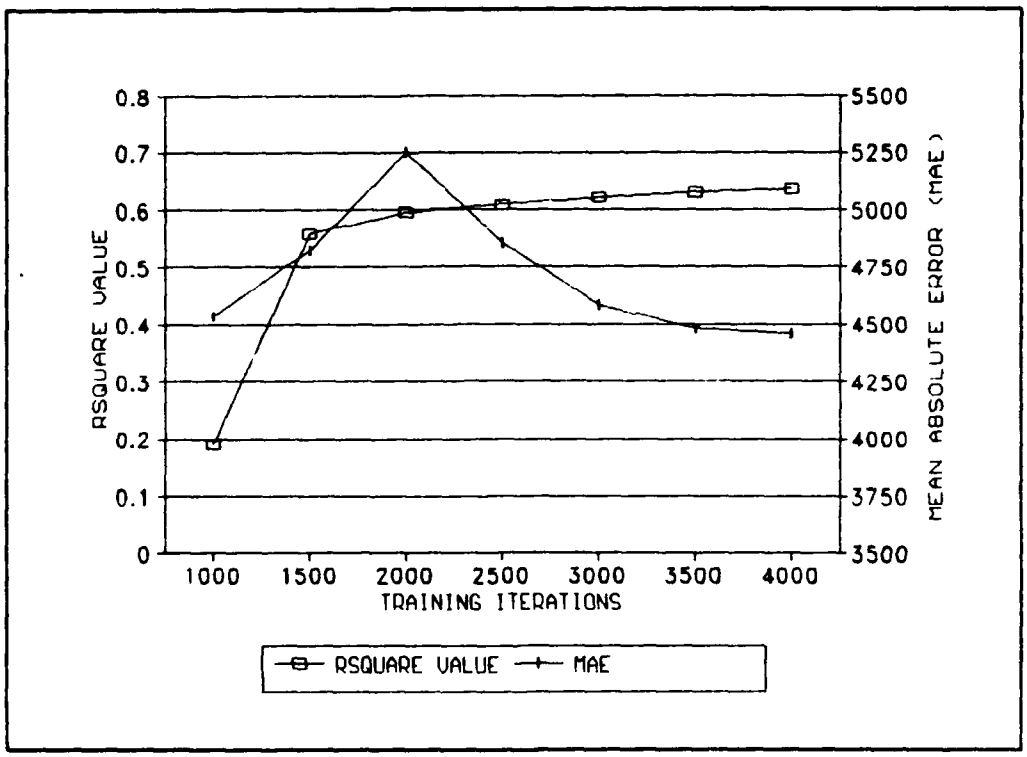

Figure 31. PACAF MSC Reduced Multivariable Network Model Training Iterations

Forecasting Evaluation. Table 18 compares the forecasting accuracy of the two network models with the DSXR model and the multiple (three variable) regression model (Appendix O is a complete output of the PACAF MSC full and reduced network models). Overall, the full multivariable network model with 4000 training iterations achieved the lowest MAE and the smallest minimum and maximum absolute error. The reduced multivariable network with 4000 training iterations (which used the same three variables as the multiple regression) performed similarly to the multiple regression model and appears to suffer from the same

extrapolation problems and subsequent overestimations in FY 1989/2 to FY 1989/4. The additional six variables used in the full network did not significantly contribute to increasing the R^2 value, but did contribute to improving forecasting accuracy. Figure 32 is a plot of the network forecasts which graphically shows the forecasting accuracy of the full network model and the overestimations of the reduced model.

Table 18

PACAF MSC Multivariable Network Forecasting Accuracy

FY/Qtr	Actual Tonnage	DSXR Model Forecasts	Multiple Regression Forecasts	Full Network Forecasts	Reduced Network Forecasts
88/3	55113	43902	50977	52321	52512
88/4	42250	40698	44943	46438	44522
89/1	43079	42714	40265	38324	39790
89/2	45086	44169	52728	45486	54536
89/3	44009	44614	47772	43804	48378
89/4	41224	38077	46229	39175	45988
	MAE:	2966	4342	2398	4457
	Minimum Error:	365	2693	205	2272
	Maximum Error:	11211	7642	4755	9450
	Rsquare:	.23	.65	.67	.64

USAFE MSC Networks. Like the PACAF MSC models, two USAFE MSC network models were developed (full multivariable network and reduced multivariable network). The full multivariable network model used all nine variables that were initially selected as independent variables (x_1 through x_9 in Table 19) and the reduced multivariable network model used the three variables from the multiple regression model (x_1 through x_3 in Table 19). The same network configurations that were used for the PACAF MSC models were used for the USAFE models (Figures 28 and 29).

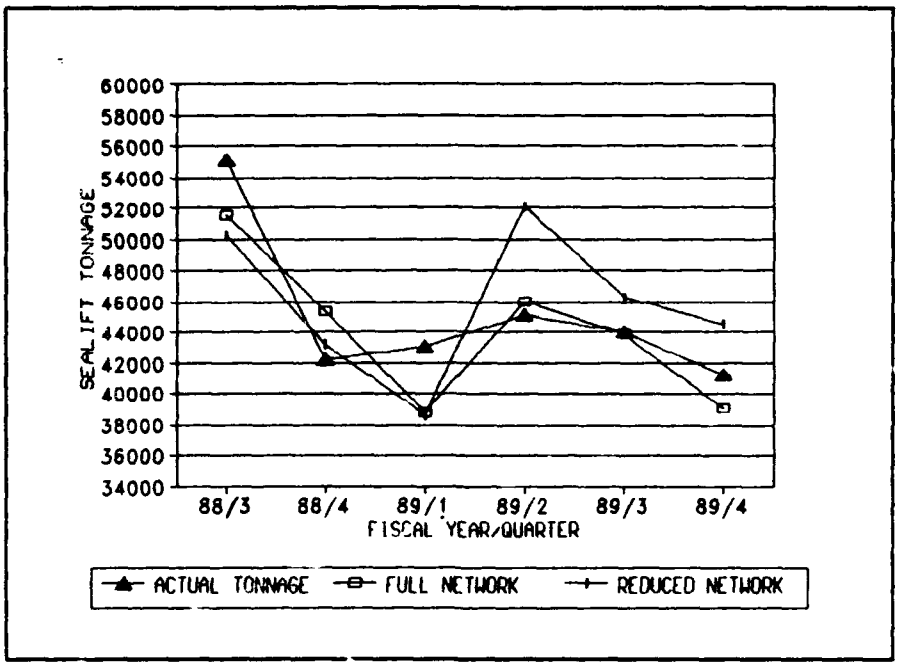

Figure 32. PACAF MSC Full and Reduced Multivariable Network Forecasts

Table 19

USAFE MSC Network Independent Variables

Output

y = quarterly sealift tonnage

Inputs

x_1 = quarterly C-130 flying hours

x_2 = quarterly F-4 flying hours

x_3 = quarterly officer population

x_4 = quarterly C-135 flying hours

x_5 = quarterly F-111 flying hours

x_6 = quarterly F-16 flying hours

x_7 = quarterly F-15 flying hours

x_8 = quarterly A-10 flying hours

x_9 = quarterly airman population.

Network Development. Figure 33 is a plot of the R^2 and MAE values with respect to the number of training iterations (full multivariable network). Similar to the PACAF MSC full multivariable network, this network makes rapid progress from 1000 to 1500 training

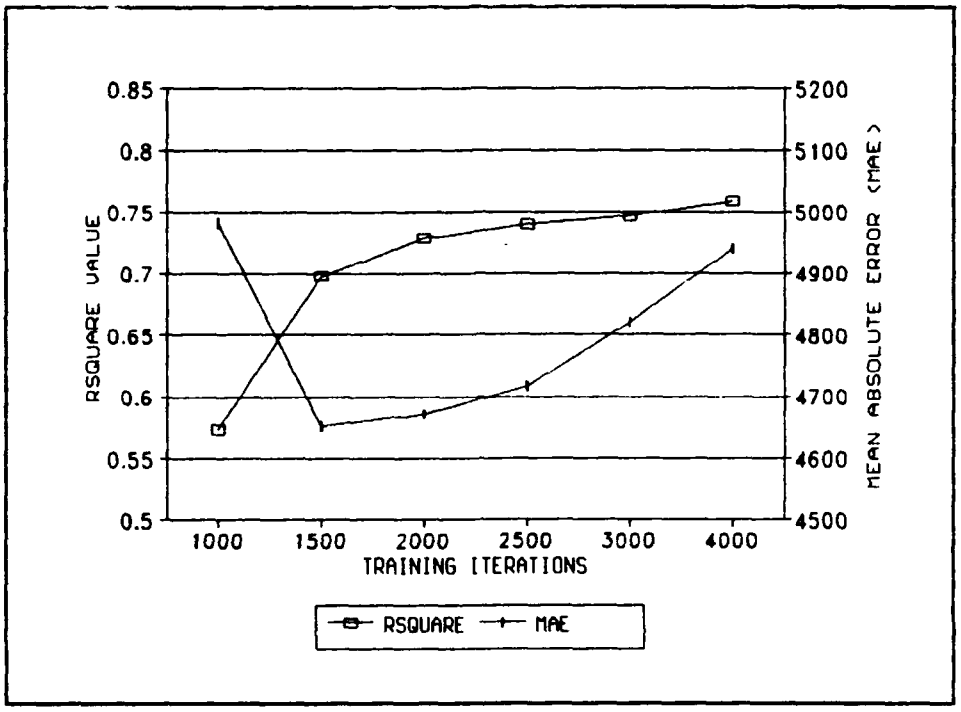

Figure 33. USAFE MSC Full Multivariable Network Model Training Iterations

iterations, but after 1500 iterations the network's progress slows down (small increases in the R^2 value). The MAE value of the forecasts starts out high, achieves a minimum at 1500 training iterations, and then begins to increase as the number of training iterations increase. The network continues to reduce the error between input and output values as the number of training iterations increase based on the higher R^2 values, but forecasting accuracy is lost based on the higher MAE values. Since the R^2 value began to plateau at .75, the training was terminated at 4000 iterations and the network was evaluated. If the training iterations

were increased beyond 4000 iterations, the network would begin to associate the unexplained error with the independent variables which is analogous to the over-fitting model problem in regression analysis.

Figure 34 is a plot of the R^2 and MAE values with respect to the number of training iterations for the reduced multivariable network. Similar to the full network, this network makes rapid progress from 500 to 1500 training iterations, but begins to plateau at an R^2 value of approximately .68 and a MAE value of approximately 3900. Training was terminated at 4000 iterations and the network was evaluated.

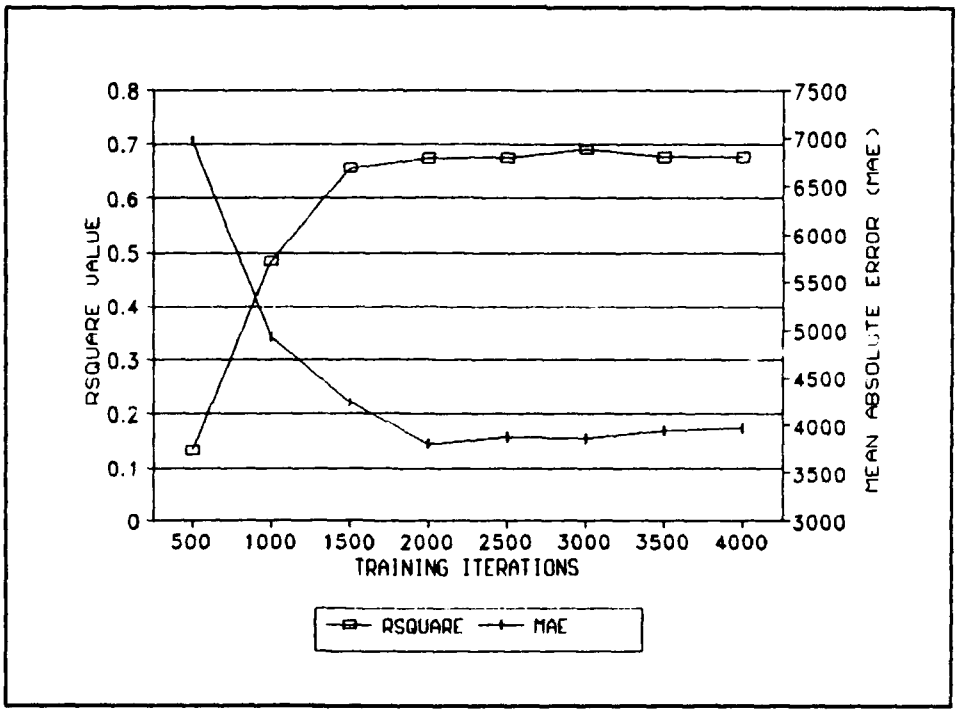

Figure 34. USAFE MSC Reduced Multivariable Network Model Training Iterations

Forecasting Evaluation. It was more difficult to achieve accurate USAFE MSC tonnage forecasts compared to the PACAF MSC tonnage forecasts because of the higher variability in the USAFE MSC tonnage

data. Table 20 compares the forecasting accuracy of the full network (4000 iterations), the reduced network (4000 iterations), the DSXR model and the multiple (three variable) regression model (Appendix O is a complete output of the full and reduced network models). The multiple regression and multivariable network models outperformed the DSXR model in MAE and minimum absolute error, while the full multivariable network outperformed the DSXR model in every category. The reduced network (which used the same three variables as the multiple regression) slightly outperformed the multiple regression model. The additional six variables used in the full network did not significantly contribute to increasing the R^2 value, but did contribute to improving forecasting accuracy compared to the DSXR model (Figure 35).

Table 20

USAFE MSC Multivariable Network Forecasting Accuracy

FY/Qtr	Actual Tonnage	DSXR Model Forecasts	Multiple Regression Forecasts	Full Network Forecasts	Reduced Network Forecasts
88/3	81479	71601	80827	76908	79684
88/4	76619	71977	77635	75211	75821
89/1	73847	66066	76831	76125	75344
89/2	63853	71427	68757	73171	67320
89/3	74806	81697	75914	81517	74269
89/4	88613	79357	74597	83334	72977
	MAE:	7670	4113	4940	3968
	Minimum Error:	4641	652	1474	537
	Maximum Error:	9877	14016	9318	15636
	Rsquare:	.16	.68	.76	.68

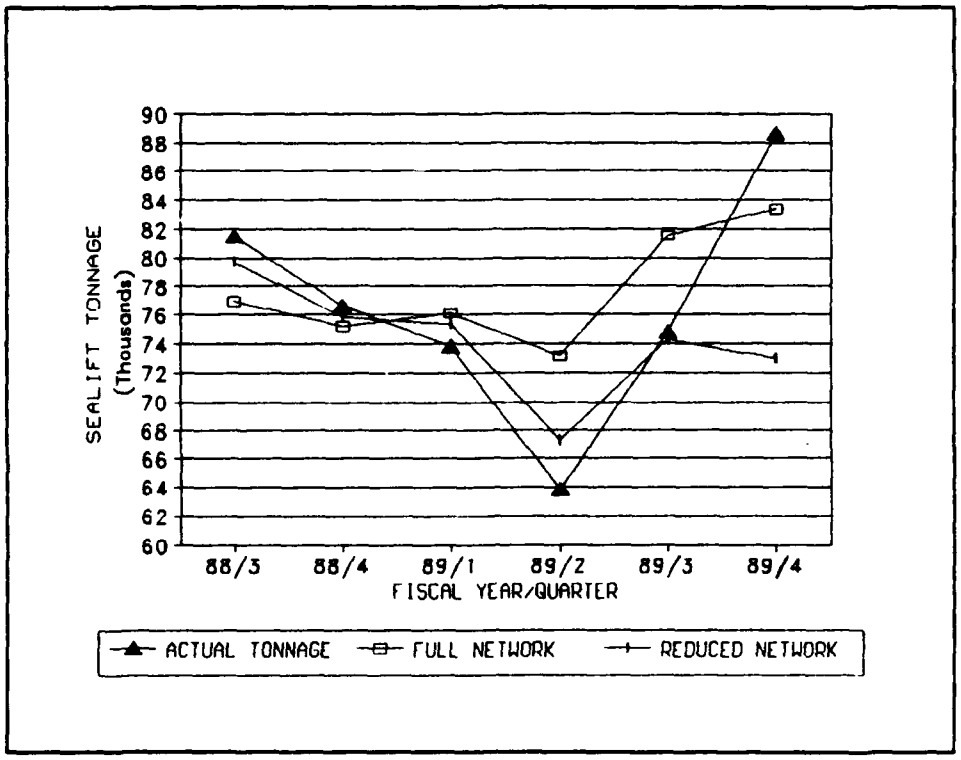

Figure 35. USAFE MSC Full and Reduced Multivariable Network Forecasts

Chapter Summary

This chapter started with an examination of the USAFE and PACAF sealift simple regression forecasting models presently used by DSXR. Multiple regression models using flying hours by aircraft type and military population variables were developed and tested for forecasting accuracy. The last section presented the development and testing of neural network forecasting models. The results of this analysis are further discussed in Chapter VI.

V. MAC SDT Forecasting Results and Analysis

This chapter is divided into five parts. The first part is an analysis of the PACAF and USAFE MAC tonnage data. Part two examines the DSXR simple regression models used to forecast MAC tonnage requirements to PACAF and USAFE. Part three and four present the development and results of the multiple regression models and the multivariable network models respectively. The last part presents the development and results of the time series forecasting networks.

In addition to the flying hour and manpower variables, other variables that have a major influence on MAC tonnage are airlift policies and budget restrictions. In September 1988 (FY 1988/4), transportation priority 2 (TP-2) cargo was restricted from the airlift system in order to save funds and as a consequence MAC tonnage to PACAF and USAFE dramatically declined. Over 90% of the total SDT cargo is sealifted, but 60% of the funds are spent on airlifting the remaining 10% of the total SDT tonnage (19).

In order to account for the decline in PACAF and USAFE MAC tonnage resulting from the TP-2 restriction, all the models (DSXR simple regression, multiple regression, and networks) were developed from data that contained the change. The forecast period for testing each models' forecasting accuracy is different from the MSC forecast period. The forecast period for these models is FY 1989/1 to 1990/1. Unlike the PACAF and USAFE MSC forecast period, the MAC forecast period has been reduced to five quarters because of the lack of data for FY 1990/2.

Data Analysis

Plots. Figure 36 graphically shows the quarterly PACAF MAC tonnage from FY 1978/1 to 1990/1. Despite the steadily increasing trend in PACAF flying hours from 1978 to 1988, the tonnage appears to be fairly consistent and fluctuates from 5500 tons to 7000 tons with a mean of approximately 6000 tons. Unlike the PACAF MSC tonnage which had large peaks in 1983 and 1985 and an increasing trend, the PACAF MAC tonnage has no apparent trend or seasonality. In the third quarter of 1988, the MAC tonnage begins to decline and reaches a lower tonnage level for 1989 as a result of the transportation priority 2 (TP-2) cargo restriction.

Figure 37 shows the quarterly USAFE MAC tonnage from FY 1978/1 to 1990/1. Unlike the USAFE flying hours which have been steadily increasing since 1978, this plot shows a slight downward trend in tonnage from 1978 to 1983 and then an increasing trend after 1983. The variability appears to be relatively constant from 1978 to 1985, but begins to increase in 1986. Similar to the PACAF data, this plot shows the sharp decrease in tonnage resulting from the TP-2 policy change. USAFE MAC tonnage does not have periods of increased tonnage requirements that are evident in the USAFE MSC tonnage plots.

Figure 38 is a plot of the PACAF MAC tonnage with respect to the PACAF total flying hours. The plot appears to be randomly distributed and does not indicate a linear relationship between flying hours and tonnage. Contrary to the PACAF MAC tonnage, the USAFE MAC tonnage (Figure 39) appears to show a slight linear relationship when flying hours are above 65,000 hours. Below 65,000 hours, the relationship does not appear linear.

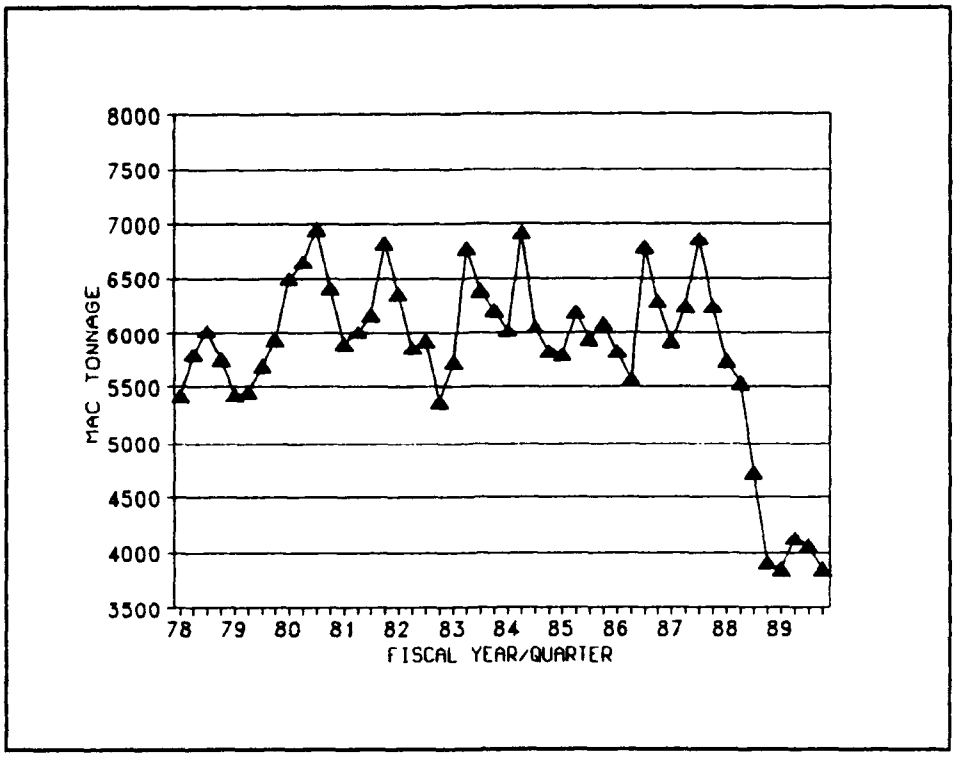

Figure 36. Quarterly PACAF MAC Tonnage

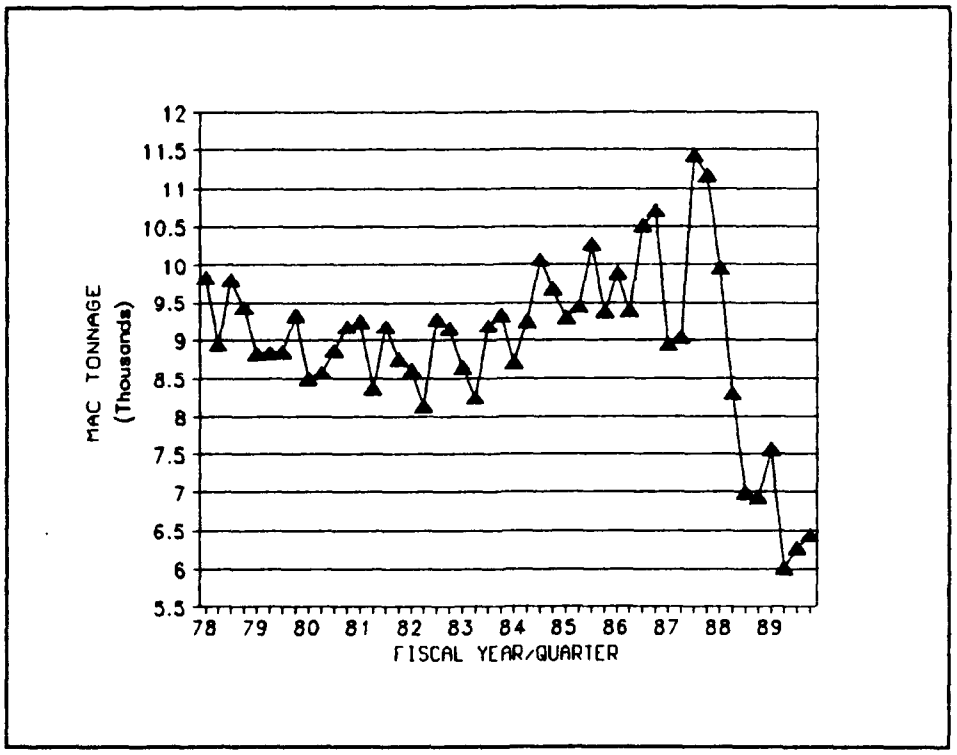

Figure 37. Quarterly USAFE MAC Tonnage

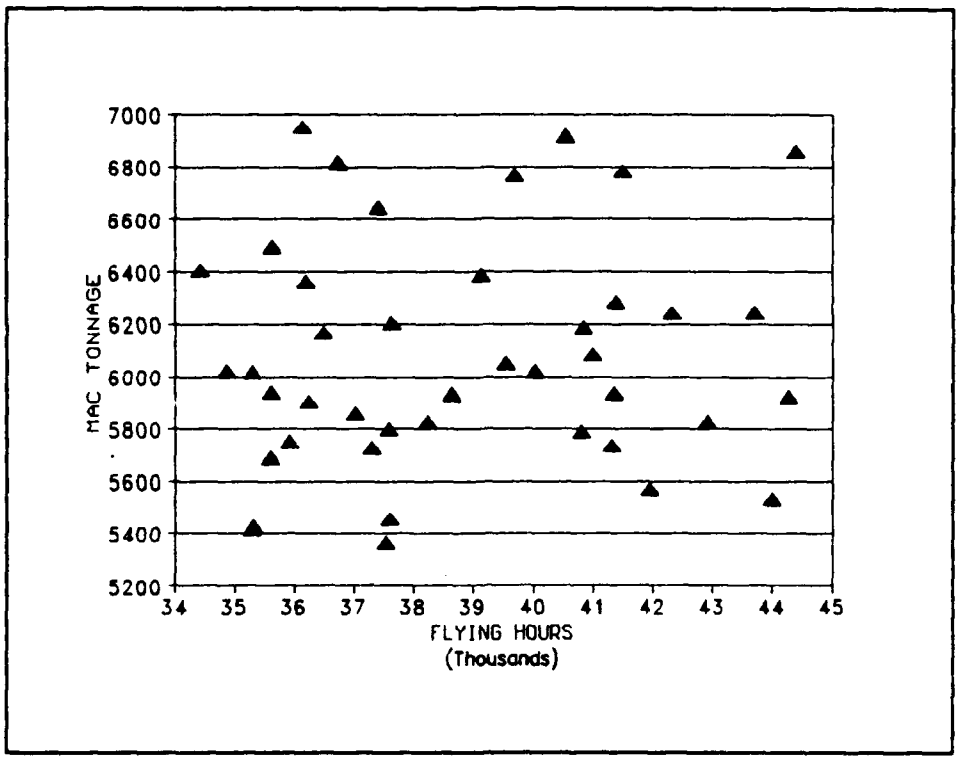

Figure 38. PACAF MAC Tonnage versus PACAF Flying Hours

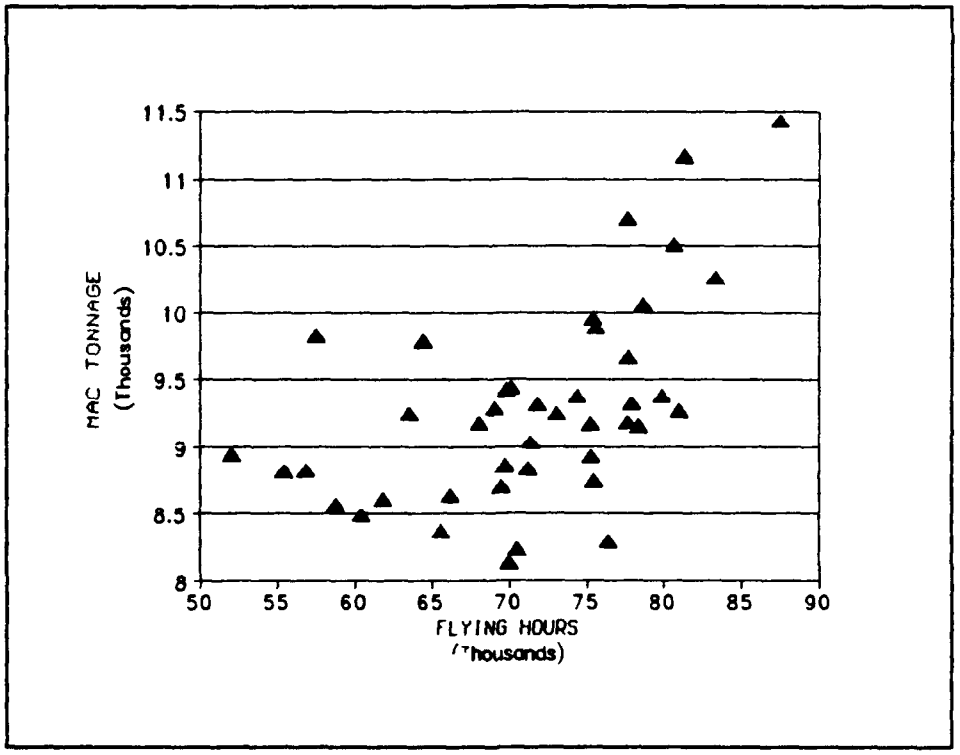

Figure 39. USAFE MAC Tonnage versus USAFE Flying Hours

Trend and Seasonal Analysis. PACAF and USAFE MAC tonnage and flying hours were analyzed using Gardner's trend and seasonal analysis methodology (Appendix P). Unlike the PACAF and USAFE MSC tonnage data sets, the PACAF and USAFE MAC tonnage data sets are not seasonal and do not exhibit a trend.

Pattern Identification. The plot of PACAF MAC tonnage with respect to fiscal year (Figure 36) appears to be random with no autoregressive (AR) or moving average (MA) aspect. Appendix Q is the SAS output of the autocorrelation analysis for this data set (FY 1978 to 1988/2). The Q-statistic indicates this series is white noise (random with no AR or MA aspect) since the value (Q = 7.71) is less than the chi square value (X^2 = 12.5916 with 6 df and 95% confidence interval). Although the Q-statistic and plot indicate the series has no patterns, the autocorrelation analysis revealed a small spike for the first lag in the ACF (r_1 = .365) and PACF which was beyond 2 standard deviations. This would indicate the possibility of an AR and/or MA aspect. The autocorrelations and partials drop to near zero after the third lag and signify the series is stationary.

The plot of USAFE MAC tonnage with respect to fiscal year (Figure 37) appears to have a 'wandering mean' and changing variance. The Q-statistic (Appendix Q) indicates this series is not white noise since the value (Q = 27.67) is greater than the chi square value (X^2 = 12.5916 with 6 df and 95% confidence interval). The autocorrelations are significantly different from zero after the second lag which means the series is not stationary. The analysis revealed spikes for the first lag and fourth lag of the ACF (r_1 = .40871, r_4 = .51296)

and PACF which were beyond 2 standard deviations which suggests an AR and/or MA aspect.

DSXR Simple Regression Model Validation and Forecasting Evaluation

DSXR PACAF MAC Model. The DSXR simple regression PACAF MAC model (dependent variable is PACAF airlift tonnage and the independent variable is the total PACAF aircraft flying hours) was developed from an altered data set (8 quarters, FY 1987/1 to FY 1988/4) that added TP-2 cargo tonnage back into quarters 2, 3 and 4 of FY 1988. A management decision was made to forecast the PACAF MAC tonnage without the change in policy because it was believed the decrease in tonnage requirements would not be sustained (19). Since the data was altered, only the forecasts (Table 21) are presented in this section.

Table 21

DSXR PACAF MAC Model Forecasting Accuracy *

FY/Qtr	Actual Tonnage	DSXR Model Forecasts
1989/1	3841	5782
2	4124	6064
3	4056	6150
4	3841	4883
1990/1	4305	4902
	MAE:	1523
	Minimum Error:	597
	Maximum Error:	2094
	Rsquare:	0.58

* based on 8 quarter regression model with:
slope = .21958
intercept = -3149.8

1. Forecasting Accuracy. The model was used to forecast for the five quarter period from fiscal year (FY) 1989/1 to 1990/1 using actual

flying hours as the independent variable. All the forecasts are overestimated because the model was developed from data that did not account for the decline in airlift resulting from the TP-2 restriction.

DSXR USAFE MAC Model. Appendix R contains the complete SAS output of the DSXR regression model used to forecast for the five quarter period from FY 1989/1 to 1990/1. The dependent variable is USAFE airlift tonnage and the independent variable is the total USAFE aircraft flying hours. Similar to the DSXR PACAF MAC model, this model was developed by using 8 quarters of data (1987/1 to 1988/4), but without altered data. Table 22 displays a portion of the SAS output. The following diagnostic output showed the following:

1. Two Tailed Test.

H_o: $\beta_1 = 0$

H_a: β_1 does not equal 0

Test Statistic: $t = 2.609$

Rejection Region: $t_{.025} < -2.447$, $t_{.025} > 2.447$

where: $\alpha = .05$, $df = 6$.

The two tailed test indicates the model is significant at the .05 level of confidence and indicates the flying hour (independent) variable contributes information for the prediction of the tonnage (dependent) variable.

2. R^2 Value. The R^2 value (.5315) is relatively low, but indicates the regression is an adequate model.

3. Residual Analysis. A plot of the residuals versus the predicted values (Appendix R, Figure 77) appears to be randomly distributed with no heteroscedasticity problem. The plot of the residuals versus flying

Table 22

DSXR USAFE MAC Model Analysis of Variance

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB>F
Model	1	10782499.865	10782499.865	6.808	0.0402
Error	6	9503160.0104	1583860.0017		
C Total	7	20285659.875			

Root MSE	1258.51500	R-square	0.5315
Dep Mean	9092.37500	Adj R-sq	0.4535
C.V.	13.84143		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	-9755.237091	7237.3126493	-1.348	0.2263
FH	1	0.244523	0.09371681	2.609	0.0402

hours (Appendix S, Figure 78) does not appear to show a problem with the assumption of linearity between tonnage and flying hours (the residuals seem to be randomly distributed). Since this model does not contain an intervention variable to account for the decline in airlift from the TP-2 restrictions, the plot of the residuals versus N (Appendix R, Figure 79) (N = successive time periods) shows how the errors become increasingly negative in period 6 through 8 (FY 1988/2 through 1988/4). The Wilk Shapiro Test for Normality was used to determine whether the residuals were normally distributed (Appendix R).

H_o: The residual distribution function is a normal distribution function.

H_a: The residual distribution function is not a normal distribution function.

Test Statistic: W = .912811

Rejection Region: $W < W_{.05} = .818$

where: $\alpha = .05$, $n = 8$.

The Wilk Shapiro Test indicates there is insufficient evidence to reject the null hypothesis based on a 95% confidence interval.

4. Outlier Detection. A plot of the residual values versus the studentized residuals (Appendix R) shows all the residuals falling within 2 standard deviations indicating no outliers are present in the data set.

5. Durbin Watson (DW) Test. Data set was too small to conduct this test.

6. Forecasting Accuracy. The model was used to forecast for the five quarter period from (FY) 1989/1 to 1990/1 using actual flying hours during this period. Table 23 shows the results of the model's forecasts.

Table 23

DSXR USAFE MAC Model Forecasting Accuracy

FY/Qtr	Actual Tonnage	DSXR Model Forecasts
1989/1	7569	6934
2	6005	8409
3	6260	11224
4	6432	10582
1990/1	5721	7003
	MAE:	2687
	Minimum Error:	635
	Maximum Error:	4964
	Rsquare:	.53

The results of the forecasts show the DSXR model overestimating four of the five quarters. The model has no intervention variable to account for the TP-2 restriction so the forecasts appear to be non-restricted TP-2 forecasts.

7. Analysis Summary. The model was a statistically valid model, but it did not account for the change in airlift from the TP-2 restrictions. The model's lack of fit with respect to the TP-2 restriction was evident in the residual plots (residuals versus N) and with the model's forecasts which were overestimated and appeared to be non-restricted TP-2 forecasts.

Multiple Regression Models

Similar to the PACAF and USAFE MSC models, the objective of the multiple regression models was to determine whether the breakout of the total flying hour variable into specific aircraft types and the addition of military population variables (Appendix F) contributed to increasing PACAF and USAFE airlift forecasting accuracy.

Data. Like the MSC data, fourteen quarters (three and a half years) of data (FY 1985 (quarter 3) to FY 1988 (quarter 4)) were used to develop the regression models. Five quarters (FY 1989/1 through FY 1990/1) were withheld to measure forecasting accuracy. In order to improve the fit of the PACAF and USAFE models, the MAC tonnage variable was transformed by taking its logarithm (LMAC).

Unlike the MSC models, an intervention variable (dummy variable) was added to both PACAF and USAFE models to account for the TP-2 restriction. From FY 1985/3 to 1988/2, the variable was 'turned on' with a '1' and from FY 1988/4 through 1990/1 the variable was 'turned off' with a '0'. Despite the fact that TP-2 restriction occurred in FY 1988/4, both data sets show major decreases in tonnage beginning in FY 1988/3. To account for this decline, the intervention variable was 'turned half on' with a '.5'.

The R^2, F-test, and Durbin Watson values were used to assess the fit of the models and the mean absolute error (MAE), minimum absolute error, and maximum absolute error were used to measure their forecasting accuracy.

PACAF MAC Multiple Regression Model Development. The methodology that was used to develop the PACAF MSC multiple regression model was used to develop the PACAF MAC multiple regression model. Out of the ten initially selected independent variables (seven aircraft, two military population variables and the TP-2 variable), six were eliminated because of multicollinearity problems or statistically non-significant t-values. The resulting four variable model was statistically significant at the 95% confidence level (Appendix S). The four variable model was;

$$y = 14.700077 - .000375x_1 + .000298x_2 - .000224x_3 + .575525x_4$$

where: y = quarterly airlift tonnage (logarithm)

x_1 = quarterly airman population

x_2 = quarterly B-52 flying hours

x_3 = quarterly F-15 flying hours

x_4 = TP-2 dummy variable.

Model Validation and Forecasting Evaluation. Appendix S is the complete SAS output of the multiple regression model. Table 24 displays a portion of the SAS output. The following diagnostic output showed the following:

1. The analysis of variance F-test.

H_o: $\beta_1 = \beta_2 = \beta_3 = \beta_3 = 0$

H_a: At least one β_i does not equal 0

Test Statistic: F = 21.417

Rejection Region: $F > F_{.05} = 3.63$

where: $\alpha = .05$, $v_1 = 4$, $v_2 = 9$.

The F test proves the model is significant at the .05 level of confidence and indicates the independent variables contribute information for the prediction of the tonnage (dependent) variable.

Table 24

PACAF MAC Multiple Regression Model Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	4	0.24888	0.06222	21.417	0.0001
Error	9	0.02615	0.00291		
C Total	13	0.27503			

Root MSE	0.05390	R-square	0.9049
Dep Mean	8.66209	Adj R-sq	0.8627
C.V.	0.62225		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	14.700077	3.45308085	4.257	0.0021
AMN	1	-0.000375	0.00022556	-1.664	0.1306
B52	1	0.000298	0.00013712	2.175	0.0576
F15	1	-0.000224	0.00008736	-2.563	0.0305
TP2	1	0.575525	0.09754821	5.900	0.0002

2. R^2 Value. The model has an excellent fit with a high R^2 value (.90).

3. Residual Analysis. A plot of the residuals versus the predicted values (Appendix S, Figure 80) appears to be randomly distributed and shows no heteroscedasticity problem. The plots of the residuals versus each of the independent variables (Appendix S, Figures 81 - 83) also appear to be randomly distributed and show no problem with the assumption

of linearity between tonnage and each of the independent variables. The Wilk Shapiro Test for Normality was used to determine whether the residuals were normally distributed.

H_o: The residual distribution function is a normal distribution function.

H_a: The residual distribution function is not a normal distribution function.

Test Statistic: $W = .986504$

Rejection Region: $W < W_{.05} = .874$

where: $\alpha = .05$, $n = 14$.

The Wilk Shapiro Test indicates there is insufficient evidence to reject the null hypothesis and accept the alternative hypothesis based on a 95% confidence interval.

4. Multicollinearity. The model contains two variables (B-52 and AMN) that are positively correlated ($r = .55$) (Appendix T). The variance inflation factors were computed to determine whether a problem of multicollinearity existed between the variables.

H_o: Variables x_1, x_2, x_3, and x_4 are more closely related to the dependent variable than each other.

H_a: Variables x_1, x_2, x_3, and x_4 are more closely related to each other than the dependent variable.

Test Statistic: $VIF = 2.65$ (x_1)

$VIF = 2.13$ (x_2)

$VIF = 2.61$ (x_3)

$VIF = 3.57$ (x_4)

Rejection Region: $VIF > 1 / (1 - R^2) = 10$.

The results prove that the independent variables are more closely related to the dependent variable than to each other.

5. Outlier Detection. The output of the studentized residuals (Appendix S) shows all the residuals falling within 2 standard deviations and no problems with outliers.

6. Durbin Watson (DW) Test. This test was used to test for the existence of first order autocorrelation.

H_o: The residuals are not negatively autocorrelated.

H_a: The residuals are negatively autocorrelated.

Test Statistic: DW = 2.409

Rejection Region: $4 - d_l < DW < 4$ (negative autocorrelation)

Acceptance Region: $2 < DW < 4 - d_u$ (no autocorrelation, neg.test)

where: $d_l = .69$, $d_u = 1.97$.

There is insufficient evidence to reject the null hypothesis and accept the alternative hypothesis based on a 95% level of confidence. The plot of the residuals versus N (Appendix S, Figure 84) (N = successive time periods) appears to be randomly distributed.

7. Forecasting Accuracy. The multiple regression model was used to forecast for the five quarter period from FY 1989/1 to 1990/1. Table 25 shows the results of the multivariable model forecasts compared to the DSXR PACAF MAC model.

Unlike the DSXR model, the multiple regression model underestimated on all five of the forecasts, but achieved the lowest mean absolute error, minimum absolute error, and maximum absolute error compared to the DSXR model. Figure 40 is a plot of the DSXR forecasts compared to the multiple regression forecasts which shows the overestimations by the DSXR

simple regression model and the underestimations by the multiple regression model.

Table 25

PACAF MAC Multiple Regression Model Forecasting Accuracy

FY/Qtr	Actual Tonnage	DSXR Model Forecasts	Multiple Regression Forecasts
1989/1	3841	5782	3258
2	4124	6064	2826
3	4056	6150	3298
4	3841	4883	3435
1990/1	4305	4902	2967
	MAE:	1523	877
	Minimum Error:	597	406
	Maximum Error:	2094	1338
	Rsquare:	0.58	0.90

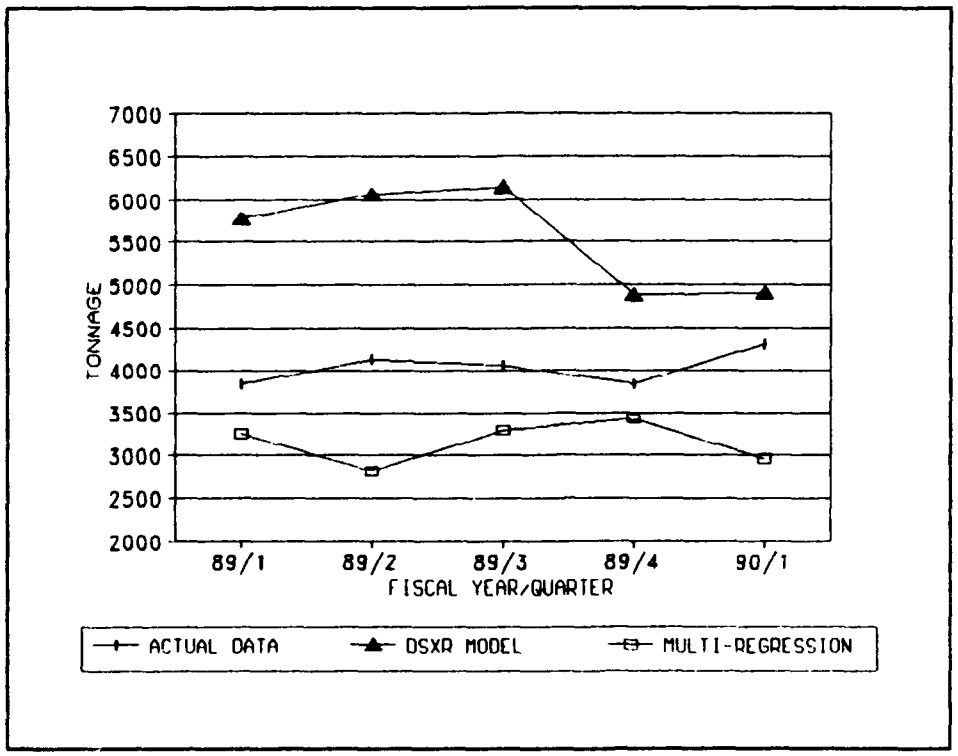

Figure 40. DSXR and Multiple Regression PACAF MAC Forecasts

8. Summary of Analysis. Overall, the model was useful based on the results of the F-test (95% confidence level) and the residual analysis revealed no problems with heteroscedasticity, nonlinearity, multicollinearity and autocorrelation. The independent variables were positively correlated ($r > +.5$) and negatively correlated ($r < -.5$) with several variables that were not included in the model (Table 26) (Appendix T). The model's forecasts were underestimated, but the forecast were more accurate than the DSXR forecasts.

Table 26

PACAF MAC Multiple Regression Model Independent Variable Correlations

Independent Variables	Positively Correlated Variables	Negatively Correlated Variables
AMN	F-16	OFF, C130, C135
B-52	(none)	(none)
F-15	(none)	(none)

USAFE MAC Multiple Regression Model. Similar to the PACAF MAC model, the USAFE MAC multiple regression model was developed from the ten initially selected independent variables (seven aircraft, two military population variables and the TP-2 variable). Six variables were eliminated because of multicollinearity problems or statistically non-significant t-values. The resulting four variable model was statistically significant at the 95% confidence level (Appendix U). The four variable model was;

$$y = 12.07 + .000712x_1 - .000396x_2 + .000061234x_3 + .22798x_4$$

where: y = quarterly airlift tonnage (logarithm)

x_1 = quarterly officer population

x_2 = quarterly airman population

x_3 = quarterly A-10 flying hours

x_4 = TP-2 dummy variable.

Model Validation and Forecasting Evaluation. Appendix U is the complete SAS output of the multiple regression model. Table 27 displays a portion of the SAS output. The following diagnostic output showed the following:

1. The analysis of variance F-test.

H_o: $\beta_1 = \beta_2 = \beta_3 = 0$

H_a: At least one β_i does not equal 0

Test Statistic: F = 38.162

Rejection Region: $F > F_{.05} = 3.63$

where: $\alpha = .05$, $v_1 = 4$, $v_2 = 9$.

The F test proves the model is significant at the .05 level of confidence and indicates the independent variables contribute information for the prediction of the tonnage (dependent) variable.

2. R^2 Value. The model has an excellent fit with a high R^2 value (.94).

3. Residual Analysis. A plot of the residuals versus the predicted values (Appendix U, Figure 85) appears to be randomly distributed and shows no heteroscedasticity problem. The plots of the residuals versus each of the independent variables (Appendix U, Figure 86 - 88) also appear to be randomly distributed and show no problem with the assumption of linearity between tonnage and each of the independent variables. The

Wilk Shapiro Test for Normality was used to determine whether the residuals were normally distributed.

H_o: The residual distribution function is a normal distribution function.

H_a: The residual distribution function is not a normal distribution function.

Test Statistic: $W = .937$

Rejection Region: $W < W_{.05} = .874$

where: $\alpha = .05$, $n = 14$.

The Wilk Shapiro Test indicates there is insufficient evidence to reject the null hypothesis and accept the alternative hypothesis based on a 95% confidence interval.

Table 27

USAFE MAC Multiple Regression Model Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	4	0.29379	0.07345	38.162	0.0001
Error	9	0.01732	0.00192		
C Total	13	0.31111			

Root MSE	0.04387	R-square	0.9443
Dep Mean	9.14729	Adj R-sq	0.9196
C.V.	0.47960		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	12.070129	2.30016919	5.247	0.0005
OFF	1	0.000712	0.00016972	4.193	0.0023
AMN	1	-0.000396	0.00010479	-3.776	0.0044
A10	1	0.000061234	0.00001052	5.820	0.0003
TP2	1	0.227948	0.06024109	3.784	0.0043

4. Multicollinearity. The airman population variable (AMN) was positively correlated ($r = .53$) with the officer population variable (OFF) and with the A-10 flying hour variable ($r = .52$) (Appendix V). The variance inflation factors were computed to determine whether a problem of multicollinearity existed between the variables.

H_o: Variables x_1, x_2, x_3, and x_4 are more closely related to the dependent variable than each other.

H_a: Variables x_1, x_2, x_3, and x_4 are more closely related to each other than the dependent variable.

Test Statistic: VIF = 3.56 (x_1)

VIF = 2.70 (x_2)

VIF = 1.41 (x_3)

VIF = 2.05 (x_4)

Rejection Region: $VIF > 1 / (1 - R^2) = 16.7$.

The results prove that both independent variables are more closely related to the dependent variable than to each other.

5. Outlier Detection. The output of the studentized residuals (Appendix W) shows all the residuals falling within 2 standard deviations and no problems with outliers.

6. Durbin Watson (DW) Test. This test was used to test for the existence of first order autocorrelation.

H_o: The residuals are not negatively autocorrelated.

H_a: The residuals are negatively autocorrelated.

Test Statistic: DW = 2.449

Rejection Region: $4 - d_l < DW < 4$ (negative autocorrelation)

Acceptance Region: $2 < DW < 4 - d_u$ (no autocorrelation, neg.test)

where: $d_l = .69$, $d_u = 1.97$.

There is insufficient evidence to reject the null hypothesis and accept the alternative hypothesis based on a 95% level of confidence. The plot of the residuals versus N (Appendix U, Figure 89) (N = automatic observation counter that creates a sequential period indicator) is randomly distributed.

7. Forecasting Accuracy. The multiple regression model was used to forecast for the five quarter period from FY 1989/1 to 1990/1. Table 28 shows the results of the multiple regression model forecasts compared to the DSXR USAFE MAC model.

Similar to the DSXR model, the multiple regression model overestimated on four of the five forecasts, but achieved the lo est mean absolute error, minimum absolute error, and maximum absolute error compared to the DSXR model. Figure 41 is a plot of the DSXR model forecasts compared to the multiple regression forecasts which shows the overestimations by the DSXR simple regression model compared to the multivariable regression model forecasts.

Table 28

USAFE MAC Multiple Regression Model Forecasting Accuracy

FY/Qtr	Actual Tonnage	DSXR Model Forecasts	Multiple Pegression Forecasts
1989/1	7569	6934	6398
2	6005	8409	6308
3	6260	11224	7019
4	6432	10582	7193
1990/1	5721	7003	6004
	MAE:	2687	655
	Minimum Error:	635	283
	Maximum Error:	4964	1171
	Rsquare:	.53	.94

8. Summary of Analysis. Overall, the model was a useful based on the results of the F-test (95% confidence level) and the residual analysis revealed no problems with heteroscedasticity, nonlinearity, multicollinearity and autocorrelation. Like all the other multiple regression models, the independent variables were positively correlated ($r > +.5$) and negatively correlated ($r < -.5$) with several variables that were not included in the model (Table 29) (Appendix V). The model's forecasts were overestimated, but the forecast were more accurate than the DSXR forecasts.

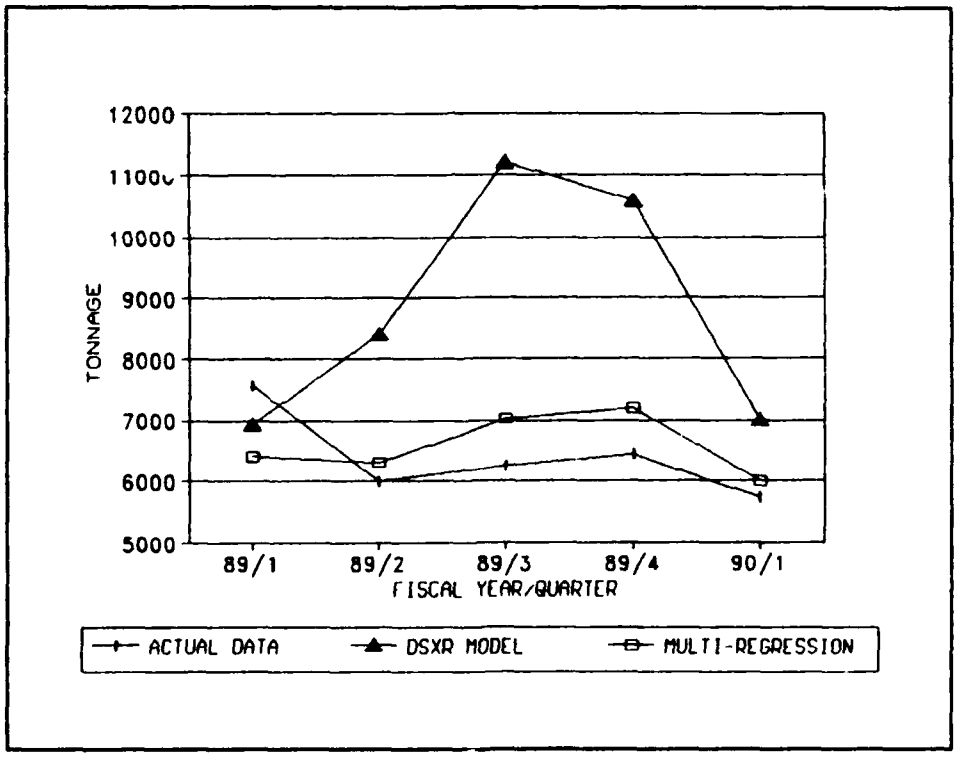

Figure 41. DSXR and Multiple Regression USAFE MAC Forecasts

Table 29

USAFE MAC Multiple Regression Model Independent Variable Correlations

Independent Variables	Positively Correlated Variables	Negatively Correlated Variables
OFF	F-4	F-16
AMN	F-4	(none)
A-10	F-111, C-135, F-4, F-15	(none)

Neural Networks

Similar to the PACAF and USAFE MSC networks, two types of neural network forecasting models were developed for the PACAF and USAFE MAC data sets. One type of network (full multivariable network or full network) used all 10 variables (7 aircraft variables, 2 military population, and the TP-2 restriction dummy variable), and the other network (reduced multivariable network or reduced network) used the same four variables used by the multivariable regression models.

Data. Similar to the regression models, fourteen quarters of data (1985/3 to 1988/4) were used to develop the networks (Appendix F, G, and J (actual data), Appendix M and N (transformed data)).

PACAF MAC Networks. The full multivariable network model consisted of ten inputs (x_1 through x_{10} in Table 30) and one output (y) with two hidden layers consisting of fifteen processing elements in the first layer and six processing elements in the second layer. The reduced multivariable network model consisted of four inputs (x_7 through x_{10} in Table 30) and one output (y) with two hidden layers consisting of ten processing elements in the first layer and five processing elements in the second layer.

Network Development. Figure 42 is a plot of the R^2 and MAE values with respect to the number of training iterations (full multivariable network). At 2500 iterations, the k^2 value begins to plateau at an approximate value of .81. The MAE values start out high, but decrease as the number of training iterations increase. Since the R^2 value became relatively stable, the training was terminated at 4000 iterations and the network was evaluated.

Figure 43 is a plot of the R^2 and MAE values with respect to the number of training iterations for the reduced multivariable network. Like the full multivariable network, this network's R^2 value increases

Table 30

PACAF MAC Network Independent Variables

Output

y = quarterly airlift tonnage

Inputs

x_1 = quarterly A-10 flying hours

x_2 = quarterly F-16 flying hours

x_3 = quarterly officer population

x_4 = quarterly C-130 flying hours

x_5 = quarterly C-135 flying hours

x_6 = quarterly F-4 flying hours

x_7 = quarterly F-15 flying hours

x_8 = quarterly B-52 flying hours

x_9 = quarterly airman population

x_{10} = TP-2 restriction dummy variable.

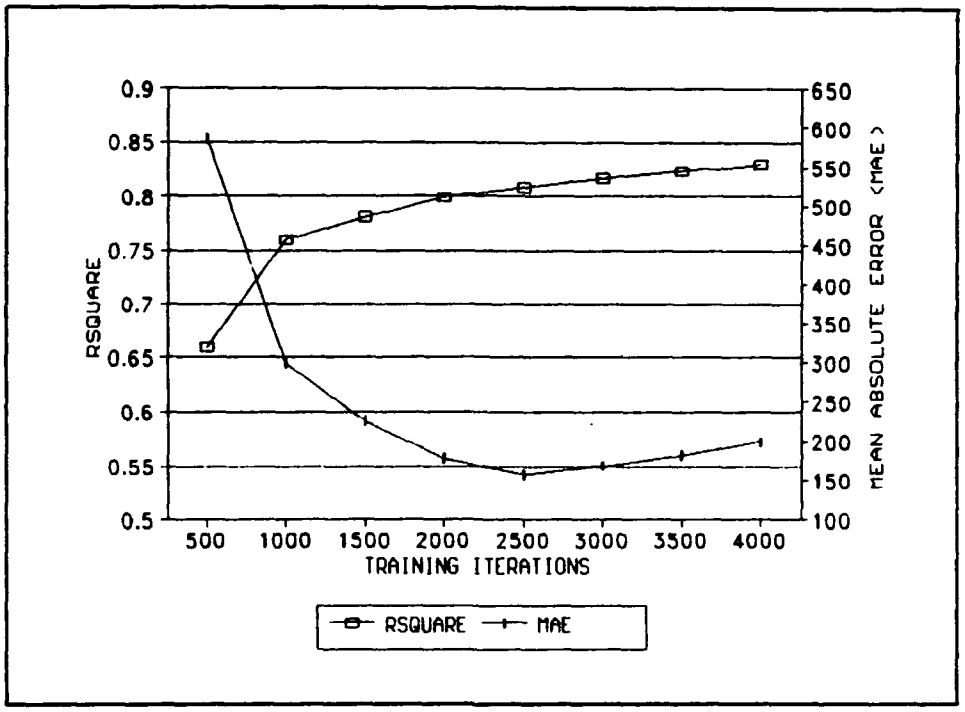

Figure 42. PACAF MAC Full Multivariable Network Training Iterations

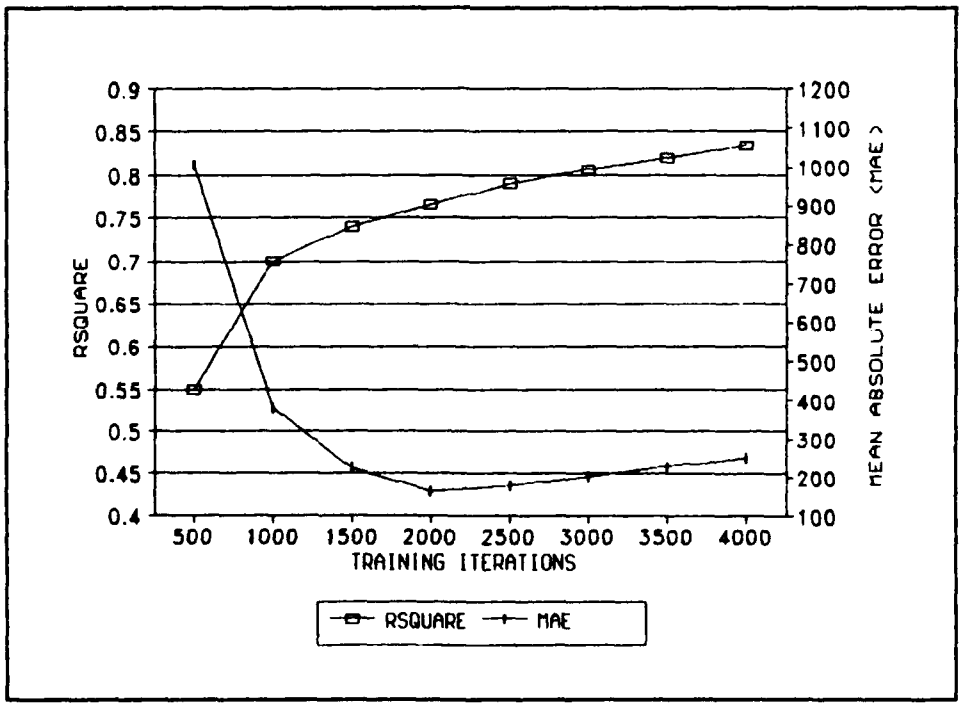

Figure 43. PACAF MAC Reduced Multivariable Network Training Iterations

with a decreasing rate as the number of training iterations increase and the MAE value starts out high and begins to decrease. Training for this network was also terminated at 4000 iterations.

Forecasting Evaluation. Table 31 compares the forecasting accuracy of the multivariable networks with the multiple (four variable) regression model (Appendix W is the output for the full and reduced network). The full multivariable network (4,000 training iterations) achieved the lowest MAE and the smallest minimum absolute error while the reduced multivariable network (4,000 training iterations) achieved the lowest maximum error. Overall, the reduced multivariable network and the full multivariable network had relatively similar forecasting capability and outperformed the multiple regression model. Figure 44 is a plot of the full and reduced multivariable network forecasts.

Table 31

PACAF MAC Multivariable Network Forecasting Accuracy

FY/Qtr	Actual Tonnage	Multiple Regression Forecasts	Full Network Forecasts	Reduced Network Forecasts
89/1	3841	3258	3890	3783
89/2	4124	2826	3905	3749
89/3	4056	3298	3947	3845
89/4	3841	3435	3829	3888
90/1	4305	2967	3698	3739
	MAE:	877	199	251
	Minimum Error:	406	12	47
	Maximum Error:	1338	607	566
	Rsquare:	.90	.83	.83

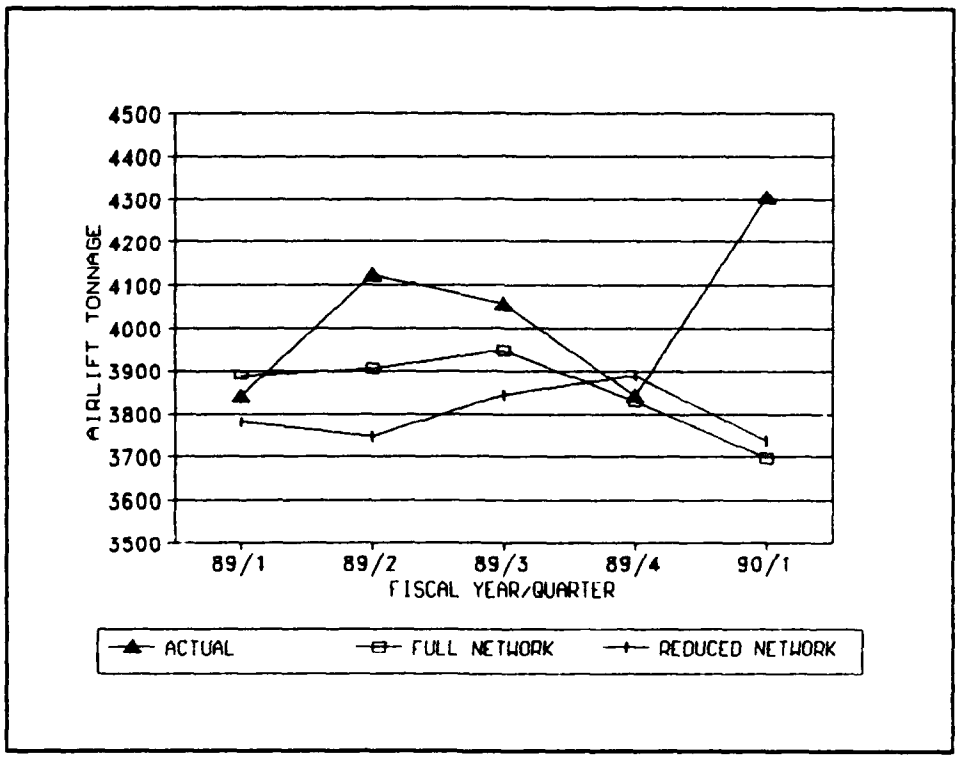

Figure 44. PACAF MAC Full and Reduced Multivariable Network Forecasts

USAFE MAC Networks. Similar to the PACAF MAC network, the full multivariable network model consisted of ten inputs (x_1 through x_{10} in Table 32) and one output (y) with two hidden layers consisting of fifteen processing elements in the first layer and six processing elements in the second layer. The reduced multivariable network consisted of four inputs (x_7 through x_{10} in Table 32) and one output (y) with two hidden layers consisting of ten processing elements in the first layer and five processing elements in the second layer.

Table 32

USAFE MAC Network Independent Variables

Output
y = quarterly sealift tonnage
Inputs
x_1 = quarterly C-130 flying hours
x_2 = quarterly F-4 flying hours
x_3 = quarterly C-135 flying hours
x_4 = quarterly F-111 flying hours
x_5 = quarterly F-16 flying hours
x_6 = quarterly F-15 flying hours
x_7 = quarterly A-10 flying hours
x_8 = quarterly officer population
x_9 = quarterly airman population.
x_{10} = TP-2 restriction dummy variable.

Network Development. Figure 45 is a plot of the R^2 and MAE values with respect to the number of training iterations for the full multivariable network. The relationship between the independent variables and airlift tonnage appears to be a stronger relationship than the relationship between the independent and airlift tonnage for the PACAF MAC data. At only 2,000 iterations, the network achieved a .92 R^2 value in contrast to the lower .79 R^2 value achieved by the PACAF MAC model. This network required significantly more training iterations than any other model. After 20,000 iterations, training was terminated and the network was then evaluated.

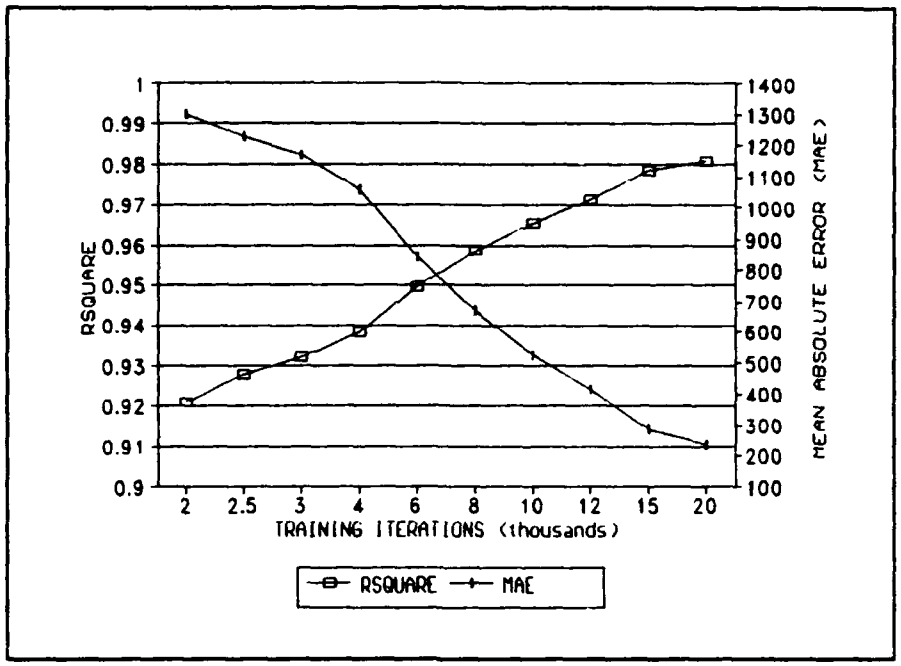

Figure 45. USAFE MAC Full Multivariable Network Training Iterations

Figure 46 is a plot of the R^2 and MAE values with respect to the number of training iterations for the reduced multivariable network. After 3000 training iterations, this network showed small increases in the R^2 value while the MAE value remained relatively constant. The training for this network was terminated at 4000 iterations.

Forecasting Evaluation. Table 33 compares the forecasting accuracy of the multivariable network models with the multiple (four variable) regression model (Appendix W is the full and reduced network output). Overall, the full multivariable network model (with 20,000 training iterations) achieved the lowest MAE and the smallest minimum and maximum absolute error. The multivariable regression model slightly outperformed the reduced multivariable network (with 4,000 training

iterations). Figure 47 is a plot of the full and reduced multivariable network forecasts.

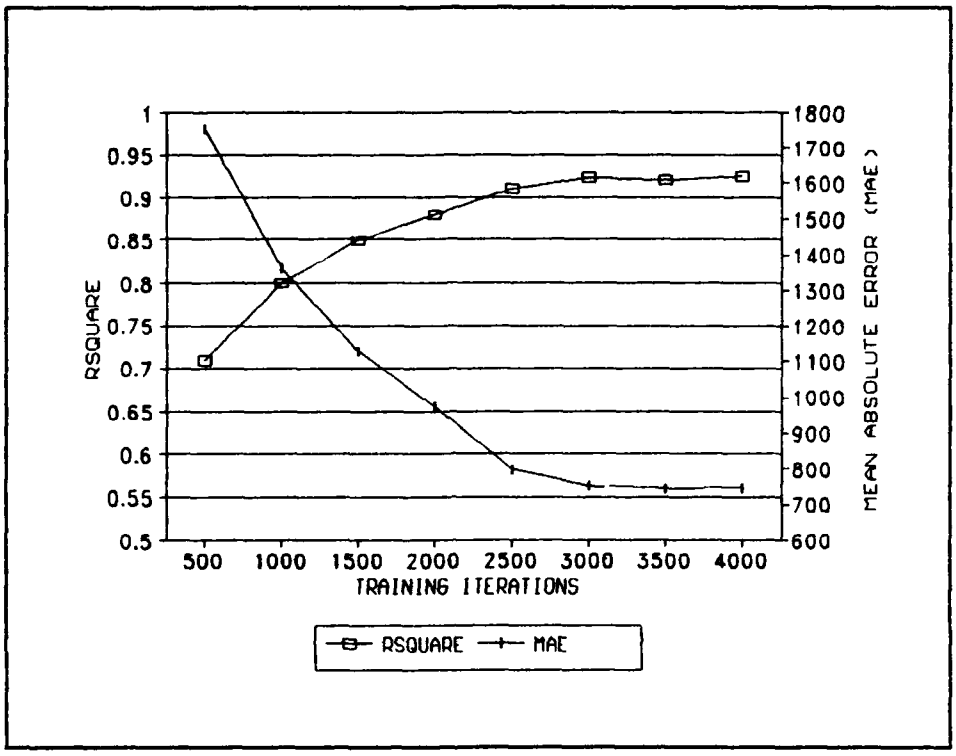

Figure 46. USAFE MAC Reduced Multivariable Network Training Iterations

Table 33

USAFE MAC Multivariable Network Forecasting Accuracy

FY/Qtr	Actual Tonnage	Multiple Regression Forecasts	Full Network Forecasts	Reduced Network Forecasts
89/1	7569	6398	6642	6498
89/2	6005	6308	6013	6462
89/3	6260	7019	6154	7096
89/4	6432	7193	6389	7275
90/1	5721	6004	5824	6247
	MAE:	655	237	746
	Minimum Error:	283	8	457
	Maximum Error:	1171	927	1071
	Rsquare:	.94	.98	.92

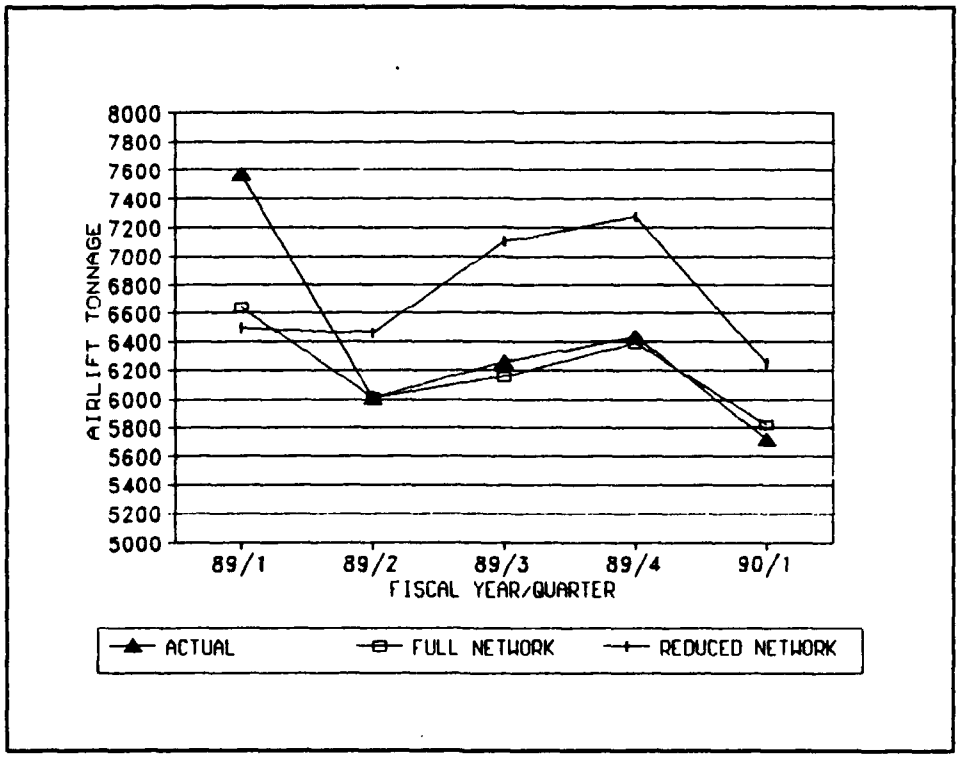

Figure 47. USAFE MAC Full and Reduced Multivariable Network Forecasts

Time Series Neural Networks

As an alternative to using aircraft flying hour and military population independent variables, neural network forecasting models were developed using a univariate time series approach. The time series forecasting methodology is based on the principle that historical changes in the data will be repeated in the future. The historical changes for the PACAF and USAFE MSC tonnages which were characterized with time periods of increased tonnage requirements (large peaks) are unlikely to be repeated in the near future because of the eminent reduction in forces at overseas locations. Unlike the MSC data, the PACAF and USAFE MAC tonnage data are applicable to this approach since the data are more stable and do not exhibit the large peaks that are evident in the MSC data.

Data. PACAF and USAFE MAC actual tonnage data from FY 1983/1 to 1988/4 (five years) (Appendix X) were used to develop the time series networks. The same transformation equations that were used for the previous networks were used for the time series networks. The forecast period (FY 1989/1 to 1990/1) also remained the same.

PACAF and USAFE MAC Time Series Network Development. The PACAF and USAFE MAC time series network configuration (Figure 48) consisted of five inputs and one output (y) with two hidden layers consisting of ten processing elements in the first layer and five processing elements in the second layer. The networks used four previous quarters of airlift tonnage and the TP-2 dummy variable (x_1 through x_5) as inputs and the output (y) was the forecast for the next quarter's tonnage at time t + 1 (Table 34).

The network training examples were presented to the time series networks in a 'shift register' manner (Table 35) and each network was developed with 10,000 training iterations. The intervention variable (x_5) was used in a similar fashion to the multiple regression intervention variables. The PACAF MAC time series network model achieved a R^2 value of .74 while the USAFE time series network model achieved a R^2 value of .68. Figures 49 (PACAF MAC time series model) and 50 (USAFE MAC time series model) display each network's pattern recognition capability by showing the actual data versus the network's output. Each network data point (Fy 1984/1 to 1988/4) represents a predicted value for tonnage at t+1 based on the previous four quarters of actual tonnage which the network was presented during the training iterations. The network data points (FY 1989/1 to 1990/1) are actual forecasts.

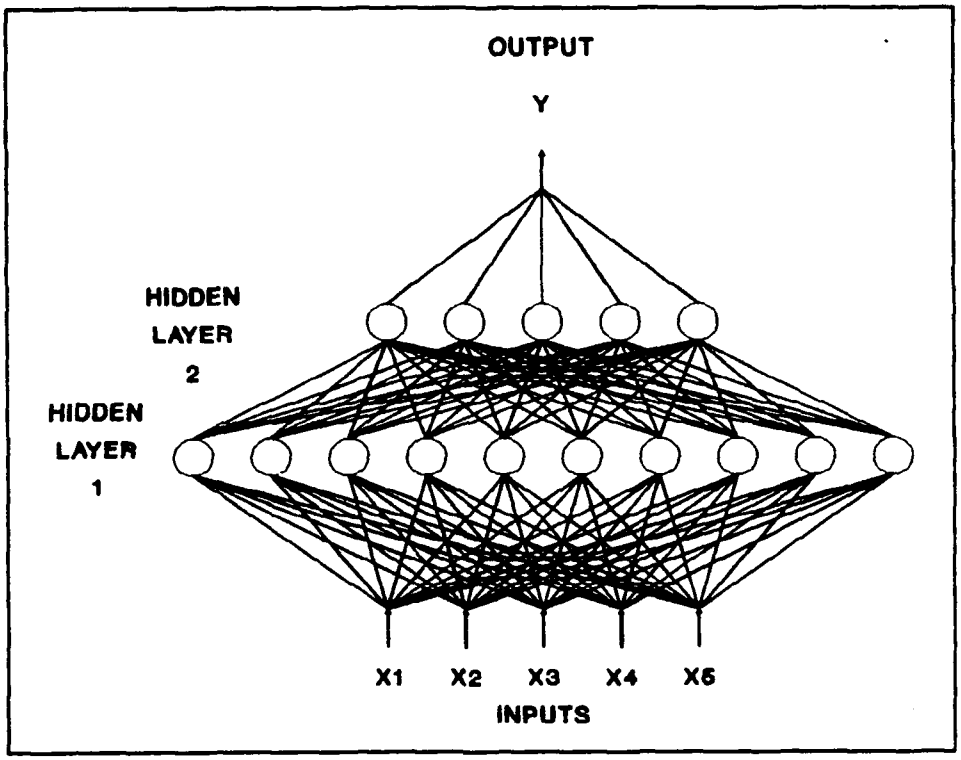

Figure 48. USAFE and PACAF MAC Time Series Network Configuration

Table 34

PACAF and USAFE MAC Time Series Network Independent Variables

Output

y = airlift tonnage at time (t+1)

Inputs

x_1 = airlift tonnage at time (t)

x_2 = airlift tonnage at time (t-1)

x_3 = airlift tonnage at time (t-2)

x_4 = airlift tonnage at time (t-3)

x_5 = TP-2 intervention (dummy) variable.

Table 35

Time Series Network Training and Forecasting Iteration Methodology

Training Iteration	Inputs (FY/Qtr) x_1	x_2	x_3	x_4	x_5	Predicted Value (FY/Qtr) y
1	83/1	83/2	83/3	83/4	.9	84/1
2	83/2	83/3	83/4	84/1	.9	84/2
.	.	.	.	.	.	.
.	.	.	.	.	.	.
18	87/2	87/3	87/4	88/1	.9	88/2
19	87/3	87/4	88/1	88/2	.45	88/3
20	87/4	88/1	88/2	88/3	0	88/4

Forecast Iteration	Inputs (FY/Qtr) x_1	x_2	x_3	x_4	x_5	Forecast (FY/Qtr) y
1	88/1	88/2	88/3	88/4	0	89/1
2	88/2	88/3	88/4	89/1	0	89/2
.	.	.	.	.	.	.
.	.	.	.	.	.	.
5	89/1	89/2	89/3	89/4	0	90/1

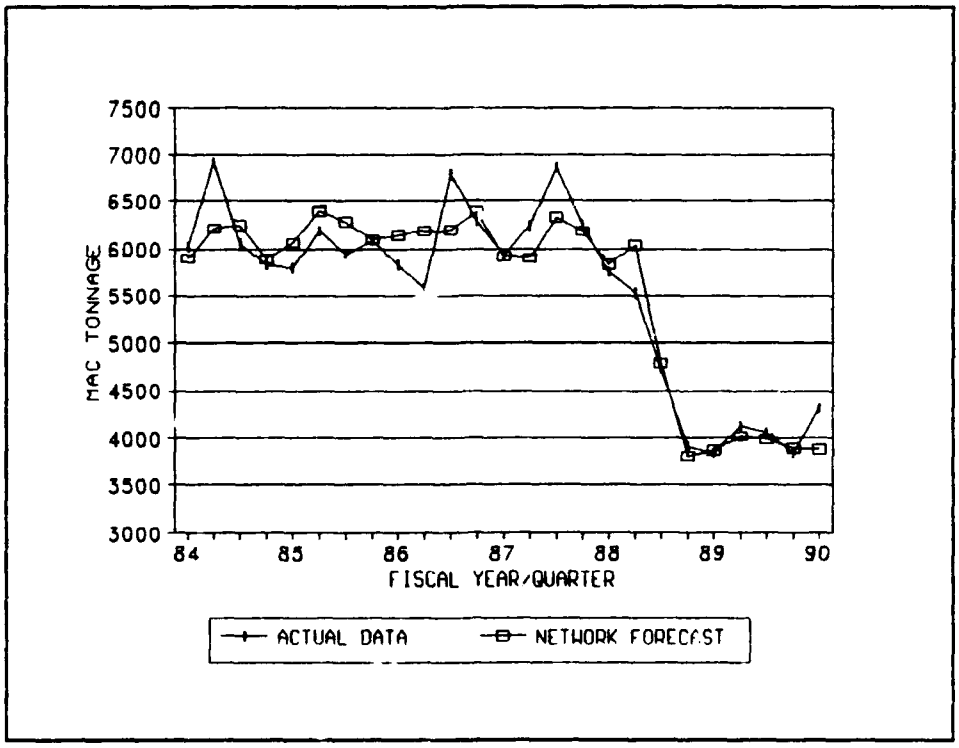

Figure 49. PACAF MAC Time Series Network Pattern Recognition

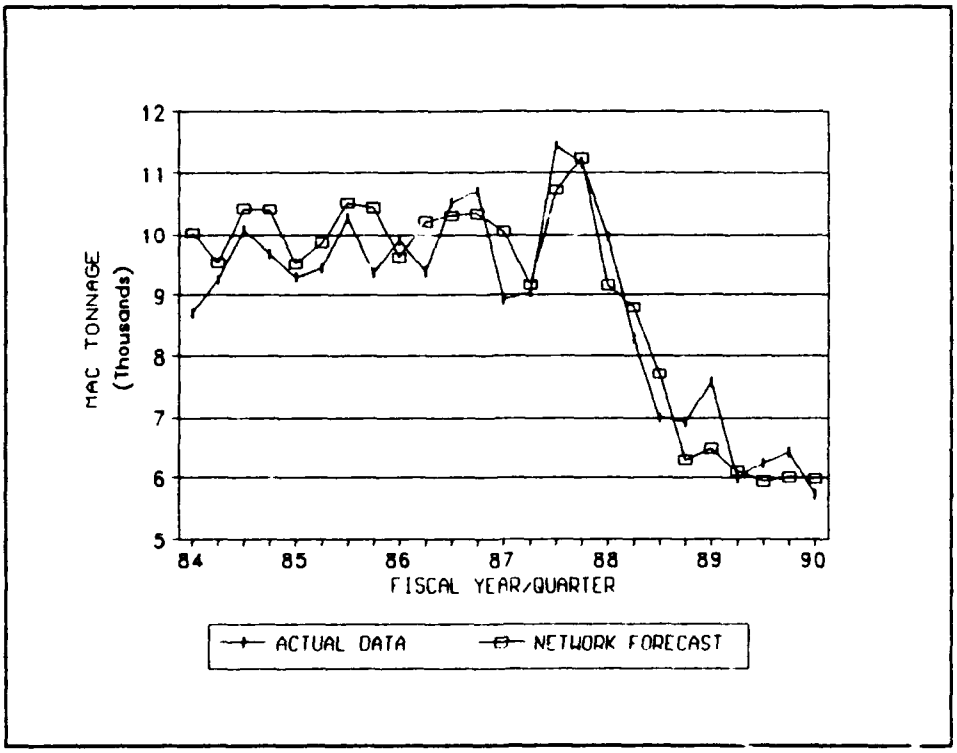

Figure 50. USAFE MAC Time Series Network Pattern Recognition

PACAF MAC Time Series Network Forecasting Evaluation. Since the PACAF MAC full multivariable network outperformed all other PACAF MAC models, the forecasting accuracy of the PACAF MAC time series network model was compared with the PACAF MAC full multivariable network (Table 36) (Appendix Y is the PACAF and USAFE time series network output). The time series network model slightly outperformed the full multivariable network by achieving the lowest MAE and the smallest maximum absolute error. Figure 51 graphically shows the forecasting accuracy of the time series network model and the multivariable network model.

Table 36

PACAF MAC Time Series Network Forecasting Accuracy

FY/Qtr	Actual Tonnage	Full Network Forecasts	Time Series Network Forecasts
1989/1	3841	3890	3868
2	4124	3905	4014
3	4056	3947	3987
4	3841	3829	3893
1990/1	4305	3698	3887
	MAE:	199	135
	Minimum Error:	12	27
	Maximum Error:	607	418
	Rsquare:	.83	.74

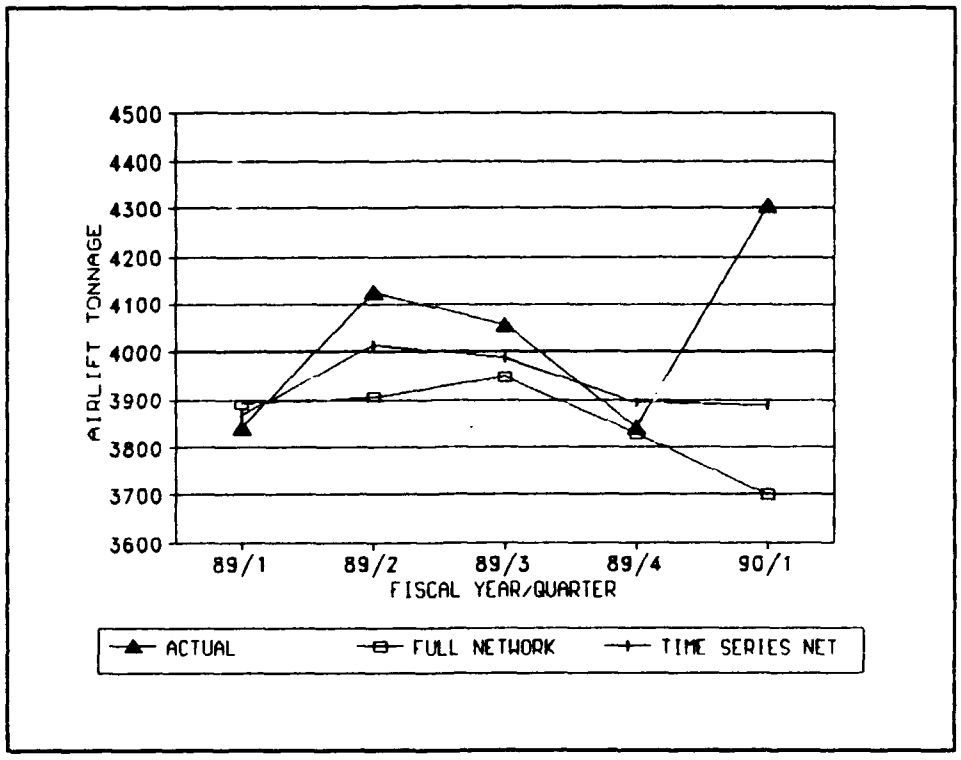

Figure 51. PACAF MAC Time Series Network Forecasts

USAFE MAC Time Series Network Forecasting Evaluation. Table 37 compares the forecasting accuracy of the USAFE MAC time network model with the USAFE MAC full multivariable network (full network outperformed all other previous models). Unlike the PACAF MAC data, the full multivariable network model achieved the lowest MAE, the smallest minimum absolute error, and the smallest maximum absolute error. The time series network outperformed the USAFE MAC multiple regression and reduced multivariable network in MAE and smallest minimum error. Figure 52 graphically shows the forecasting accuracy of the time series network model and the full multivariable network model.

Table 37

USAFE MAC Time Series Network Forecasting Accuracy

FY/Qtr	Actual Tonnage	Full Network Forecasts	Time Series Network Forecasts
1989/1	7569	6642	6482
2	6005	6013	6115
3	6260	6154	5946
4	6432	6389	6012
1990/1	5721	5824	5986
	MAE:	237	439
	Minimum Error:	8	110
	Maximum Error:	927	1087
	Rsquare:	.98	.68

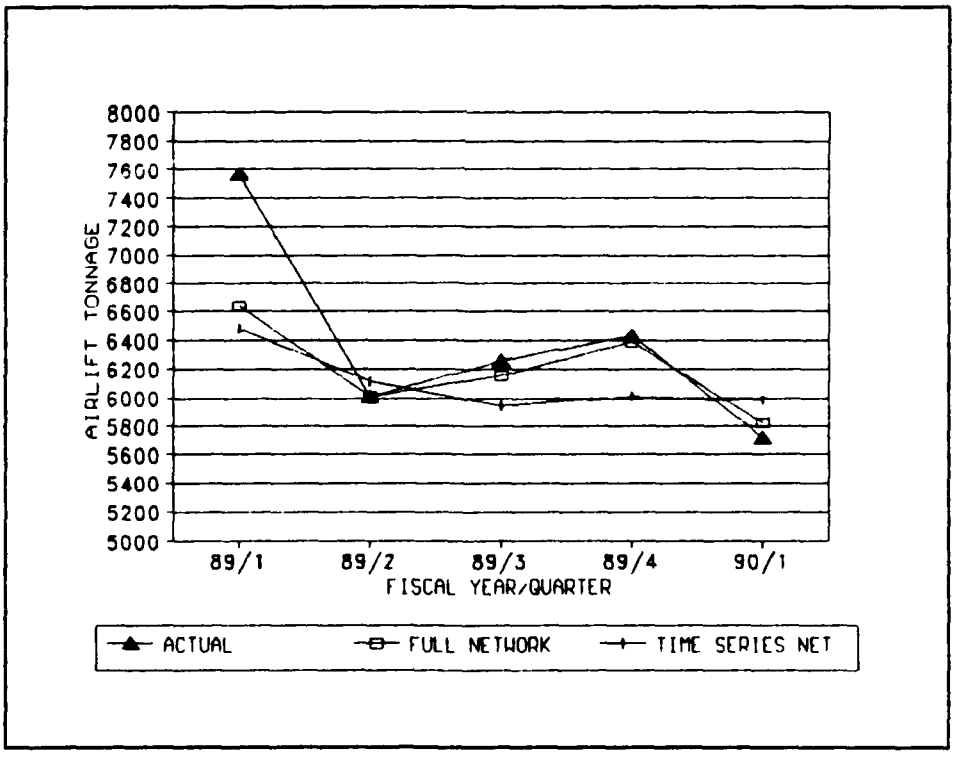

Figure 52. USAFE MAC Time Series Network Forecasts

Chapter Summary

This chapter started with an examination of the USAFE and PACAF MAC simple regression forecasting models presently used by DSXR. Multiple regression models and multivariable (full and reduced) neural networks using flying hours by aircraft type and military population variables were formulated and tested for forecasting accuracy. The last section presented the development and testing of time series neural network forecasting models. The results of this analysis are further discussed in Chapter VI.

VI. Research Conclusions and Findings

This chapter is divided into three parts. The first part presents the research conclusions and addresses the two research objectives. The second part presents other research findings and part three addresses areas for future study.

Research Conclusions

Research Objective 1. The first objective of this research was to develop multiple regression and neural network models that were statistically more accurate and reliable than the models currently used by DSXR. Figure 53 shows the rank order (1 = best performer, 5 = worst performer) of the full multivariable network, reduced multivariable network, multiple regression, DSXR simple regression, and the 12 quarter average in terms of MAE, minimum absolute error, and maximum absolute error for the PACAF and USAFE MSC tonnage data sets. The average rank for each model was determined by summing the rank order for each category and dividing by three. For the PACAF MSC data set, the full multivariable network outperformed all other models in every category. The reduced multivariable network and multiple regression model suffered from extrapolation problems and consequently overestimated the forecasts for FY 1989/2 to 1989/4. Since the PACAF MSC data was relatively stable after 1986, the simple 12 quarter average provided accurate forecasts in terms of MAE, but achieved the largest maximum absolute error out of the five models. The DSXR simple regression model also performed fairly well, but its maximum absolute error was almost three times the full multivariable network maximum absolute error. The full multivariable

network was the only network that outperformed the DSXR model in every category.

	MSC							
	PACAF				USAFE			
MODEL	MAE	MIN ERR	MAX ERR	AVE RANK	MAE	MIN ERR	MAX ERR	AVE RANK
Full Multivariable Network	1	1	1	1.0	3	3	1	2.3
Reduced Multivariable Network	5	4	3	4.0	1	1	4	2.0
Multiple Regression	4	5	2	3.7	2	2	3	2.3
DSXR Simple Regression	3	2	4	3.0	5	5	2	4.0
12 Quarter Average	2	3	5	3.3	4	4	5	4.3

Figure 53. PACAF and USAFE MSC Model Rank Order

For the USAFE MSC data set, the reduced multivariable network outperformed all other models in MAE and minimum absolute error, while the full multivariable network achieved the lowest maximum absolute error. The multivariable networks (reduced and full), the multiple regression model, and the 12 quarter average outperformed the DSXR model in MAE and minimum absolute error. Like the PACAF MSC results, the full multivariable network was the only network that outperformed the DSXR model in every category.

Figure 54 shows the rank order of the full multivariable network, reduced multivariable network, multiple regression, time series network, and the DSXR simple regression in terms of MAE, minimum absolute error, and maximum absolute error for the PACAF and USAFE MAC tonnage data sets. For the PACAF MAC data set, the time series network achieved the lowest MAE, but all three networks achieved comparable MAE's and minimum and

maximum absolute errors. All the models outperformed the DSXR model in every category.

Similar to the PACAF MSC results, the full multivariable network outperformed all other models in every category for the USAFE MAC data set. The time series network also performed very well while the reduced multivariable network and the multiple regression model achieved similar forecasting accuracy.

	MAC							
	PACAF				USAFE			
MODEL	MAE	MIN ERR	MAX ERR	AVE RANK	MAE	MIN ERR	MAX ERR	AVE RANK
Full Multivariable Network	2	1	3	2.0	1	1	1	1.0
Reduced Multivariable Network	3	3	2	2.7	4	4	2	3.3
Time Series Network	1	2	1	1.3	2	2	3	2.3
Multiple Regression	4	4	4	4.0	3	3	4	3.3
DSXR Simple Regression	5	5	5	5.0	5	5	5	5.0

Figure 54. PACAF and USAFE MAC Model Rank Order

Research Objective 2. The second objective of this research was to determine whether the multiple regression models or the neural network models were statistically more accurate. Compared to the full multivariable networks, the multiple regression models were less accurate in terms of MAE, minimum absolute error, and maximum absolute error for all the data sets except for the USAFE MSC data set. The multiple regression model had a lower MAE (4113) and minimum absolute error (652) for the USAFE MSC data set compared to the full multivariable network (MAE = 4940, minimum absolute error = 1474).

In comparison to the reduced multivariable networks, the multiple regression models performed with similar forecasting accuracy for the PACAF and USAFE MSC data sets and the USAFE MAC data set. For the PACAF data set, the reduced multivariable network significantly outperformed the multivariable network in every category (71% reduction in MAE, 88% reduction in minimum absolute error, 54% reduction in maximum absolute error). Overall, the multiple regression and the reduced multivariable networks achieved relatively comparable forecasting accuracy.

Findings

Multivariable Model Approach. The multivariable model development approach was based on the breakout of the total flying hour variable and the addition of the military population variables. The approach appears to increase forecasting accuracy compared to DSXR's simple regression technique. As illustrated in Figures 51 and 52, the full multivariable network outperformed the DSXR simple linear regression model in every category and for all the data sets. The reduced multivariable network and the multiple regression model outperformed the DSXR simple regression model in every category of the PACAF and USAFE MAC data sets and in MAE and minimum absolute error for the USAFE MSC data set.

A potential problem with the multivariable approach was discovered with the PACAF MSC data set. The DSXR model outperformed the multiple regression and reduced multivariable network in MAE (33% reduction compared to the multiple regression MAE and a 22% reduction compared to the reduced multivariable network MAE) and minimum absolute error. The reduced multivariable network and the multiple regression model suffered from extrapolation problems and significantly overestimated three out of

the six forecast periods (three of six forecasts were based on officer population values which were outside the data set that was used to develop the models).

There seems to be a tradeoff between a small data set and the extrapolation problem. In order to model the current relationships between SDT tonnage and the aircraft flying hours and military populations, the data sets that were used to develop the models were small (14 quarters). With a small data set, the potential of forecasting with variables outside of the data set (extrapolation) is significant. When extrapolation is encountered, the forecasts may have to be subjectively altered or the data set may have to be increased (if possible) so that no variables fall outside of the data set that was used to develop the model.

Unlike the reduced network and multiple regression model, the PACAF MSC full multivariable network did not suffer from extrapolation problems. The PACAF MSC full multivariable network used nine independent variables in contrast to the three variables used for the reduced network and the multiple regression model. This means the full multivariable network forecasts were based on six additional independent variables which seemed to diminish the extrapolation effect so that forecasting accuracy was not degraded with two of nine independent variables falling outside the data set.

Aircraft Flying Hour Variables. Table 38 shows the aircraft flying hour variables that were used in the multiple regression models. The '+' and '-' sign indicate the direction of the variable coefficient. The DSXR personnel believed the relationship between flying hours and tonnage requirements were not always positively correlated and that the F-16

aircraft seemed to require less logistical support compared to other aircraft. The apriori expectations regarding the direction of the relationships between flying hours by type of aircraft and tonnage were dependent on the type of aircraft.

Table 38

Multiple Regression Aircraft Variables

PACAF MSC	PACAF MAC	USAFE MSC	USAFE MAC
+ A-10	+ B-52	+ C-130	+ A-10
- F-16	- F-15	+ F-4	

The results (Table 38) show the older aircraft (F-4, C-130, A-10, and the B-52) with positive coefficients which indicate SDT tonnage requirements have been increasing (decreasing) with increased (decreased) flying hours. The relatively newer aircraft (F-15 and F-16) have negative coefficients which indicate SDT tonnage requirements have been decreasing (increasing) with increased (decreased) flying hours. Figures 55 and 56 show the percentage of the total aircraft flying hours that were flown by the F-16, F-15, F-4, and A-10 in PACAF and USAFE respectively. In PACAF and USAFE, the percentage of F-16 flying hours has been increasing since 1982 and has exceeded all other aircraft, but SDT requirements have not proportionately increased. Contrary to the F-16 flying hours, the percentage of F-4 flying hours has been decreasing since 1982. The general trend in PACAF and USAFE SDT requirements is inversely related to the increasing trend in F-16 flying hours, but directly related to the phaseout of the F-4 aircraft. The direction of the relationship between aircraft flying hours depends on the time period

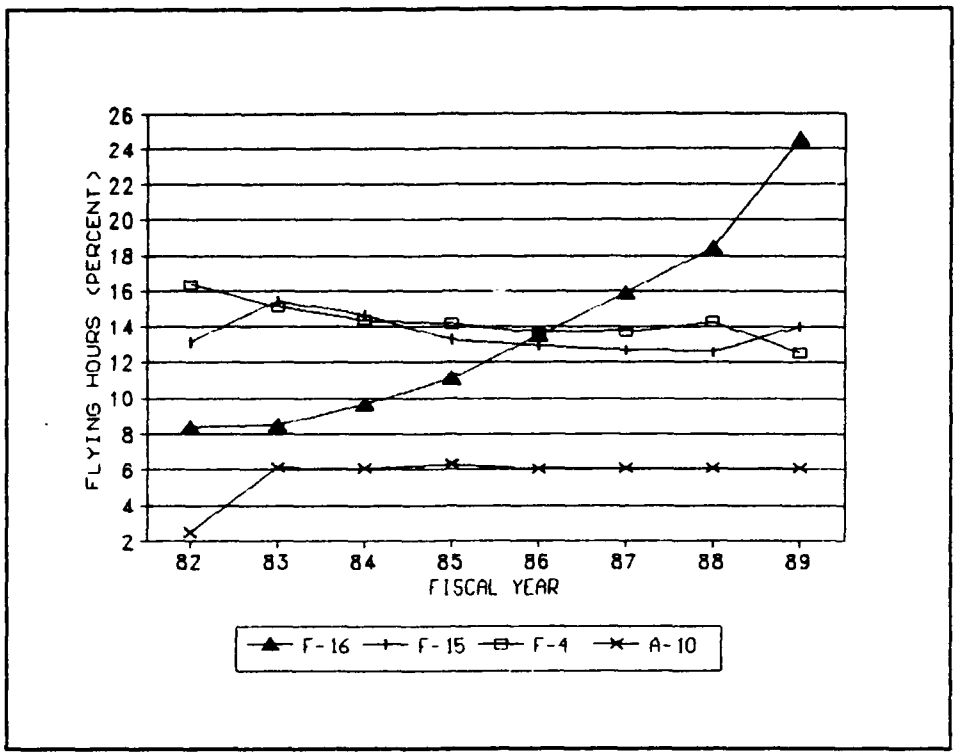

Figure 55. PACAF Aircraft Flying Hours by Percent

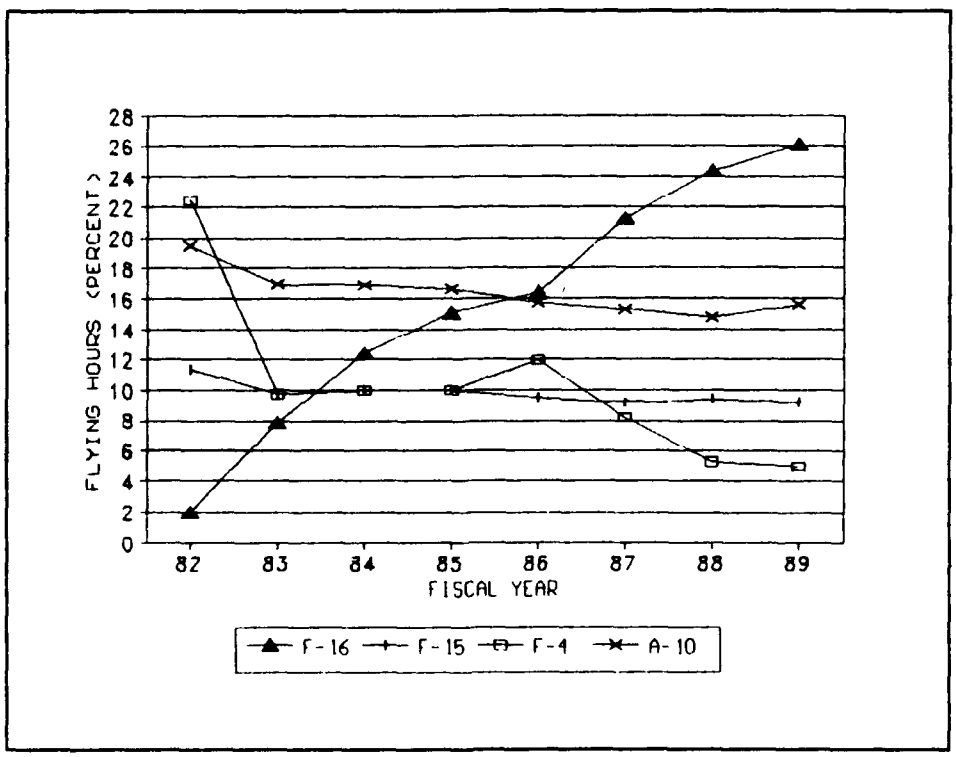

Figure 56. USAFE Aircraft Flying Hours by Percent

which is directly related to the stage of the aircraft's life cycle. With the completion of the phaseout of the F-4, the relationship between tonnage and F-16 and F-15 flying hours should change.

Military Population Variables. The military population variables significantly contributed to improving forecasting accuracy. Every multiple regression model that was developed contained at least one military population variable that was statistically significant. Table 39 shows the military population variables (officer manpower (OFF) and airman manpower (AMN)) that were used in the multiple regression models. A 'Yes' indicates the variable was used in the multiple regression model, while a 'No' indicates the variable was not used. The '+' and '-' sign indicate the direction of the variable coefficient.

Table 39

Multiple Regression Military Population Variables

	PACAF		USAFE	
	MSC	MAC	MSC	MAC
OFF	- YES	NO	- YES	+ YES
AMN	NO	- YES	NO	- YES

The apriori relationship between tonnage and military population was expected to be positively correlated. An increased (decreased) military population should require increased (decreased) SDT tonnage requirements, but this was not the case. For the PACAF and USAFE MSC data sets, the coefficients for the officer population variables were negative. From 1981 to 1986, the PACAF and USAFE officer manpower strength had been increasing, but since 1987, the officer strength appears to be

increasing, but since 1987, the officer strength appears to be decreasing. The residual analysis indicated a nonlinear relationship between tonnage and officer population. Second order regression models were developed by adding a squared officer population variable which improved the fit compared to the first order regression models, but did not improve forecasting accuracy. The second order relationship between tonnage and officer population was a positive nonlinear relationship (officer variable (OFF) was positive while the squared officer variable (SOFF) was negative). An increased (decreased) officer population required an increased (decreased) SDT tonnage requirements, but at a nonlinear decreasing (increasing) rate.

For the PACAF and USAFE MAC data sets, the coefficients for the airman population variables were also negative. Unlike the PACAF and USAFE MSC multiple regression models, these models did not require higher order terms to improve fit.

The military population variables could be measuring some other phenomena or interacting with other variables that affect SDT tonnage requirements. For example, the phaseout of the F-4 can be measured by a reduction in manpower requirements, but SDT requirements will increase or decrease depending on the activity level of the other remaining aircraft. Another example is the two level maintenance concept which is based on reducing manpower requirements. When it is initiated, manpower requirements should decrease, but SDT requirements will probably increase. The military manpower variables may change their direction of relationship with SDT tonnage depending on the fiscal year, manpower buildups and reductions and other phenomena, but the variables appear to contribute to improving forecasting accuracy.

Financial Implications. The objective for improving SDT forecasting accuracy is to minimize the financial implications resulting from SDT overestimations and underestimations. Overestimations divert scarce funds from other programs while underestimations degrade logistical support capability.

Figure 40 shows the cost per measurement ton (MSC) and cost per short ton (MAC) for fiscal years 1988, 1989, and 1990. These values were used to calculate the over/underestimations for each of the models in dollars (Figure 41). For the MSC data, the total over/underestimations for each model is the six quarter forecast period (FY 1988/3 to 1989/4) based on FY 1988 and 1989 cost per measurement ton values. For the MAC data, the total over/underestimations for each model is the five quarter forecast period (FY 1989/1 to 1990/1) based on FY 1989 and 1990 cost per ton values. The 'Difference' entry for each model shows the total reduction in over/underestimations compared to the DSXR model. The PACAF MSC reduced multivariable network and multiple regression model were the only models that did not reduce the total over/underestimations (figures are in parentheses in Table 41) compared to the DSXR model because of the overestimations resulting from the extrapolation problems.

Table 42 shows the total difference in over/underestimations for each model compared to the DSXR model. The largest difference was realized by the full multivariable network model and most of the difference came from the MAC data ($35.7 million) compared to the MSC data ($1.3 million). The MAC difference was large because the DSXR MAC models suffered from a lack of an intervention term and substantially overestimated the forecasts. The multiple regression models also achieved substantial difference ($24.6 million).

Table 40

Overseas SDT Costs (Dollars)

Fiscal Year	PACAF MSC*	PACAF MAC**	USAFE MSC*	USAFE MAC**
1988	45.43	2,208.00	51.64	1,483.00
1989	73.12	2,519.00	87.13	1,547.00
1990	82.19	2,463.00	91.96	1,694.00

* Cost per measurement ton ** Cost per short ton

Table 41

Financial Implications (Dollars)

	PACAF MSC	PACAF MAC	USAFE MSC	USAFE MAC
DSXR Models				
Underestimations:	903,672	0	2,234,247	982,345
Overestimations:	44,238	19,146,234	1,260,335	19,990,054
Total:	947,909	19,146,234	3,494,582	20,972,399
Full Networks				
Underestimations:	639,339	2,351,501	696,006	1,664,572
Overestimations:	219,509	123,431	1,522,380	186,858
Total:	858,847	2,474,932	2,218,386	1,851,430
Difference:	89,062	16,671,302	1,276,197	19,120,969
Reduced Networks				
Underestimations:	358,655	3,016,294	1,501,847	1,656,837
Overestimations:	1,462,006	118,393	391,305	4,195,436
Total:	1,820,661	3,134,687	1,893,152	5,852,273
Difference:	(872,752)	16,011,547	1,601,430	15,120,126
Time Series Networks				
Underestimations:	na	1,480,435	na	2,817,087
Overestimations:	na	199,001	na	619,080
Total:	na	1,679 436	na	3,436,167
Difference:	na	17,466,798	na	17,536,232
Multiple Regressions				
Underestimations:	393,658	10,965,849	1,254,883	1,811,537
Overestimations:	1,322,242	0	836,288	3,299,583
Total:	1,715,900	10,965,849	2,091,171	5,111,120
Difference:	(767,991)	8,180,385	1,403,411	15,861,279

Table 42

Total Difference in Underestimations and Overestimations (Dollars)

	MAC	MSC	Total
Full Network Model:	35,792,271	1,365,258	37,157,529
Time Series Model:	35,003,030	na	35,003,030
Reduced Network Model:	31,131,673	728,678	31,860,351
Multiple Regression Model:	24,041,664	635,420	24,677,084

Programmed versus Actual Flying Hours. DSXR produces forecasts with the simple regression models that were developed from historical tonnage and flying hour data by computing future tonnage requirements with programmed aircraft flying hour data. The forecasting accuracy of any model using the total flying hour variable or the flying hours by type of aircraft is not altered if the difference between programmed and actual flying hours is negligible and the model accurately describes the relationship between SDT tonnage and flying hours. Table 43 shows the PACAF and USAFE programmed and actual flying hours from FY 1988/3 to FY 1990/1. In 11 out of 14 cases, the programmed flying hours were overestimated. The financial overestimations and underestimations of the DSXR PACAF and USAFE MSC and MAC models using actual flying hours were compared to the same models using programmed flying hours (Table 44).

Table 44 shows how the use of the programmed flying hours actually improved forecasting accuracy for the MSC data ($708,244 savings), but significantly degraded DSXR MAC forecasting accuracy ($15,346,060 loss in over/underestimations). The forecasting accuracy of the multivariable models (full and reduced multivariable networks and the multiple regressions) will be altered to some extent if the programmed flying

hours are different from the actual flying hours. The time series networks and the 12 quarter averages are not affected by the discrepancies between programmed and actual flying hours.

Table 43

Programmed and Actual Flying Hours

	PACAF			USAFE		
FY/Qtr	Programmed Flying Hours	Actual Flying Hours	Difference	Programmed Flying Hours	Actual Flying Hours	Difference
1988/3	46252	41724	4528	79529	74331	5198
4	38962	38897	65	80236	74752	5484
1989/1	42735	40676	2059	81012	68136	12876
2	42727	41960	767	81105	74136	6969
3	42183	42353	-170	81489	85631	-4142
4	42555	36584	5971	82321	83012	-691
1990/1	39677	36672	3005	78407	68417	9990

Table 44

Programmed Flying Hours and Actual Flying Hours
Financial Implications (Dollars)

	PACAF		USAFE	
	MSC	MAC	MSC	MAC
DSXR Model with Actual Flying Hours:				
Underestimations:	903,672	0	2,234,247	982,345
Overestimations:	44,238	19,146,234	1,260,335	19,990,054
Total:	947,909	19,146,234	3,494,582	20,972,399
DSXR Model with Programmed Flying Hours:				
Underestimations:	346,853	0	1,130,556	0
Overestimations:	438,777	27,163,362	1,818,061	28,301,331
Total:	785,630	27,163,362	2,948,618	28,301,331
Difference:	(162,279)	8,017,128	(545,965)	7,328,932

Neural Network Findings. The neural network models have some advantages over conventional forecasting techniques. The number of

network training iterations can determine the strength of the relationship between the independent variables (inputs) and dependent variables (outputs). Networks can be thought of as adaptable reactive systems in which the number of training iterations determine how much of the variation in the dependent variable is modeled or explained by the independent variables. The problem is determining the optimum number of training iterations for a particular application. This is the same as the problems of undertraining and overtraining a network. Forecasting accuracy can be degraded when a network's R^2 value is too high (overtraining) or too low (undertraining). In this research, all of the networks made rapid progress (achieved significant increases in R^2 value) between 0 and 2000 training iterations. The R^2 value began to plateau after 1500 - 2000 training iterations and training beyond this point resulted in small increases in the R^2 value. A training heuristic that was used in this research was to continue training the networks beyond the 1500 - 2000 iteration range until the R^2 value increased at least another .05. All the networks (except for the time series networks and the USAFE MAC full multivariable network) required an additional 2000 - 2500 training iterations and training was terminated at 4000 iterations. The time series networks and the USAFE MAC full multivariable network were 'slower' in adjusting their weights compared to the other networks and required significantly more training iterations to increase the R^2 value.

Unlike regression analysis which fits one line or hyperplane to the data that minimizes the sum of squared error, the networks can produce an unlimited number of lines or hyperplanes which can be controlled through the number of hidden nodes the network uses. Depending on the number of

hidden nodes and the number of training iterations, networks can achieve near 100% learning convergence which means the change in the dependent variables is completely explained or caused by a change in the independent variables. This is similar to a regression model achieving an R^2 value of 1. This would be useful for problems where all the variables are known and the data is not distorted or noisy (no unexplained or random error).

Usually the opposite is true, the problem lacks many of the required variables or the relationship between the variables is not well defined. This type of data is distorted and noisy (large amount of unexplained error and random error). In this case, it is important not to model the noise or distortions in the dependent variable with the independent variables. A model that shapes a line or hyperplane that minimizes the sum of squared error between the variables may not be appropriate for this type of data. By varying the form and size of a network, a researcher can make the network less reactive to the data. This is similar to the principles of exponential smoothing where the α value determines the degree of smoothing. Low α values give considerable smoothing while high α values make the model more reactive to the historical data. For a network, a small number of hidden nodes causes the network to filter out or smooth most of the distortions while the prominent features or relationships are modeled or explained. Increasing the number of hidden nodes increases the degree to which the model extracts the mathematical relationships between the independent and dependent variables.

Future Research

Neural Networks. Although some experimentation was required to develop the networks used in this research, more experimentation is needed to find alternate network configurations and training algorithms that can improve forecasting capability.

The back-propagation network was used in this research, but other networks such as the Kohonen networks may be capable of producing more accurate forecasts. The Kohonen networks can be used as time series forecasting models similar to the back-propagation networks developed for the MAC data sets.

Combined Forecasting Techniques. Time series, simple averages, simple and multiple regressions, and neural network forecasting techniques all have particular strengths and weaknesses and a combination of the forecasts produced by several of these models may improve forecasting accuracy. Prior research from a variety of other applications has indicated the technique of combining forecasts increases forecasting accuracy, but experimentation may be required to find the correct combination of forecasting models (2:183-186). For example, a simple combined forecasting technique was developed by combining the forecasts based on a 20 quarter average of the ratio between tonnage and flying hours and a 14 quarter average of the tonnage. The forecasts produced by this combined technique were more accurate than the forecasts produced by each of the separate techniques and more accurate than the DSXR simple regression forecasts.

Inventory Models. DSXR has been traditionally using models that rely on the flying hour program, but some experimental multiple regression models were developed in this research that used aircraft

inventory levels. These models were similar to the multiple regression models using flying hours and had comparable forecasting accuracy. A combination of the two techniques may improve forecasting accuracy.

Appendix A: DSXR PACAF MSC Model SAS Regression Output

DEP VARIABLE: TON

ANALYSIS OF VARIANCE

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB>F
MODEL	1	339797123.86	339797123.86	9.511	0.0042
ERROR	32	1143223160	35725723.76		
C TOTAL	33	1483020284			

ROOT MSE	5977.1	R-SQUARE	0.2291
DEP MEAN	41325.85	ADJ R-SQ	0.2050
C.V.	14.46334		

PARAMETER ESTIMATES

VARIABLE	DF	PARAMETER ESTIMATE	STANDARD ERROR	T FOR H0: PARAMETER=0	PROB > |T|
INTERCEP	1	-3382.12	14532.78074	-0.233	0.8175
FH	1	1.13325080	0.36745715	3.084	0.0042

OBS	ACTUAL	PREDICT VALUE	STD ERR PREDICT	LOWER95% MEAN	UPPER95% MEAN	LOWER95% PREDICT	UPPER95 PREDICT
1	33145.0	36986.5	1740.8	33440.6	40532.5	24305.8	49667.3
2	30312.0	39004.9	1271.7	36414.6	41595.1	26557.5	51452.2
3	35918.0	37565.6	1592.9	34321.0	40810.2	24965.8	50165.4
4	32220.0	35634.6	2111.0	31334.7	39934.5	22722.7	48546.5
5	35198.0	37682.4	1564.1	34496.4	40868.3	25097.5	50267.2
6	30649.0	36612.6	1840.2	32864.2	40360.9	23873.7	49351.4
7	35193.0	37958.9	1497.6	34908.5	41009.3	25407.7	50510.1
8	35396.0	38228.6	1435.0	35305.5	41151.7	25707.7	50749.4
9	37343.0	37634.8	1575.8	34425.0	40844.6	25043.9	50225.6
10	41379.0	38556.1	1362.8	35780.1	41332.1	26068.7	51043.4
11	43392.0	40401.0	1068.0	38225.5	42576.5	28033.3	52768.7
12	42968.0	39153.3	1243.8	36619.8	41686.8	26717.6	51589.0
13	43039.0	38880.2	1296.0	36240.4	41520.0	26422.4	51338.0
14	49651.0	41583.0	1028.4	39488.1	43677.9	29229.2	53936.8
15	46352.0	40960.8	1031.9	38859.0	43062.7	28605.9	53315.8
16	35398.0	39247.4	1226.8	36748.5	41746.2	26818.7	51676.0
17	38462.0	41968.3	1046.0	39837.7	44099.0	29608.4	54328.2
18	41800.0	42551.9	1099.5	40312.4	44791.4	30172.8	54931.1
19	48352.0	41407.3	1025.4	39318.7	43496.0	29054.6	53760.1
20	49203.0	39947.7	1118.2	37670.0	42225.5	27561.6	52333.8
21	47567.0	42856.8	1138.9	40536.9	45176.7	30462.8	55250.7
22	49835.0	42886.2	1143.1	40557.8	45214.7	30490.7	55281.8
23	59435.0	43471.0	1238.8	40947.7	45994.3	31037.4	55904.6

OBS	ACTUAL	PREDICT VALUE	STD ERR PREDICT	LOWER95% MEAN	UPPER95% MEAN	LOWER95% PREDICT	UPPER95% PREDICT
24	48235.0	43061.9	1169.5	40679.8	45444.0	30656.2	55467.6
25	49040.0	45240.0	1631.4	41916.9	48563.1	32619.8	57860.2
26	40829.0	44148.7	1374.2	41349.5	46947.9	31656.2	56641.2
27	42134.0	43620.6	1266.6	41040.5	46200.6	31175.3	56065.8
28	34675.0	43502.7	1244.6	40967.6	46037.8	31066.7	55938.7
29	42681.0	46786.9	2046.0	42619.3	50954.5	33918.5	59655.3
30	39408.0	44579.3	1470.9	41583.1	47575.5	32041.2	57117.4
31	35796.0	46912.7	2081.4	42672.9	51152.4	34020.7	59804.6
32	39293.0	46142.1	1868.0	42337.0	49947.1	33386.5	58897.7
33	42387.0	43430.2	1231.4	40921.9	45938.5	30999.6	55860.8
34	48394.0	46474.1	1958.9	42483.9	50464.3	33662.0	59286.2
35	.	43901.6	1322.2	41208.3	46594.9	31432.4	56370.8
36	.	40697.9	1045.1	38569.2	42826.7	28338.4	53057.5
37	.	42714.0	1119.5	40433.6	44994.4	30327.4	55100.6
38	.	44169.1	1378.7	41360.9	46977.3	31674.5	56663.6
39	.	44614.4	1479.1	41601.6	47627.3	32072.3	57156.6
40	.	38076.7	1469.9	35082.6	41070.8	25539.1	50614.4
41	.	38175.3	1447.2	35227.5	41123.1	25648.7	50702.0

OBS	RESIDUAL	STD ERR RESIDUAL	STUDENT RESIDUAL	-2 -1 0 1 2	COOK'S D
1	-3841.5	5718.0	-.671835	| * | |	0.021
2	-8692.9	5840.3	-1.4884	| ** | |	0.053
3	-1647.6	5760.9	-0.286	| | |	0.003
4	-3414.6	5591.9	-.610626	| * | |	0.027
5	-2484.4	5768.8	-.430652	| | |	0.007
6	-5963.6	5686.8	-1.0487	| ** | |	0.058
7	-2765.9	5786.5	-0.47799	| | |	0.008
8	-2832.6	5802.3	-.488184	| | |	0.007
9	-291.76	5765.6	-.050602	| | |	0.000
10	2822.9	5819.7	0.4851	| | |	0.006
11	2991.0	5880.9	0.5086	| | * |	0.004
12	3814.7	5846.3	0.6525	| | * |	0.010
13	4158.8	5834.9	0.7127	| | * |	0.013
14	8068.0	5888.0	1.3703	| | ** |	0.029
15	5391.2	5887.4	0.9157	| | * |	0.013
16	-3849.4	5849.9	-.658029	| * | |	0.010
17	-3506.3	5884.9	-.595818	| * | |	0.006
18	-751.93	5875.1	-.127986	| | |	0.000
19	6944.7	5888.5	1.1794	| | ** |	0.021
20	9255.3	5871.6	1.5763	| | *** |	0.045
21	4710.2	5867.6	0.8028	| | * |	0.012
22	6948.8	5866.8	1.1844	| | ** |	0.027
23	15964.0	5847.3	2.7301	| | ***** |	0.167
24	5173.1	5861.6	0.8825	| | * |	0.016
25	3800.0	5750.1	0.6609	| | * |	0.018
26	-3319.7	5817.0	-.570689	| * | |	0.009

OBS	RESIDUAL	STD ERR RESIDUAL	STUDENT RESIDUAL	-2 -1 0 1 2	COOK'S D
27	-1486.6	5841.3	-.254494	¦ ¦ ¦	0.002
28	-8827.7	5846.1	-1.51	¦ ***¦ ¦	0.052
29	-4105.9	5616.0	-.731106	¦ *¦ ¦	0.035
30	-5171.3	5793.3	-.892641	¦ *¦ ¦	0.026
31	-11117	5603.0	-1.9841	¦ ***¦ ¦	0.272
32	-6849.1	5677.7	-1.2063	¦ **¦ ¦	0.079
33	-1043.2	5848.9	-.178359	¦ ¦ ¦	0.001
34	1919.9	5647.0	0.3400	¦ ¦ ¦	0.007
35	.	.	.		.
36	.	.	.		.
37	.	.	.		.
38	.	.	.		.
39	.	.	.		.
40	.	.	.		.
41	.	.	.		.

SUM OF RESIDUALS 2.00089E-11
SUM OF SQUARED RESIDUALS 1143223160
PREDICTED RESID SS (PRESS) 1288831249

DURBIN-WATSON D 0.655
(FOR NUMBER OF OBS.) 34
1ST ORDER AUTOCORRELATION 0.664

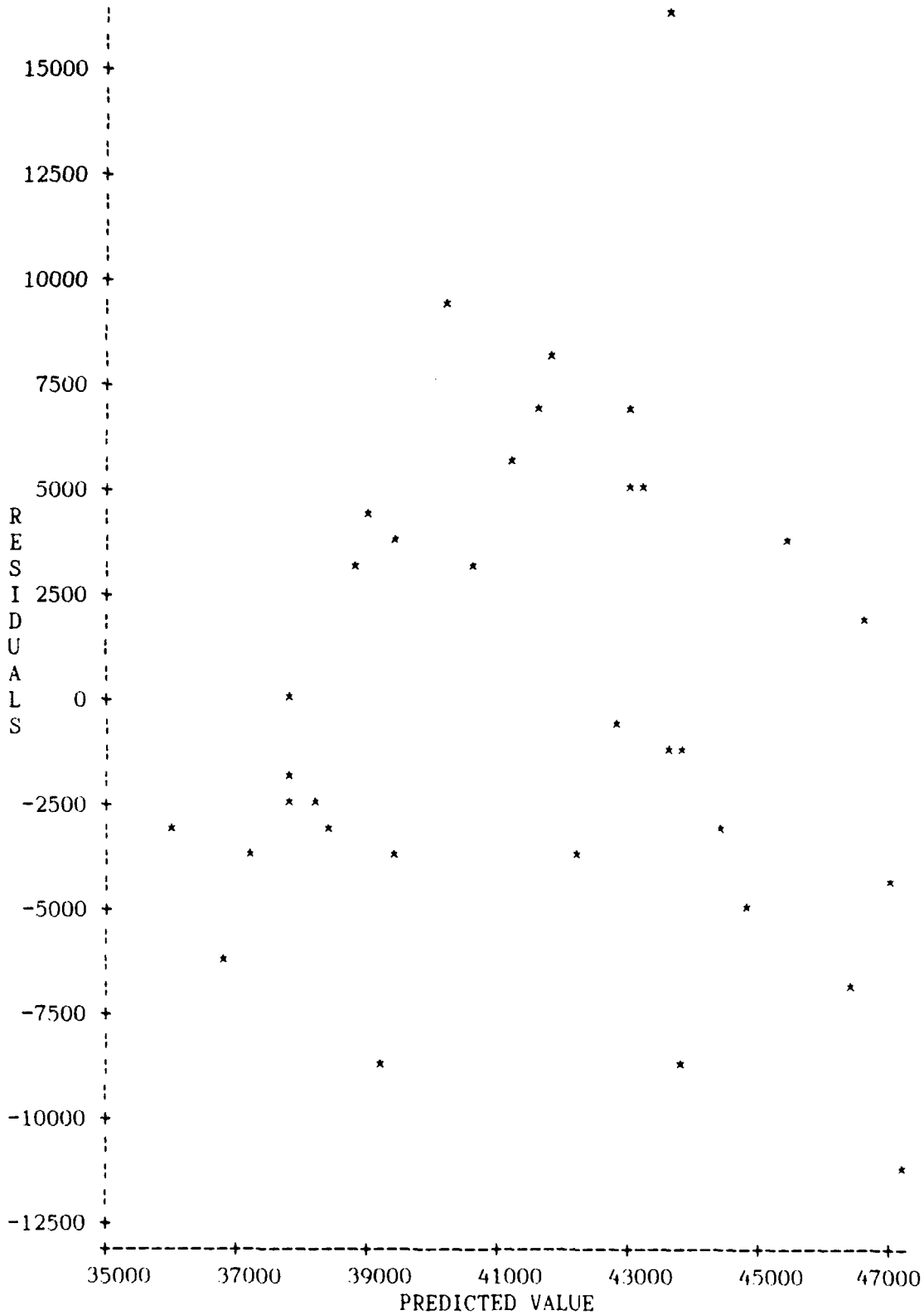

Figure 57. DSXR PACAF MSC Model Residuals vesus Predicted Values

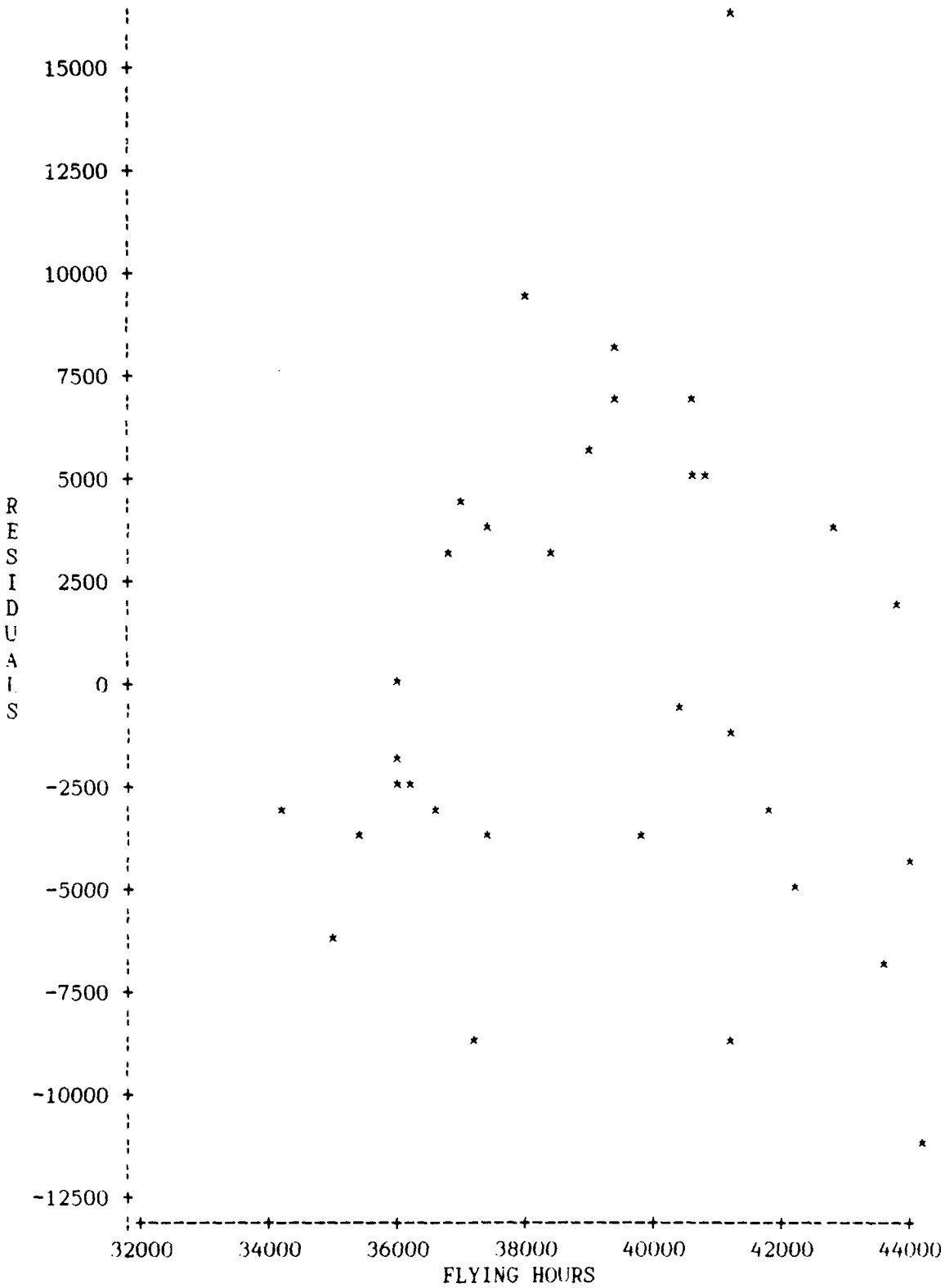

Figure 58. DSXR PACAF MSC Model Residuals versus Flying Hours

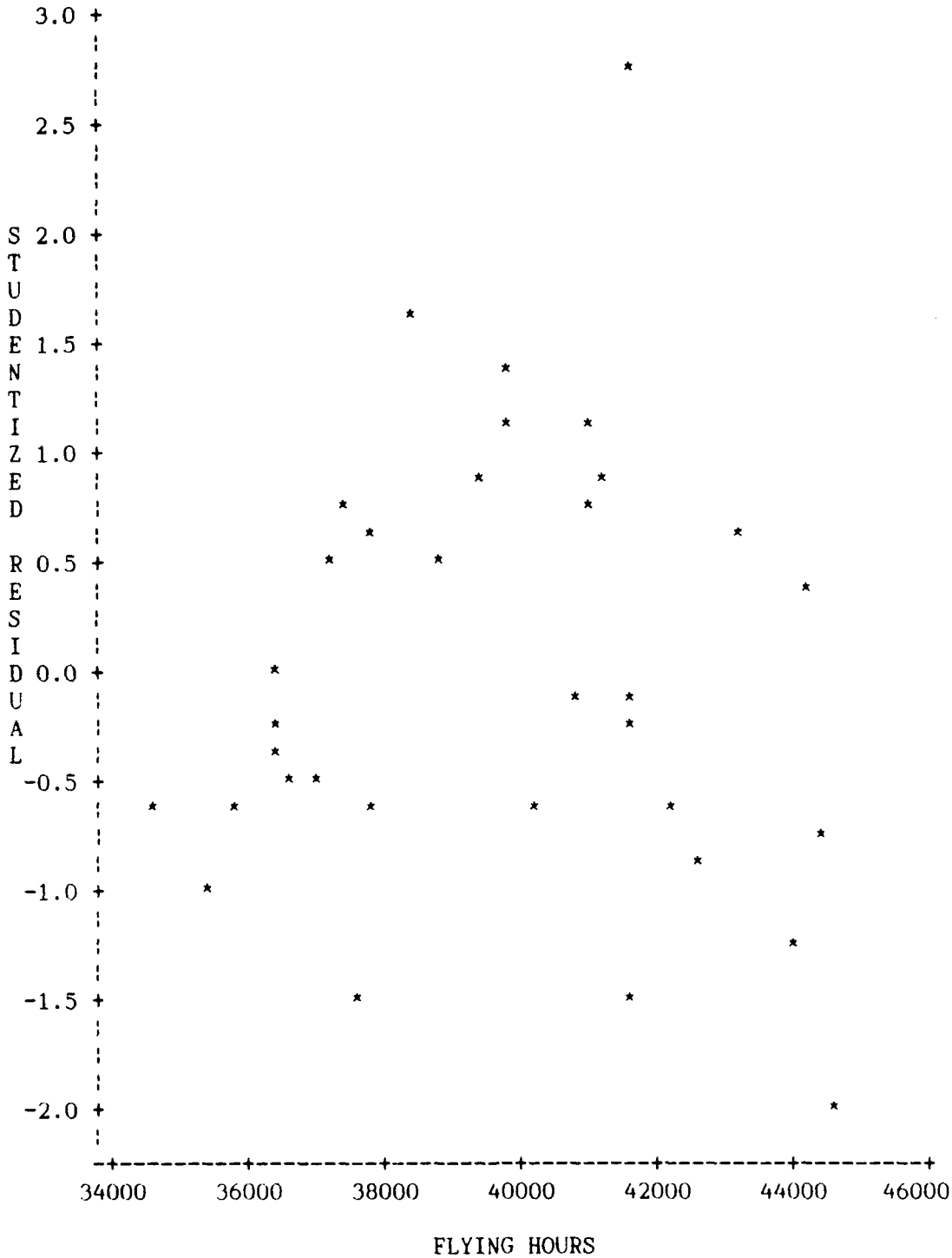

Figure 59. DSXR PACAF MSC Model Studentized Residuals versus Flying Hours

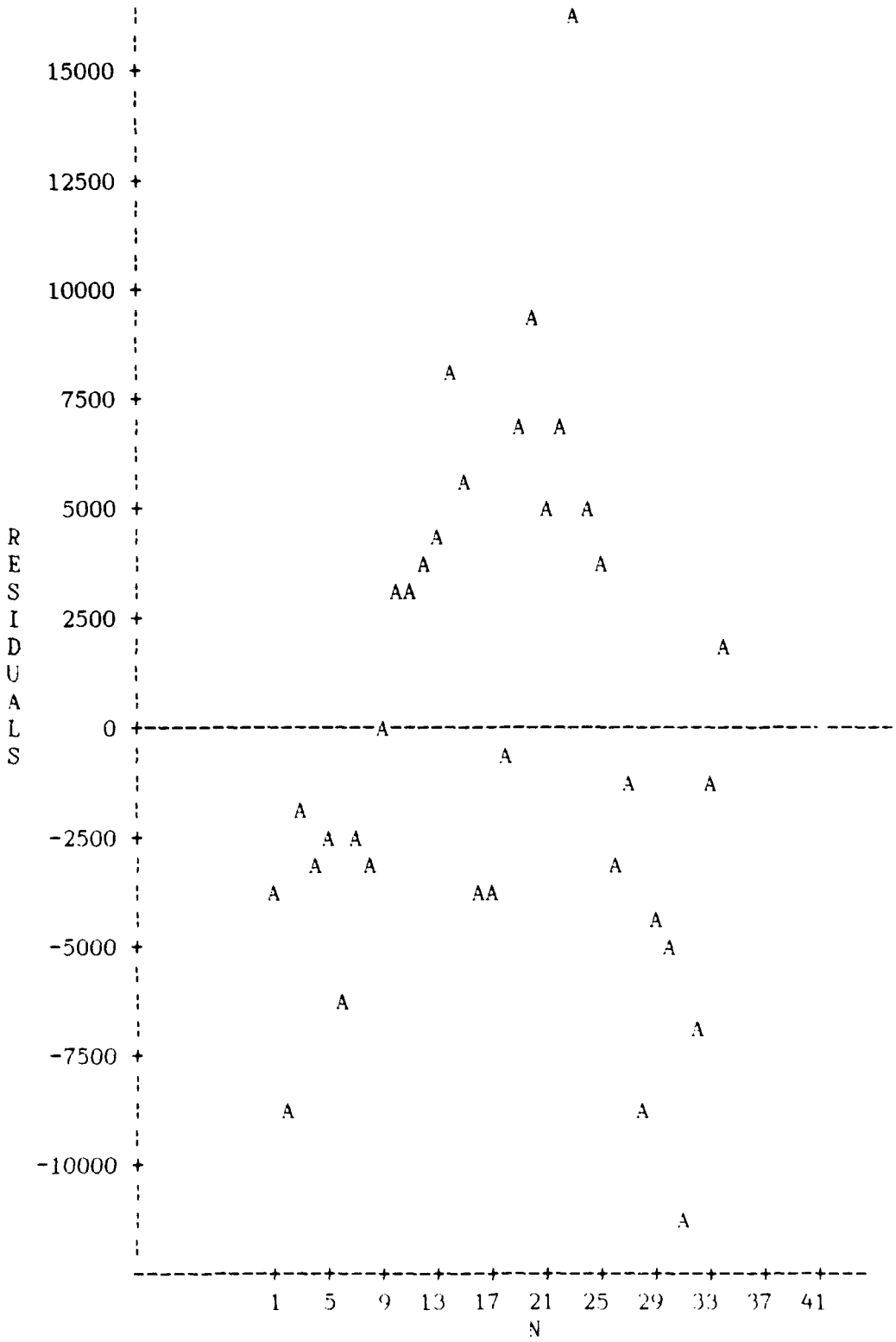

Figure 60. DSXR PACAF MSC Model Residuals versus Time (N)

UNIVARIATE

VARIABLE=RESIDUAL RESIDUALS

MOMENTS

N	34	SUM WGTS	34
MEAN	5.885E-13	SUM	2.001E-11
STD DEV	5885.84	VARIANCE	34643126
SKEWNESS	0.463256	KURTOSIS	0.239901
USS	1143223160	CSS	1143223160
CV	99.99	STD MEAN	1009.41
T:MEAN=0	5.830E-16	PROB>|T|	1
SGN RANK	-2.5	PROB>|S|	0.972723
NUM ^= 0	34		
W:NORMAL	0.97273	PROB<W	0.607

QUANTILES(DEF=4)

100% MAX	15964	99%	15964
75% Q3	4296.66	95%	10932.5
50% MED	-1264.9	90%	7508.38
25% Q1	-3843.5	10%	-7771
0% MIN	-11117	5%	-9400
		1%	-11117
RANGE	27080.7		
Q3-Q1	8140.15		
MODE	-11117		

EXTREMES

LOWEST	HIGHEST
-11117	6944.65
-8827.7	6948.76
-8692.9	8068
-6849.1	9255.28
-5963.6	15964

MISSING VALUE	.
COUNT	7
% COUNT/NOBS	17.07

Appendix B: DSXR USAFE MSC Model SAS Regression Output

DEP VARIABLE: TON

ANALYSIS OF VARIANCE

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB>F
MODEL	1	1200722483	1200722483	5.926	0.0207
ERROR	32	6484081098	202627534.30		
C TOTAL	33	7684803580			

ROOT MSE	14234.73	R-SQUARE	0.1562
DEP MEAN	70816.56	ADJ R-SQ	0.1299
C.V.	20.10085		

PARAMETER ESTIMATES

VARIABLE	DF	PARAMETER ESTIMATE	STANDARD ERROR	T FOR H0: PARAMETER=0	PROB > |T|
INTERCEP	1	5192.00148	27068.71949	0.192	0.8491
FH	1	0.89342350	0.36701624	2.434	0.0207

OBS	ACTUAL	PREDICT VALUE	STD ERR PREDICT	LOWER95% MEAN	UPPER95% MEAN	LOWER95% PREDICT	UPPER95% PREDICT
1	42971.0	59209.3	5356.8	48297.8	70120.7	28229.1	90189.4
2	50901.0	57732.5	5903.3	45707.8	69757.1	26342.9	89122.0
3	56271.0	67494.9	2796.7	61798.2	73191.6	37945.6	97044.2
4	49844.0	66006.4	3140.7	59609.0	72403.8	36314.1	95698.8
5	53377.0	61965.5	4379.5	53044.8	70886.2	31629.2	92301.8
6	53945.0	63789.9	3780.5	56089.4	71490.3	33789.7	93790.0
7	60785.0	72433.7	2530.0	67280.3	77587.2	42984.3	101883.1
8	54941.0	72624.9	2551.8	67427.2	77822.7	43167.7	102082.1
9	55855.0	60475.3	4899.7	50495.0	70455.5	29810.7	91139.8
10	57543.0	67754.0	2746.4	62159.9	73348.1	38224.3	97283.7
11	63076.0	77535.2	3684.7	70029.7	85040.7	47584.5	107485.8
12	65851.0	75205.1	3034.8	69023.6	81386.7	45558.5	104851.7
13	91436.0	64351.8	3607.3	57004.1	71699.5	34440.3	94263.3
14	93263.0	68174.8	2671.6	62733.0	73616.6	38673.5	97676.0
15	83737.0	74545.8	2882.1	68675.2	80416.4	44962.4	104129.1
16	83617.0	74829.0	2945.6	68829.1	80828.9	45219.7	104438.3
17	88315.0	67271.5	2842.6	61481.4	73061.7	37704.1	96839.0
18	86968.0	70495.0	2444.8	65515.1	75474.9	41075.5	99914.5
19	101701.0	75460.7	3098.3	69149.7	81771.6	45786.8	105134.5
20	85521.0	74612.8	2896.8	68712.2	80513.4	45023.5	104202.1
21	96200.0	66862.3	2932.3	60889.5	72835.2	37258.6	96466.1
22	71083.0	67838.0	2730.7	62275.7	73400.2	38314.3	97361.7
23	83702.0	79646.3	4372.3	70740.4	88552.3	49314.4	109978.3

OBS	ACTUAL	PREDICT VALUE	STD ERR PREDICT	LOWER95% MEAN	UPPER95% MEAN	LOWER95% PREDICT	UPPER95% PREDICT
24	77001.0	76539.0	3389.1	69635.8	83442.3	46733.6	106344.5
25	78732.0	72691.9	2559.9	67477.6	77906.3	43231.8	102152.1
26	80227.0	71673.4	2466.5	66649.4	76697.5	42246.4	101100.5
27	73784.0	77267.2	3603.0	69928.2	84606.2	47357.8	107176.5
28	56587.0	74546.7	2882.3	68675.7	80417.7	44963.3	104130.1
29	57977.0	72473.9	2534.4	67311.6	77636.3	43023.0	101924.9
30	64120.0	68953.8	2558.4	63742.7	74165.0	39494.3	98413.4
31	72994.0	83383.5	5710.6	71751.5	95015.5	52142.3	114624.8
32	74368.0	77853.2	3783.6	70146.4	85560.1	47851.5	107855.0
33	70603.0	72585.6	2547.1	67397.4	77773.9	43130.1	102041.1
34	70467.0	73479.9	2675.2	68030.8	78929.1	43977.3	102982.5
35	.	71601.1	2462.4	66585.3	76616.8	42175.4	101026.7
36	.	71977.2	2487.4	66910.6	77043.7	42542.9	101411.5
37	.	66066.3	3125.3	59700.3	72432.3	36380.7	95751.9
38	.	71426.8	2454.1	66428.1	76425.6	42004.1	100849.6
39	.	81696.7	5092.8	71323.1	92070.4	50901.9	112491.6
40	.	79356.9	4274.1	70650.8	88062.9	49083.0	109630.7
41	.	66022.5	3136.6	59633.6	72411.5	36332.0	95713.1

OBS	RESIDUAL	STD ERR RESIDUAL	STUDENT RESIDUAL	-2 -1 0 1 2	COOK'S D
1	-16238	13188.3	-1.2313	¦ **¦ ¦	0.125
2	-6831.5	12952.9	-.527406	¦ *¦ ¦	0.029
3	-11224	13957.3	-0.80416	¦ *¦ ¦	0.013
4	-16162	13883.9	-1.1641	¦ **¦ ¦	0.035
5	-8588.5	13544.3	-.634105	¦ *¦ ¦	0.021
6	-9844.9	13723.5	-0.71737	¦ *¦ ¦	0.020
7	-11649	14008.1	-.831572	¦ *¦ ¦	0.011
8	-17684	14004.1	-1.2628	¦ **¦ ¦	0.026
9	-4620.3	13364.9	-.345701	¦ ¦ ¦	0.008
10	-10211	13967.3	-.731064	¦ *¦ ¦	0.010
11	-14459	13749.6	-1.0516	¦ **¦ ¦	0.040
12	-9354.1	13907.5	-.672598	¦ *¦ ¦	0.011
13	27084.2	13770.1	1.9669	¦ ¦*** ¦	0.133
14	25088.2	13981.8	1.7944	¦ ¦*** ¦	0.059
15	9191.2	13939.9	0.6593	¦ ¦* ¦	0.009
16	8788.0	13926.6	0.6310	¦ ¦* ¦	0.009
17	21043.5	13948.0	1.5087	¦ ¦*** ¦	0.047
18	16473.0	14023.2	1.1747	¦ ¦** ¦	0.021
19	26240.3	13893.5	1.8887	¦ ¦*** ¦	0.089
20	10908.2	13936.9	0.7827	¦ ¦* ¦	0.013
21	29337.7	13929.4	2.1062	¦ ¦**** ¦	0.098
22	3245.0	13970.4	0.2323	¦ ¦ ¦	0.001
23	4055.7	13546.6	0.2994	¦ ¦ ¦	0.005
24	461.9848	13825.4	.0334157	¦ ¦ ¦	0.000
25	6040.1	14002.7	0.4314	¦ ¦ ¦	0.003
26	8553.6	14019.4	0.6101	¦ ¦* ¦	0.006

OBS	RESIDUAL	STD ERR RESIDUAL	STUDENT RESIDUAL	-2 -1 0 1 2	COOK'S D
27	-3483.2	13771.2	-0.25293		0.002
28	-17960	13939.9	-1.2884	**	0.035
29	-14497	14007.3	-1.035	**	0.018
30	-4833.8	14002.9	-.345202		0.002
31	-10390	13039.0	-.796802	*	0.061
32	-3485.2	13722.7	-.253977		0.002
33	-1982.6	14005.0	-.141565		0.000
34	-3012.9	13981.1	-.215501		0.001
35	.	.	.		.
36	.	.	.		.
37	.	.	.		.
38	.	.	.		.
39	.	.	.		.
40	.	.	.		.
41	.	.	.		.

SUM OF RESIDUALS -3.81988E-11
SUM OF SQUARED RESIDUALS 6484081098
PREDICTED RESID SS (PRESS) 7236550336

DURBIN-WATSON D 0.644
(FOR NUMBER OF OBS.) 34
1ST ORDER AUTOCORRELATION 0.657

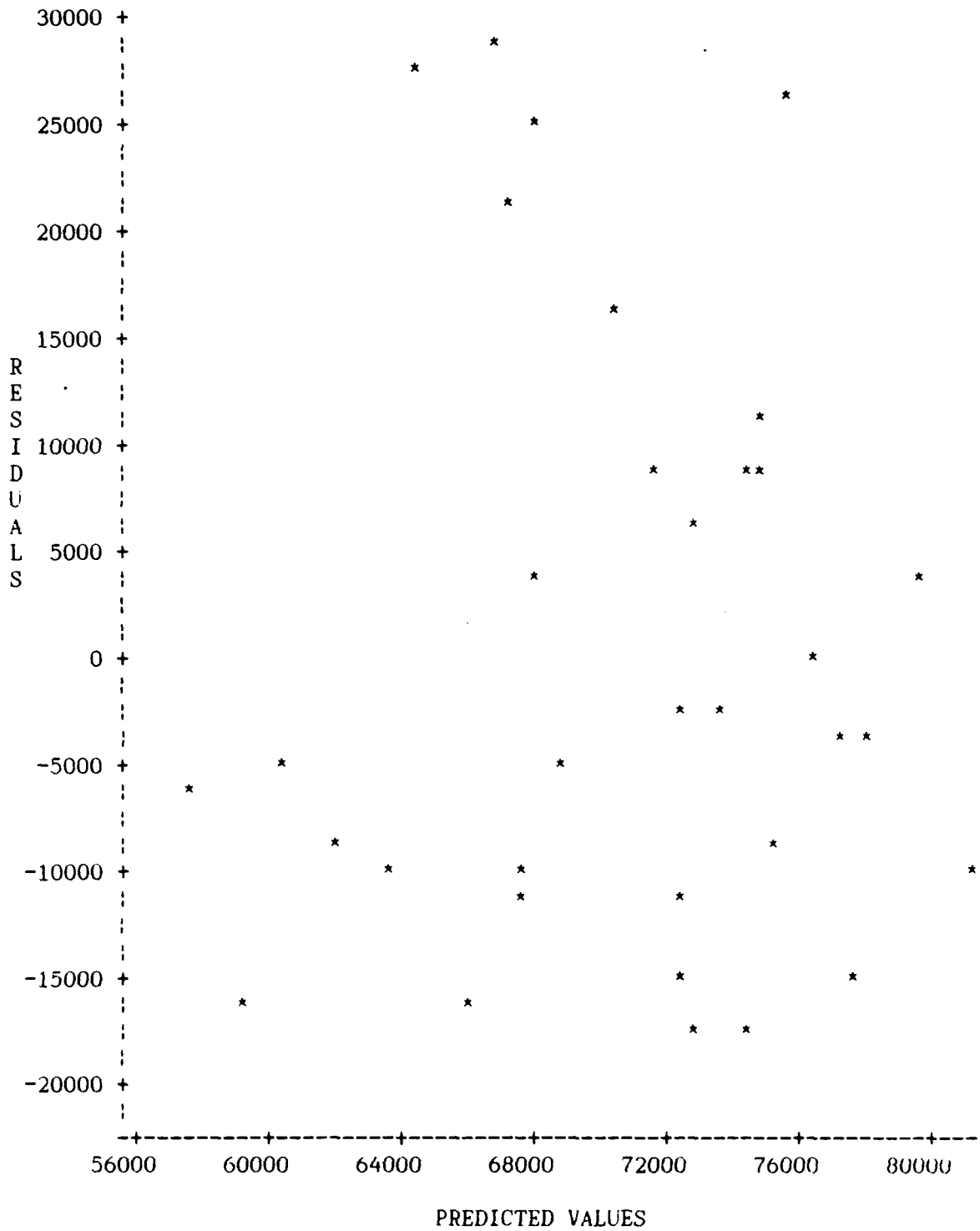

Figure 61. DSXR USAFE MSC Model Residuals versus Predicted Values

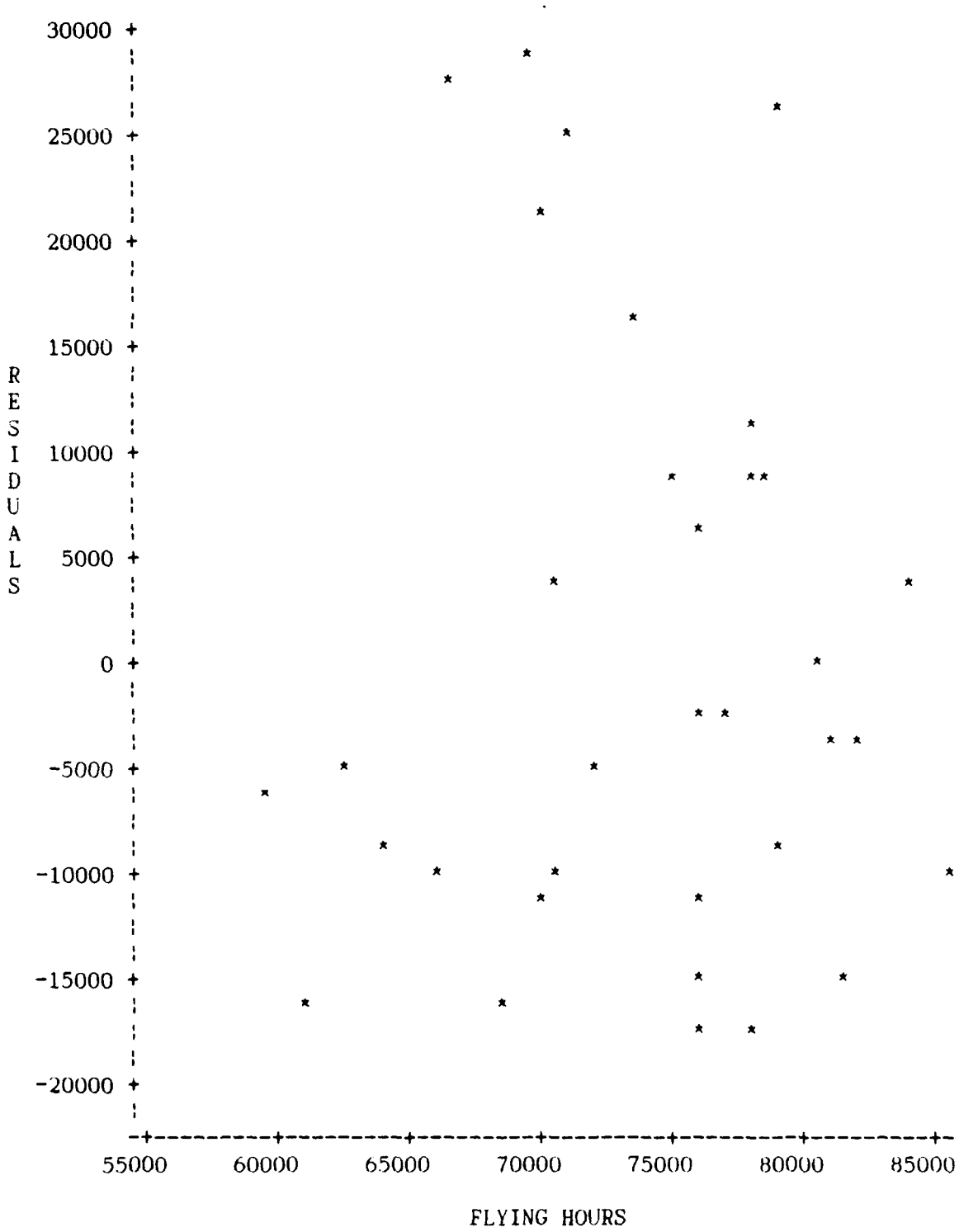

Figure 62. DSXR USAFE MSC Model Residuals versus Flying Hours

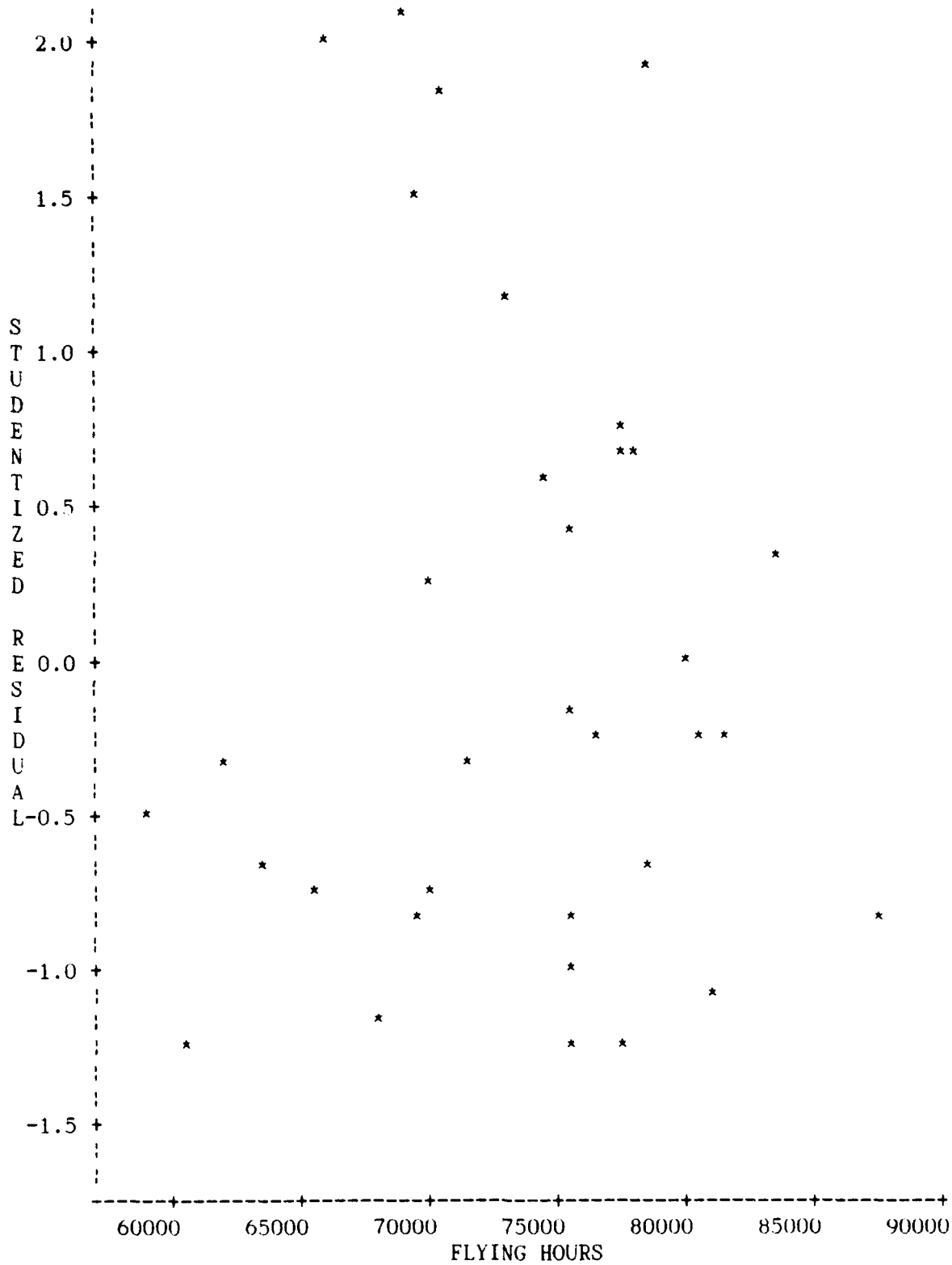

Figure 63. DSXR USAFE MSC Model Studentized Residuals versus Flying Hours

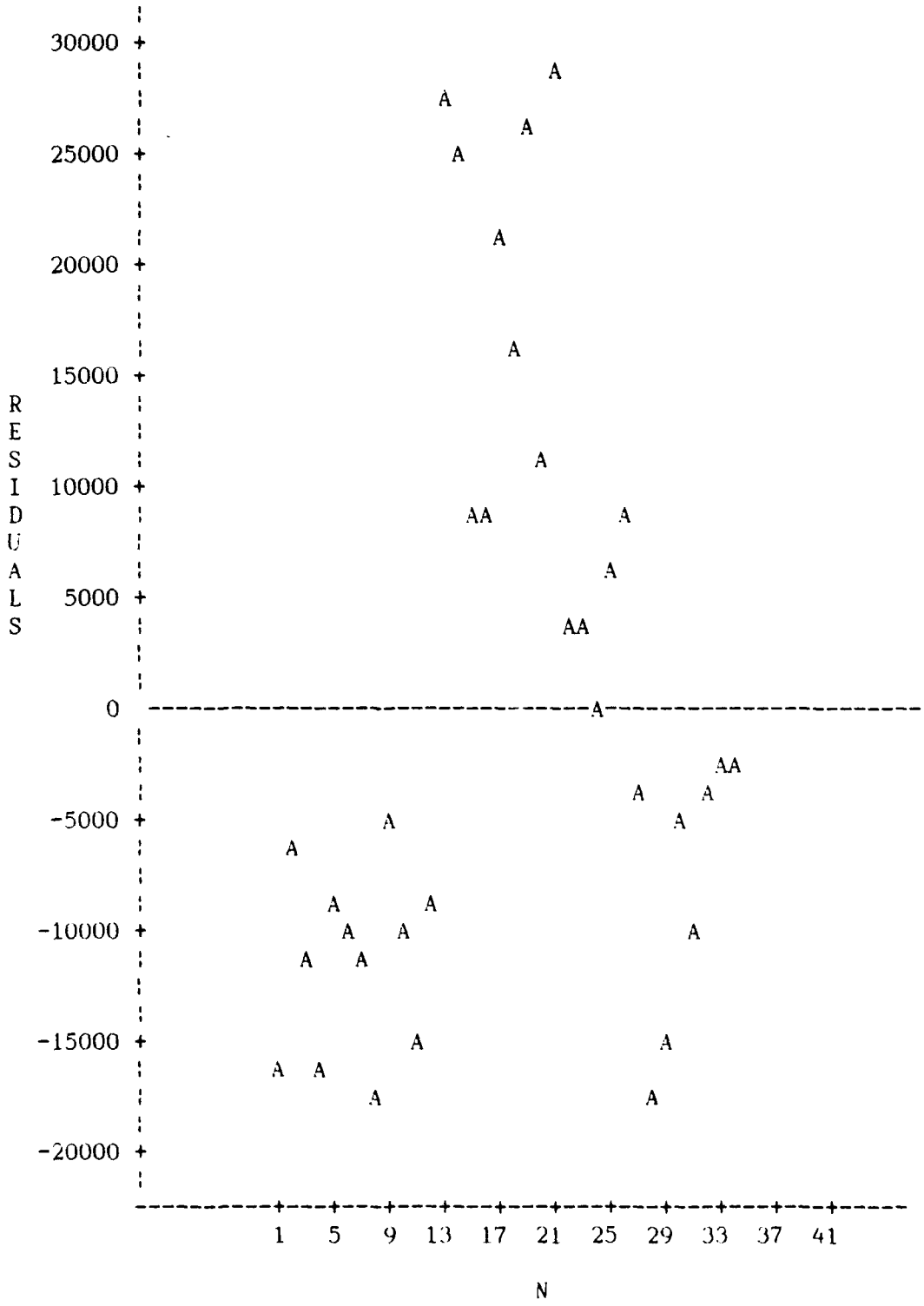

Figure 64. DSXR USAFE MSC Model Residuals versus Time (N)

UNIVARIATE

VARIABLE=RESIDUAL RESIDUALS

MOMENTS

N	34	SUM WGTS	34
MEAN	-1.123E-12	SUM	-3.820E-11
STD DEV	14017.4	VARIANCE	196487306
SKEWNESS	0.72867	KURTOSIS	-0.502932
USS	6484081098	CSS	6484081098
CV	-99999	STD MEAN	2403.96
T:MEAN=0	-4.674E-16	PROB>|T|	1
SGN RANK	-27.5	PROB>|S|	0.644364
NUM ^= 0	34		
W:NORMAL	0.907887	PROB<W	<.01

QUANTILES(DEF=4)

100% MAX	29337.7	99%	29337.7
75% Q3	8888.8	95%	27647.5
50% MED	-3484.2	90%	25664.3
25% Q1	-10598	10%	-16200
0% MIN	-17960	5%	-17753
		1%	-17960
RANGE	47297.3		
Q3-Q1	19486.9		
MODE	-17960		

EXTREMES

LOWEST	HIGHEST
-17960	21043.5
-17684	25088.2
-16238	26240.3
-16162	27084.2
-14497	29337.7

MISSING VALUE	.
COUNT	7
% COUNT/NOBS	17.07

Appendix C: Trend and Seasonal Analysis
(PACAF, USAFE MSC and Flying Hours)

TREND AND SEASONAL ANALYSIS
PACAF MSC TONNAGE

	Actual	DBQ	DBQ-1	DBQ-2
Variance:	49027789	51731495	44276024	1.02E+08
Index:	100%	106%	90%	209%
Trend:	None	None	Moderate	Strong
Seasonal:	No	Yes	Yes	Yes
			++++++++	

Year	Qtr	Actual Data	Diffs Between Same Qtr. Each Yr.	First Diffs Between Diffs	Second Diffs Between Diffs
1978	1	31163			
	2	30967			
	3	32924			
	4	32018			
1979	1	33469	2306		
	2	30548	-419	-2725	
	3	33046	122	541	3266
	4	34991	2973	2851	2310
1980	1	33145	-324	-3297	-6148
	2	30312	-236	88	3385
	3	35818	2772	3008	2920
	4	32220	-2771	-5543	-8551
1981	1	35198	2053	4824	10367
	2	30649	337	-1716	-6540
	3	35193	-625	-962	754
	4	35396	3176	3801	4763
1982	1	37343	2145	-1031	-4832
	2	41379	10730	8585	9616
	3	43392	8199	-2531	-11116
	4	42968	7572	-627	1904
1983	1	43039	5696	-1876	-1249
	2	49651	8272	2576	4452
	3	46352	2960	-5312	-7888
	4	35398	-7570	-10530	-5218
1984	1	38462	-4577	2993	13523
	2	41800	-7851	-3274	-6267
	3	48352	2000	9851	13125
	4	49203	13805	11805	1954
1985	1	47567	9105	-4700	-16505
	2	49835	8035	-1070	3630
	3	59435	11083	3048	4118
	4	48235	-968	-12051	-15099

1986	1	49040	1473	2441	14492
	2	40829	-9006	-10479	-12920
	3	42134	-17301	-8295	2184
	4	34675	-13560	3741	12036
1987	1	42681	-6359	7201	3460
	2	39408	-1421	4938	-2263
	3	35796	-6338	-4917	-9855
	4	39293	4618	10956	15873
1988	1	42387	-294	-4912	-15868
	2	48394	8986	9280	14192
	3	55113	19317	10331	1051
	4	42250	2957	-16360	-26691
1989	1	43079	692	-2265	14095
	2	45086	-3308	-4000	-1735
	3	44009	-11104	-7796	-3796
	4	41224	-1026	10078	17874

TREND AND SEASONAL ANALYSIS
PACAF FLYING HOURS

	Actual	DBQ	DBQ-1	DBQ-2
Variance:	8267635	2701221	4219453	12887688
Index:	100%	33%	51%	156%
Trend:	None	None	Moderate	Strong
Seasonal:	No	Yes ++++++++	Yes	Yes

Year	Qtr	Actual Data	Diffs Between Same Qtr. Each Yr.	First Diffs Between Diffs	Second Diffs Between Diffs
1978	1	35287			
	2	37575			
	3	34865			
	4	35922			
1979	1	35310	23		
	2	37593	18	-5	
	3	35600	735	717	722
	4	35615	-307	-1042	-1759
1980	1	35622	312	619	1661
	2	37403	-190	-502	-1121
	3	36133	533	723	1225
	4	34429	-1186	-1719	-2442
1981	1	36236	614	1800	3519
	2	35292	-2111	-2725	-4525
	3	36480	347	2458	5183
	4	36718	2289	1942	-516
1982	1	36194	-42	-2331	-4273
	2	37007	1715	1757	4088
	3	38635	2155	440	-1317
	4	37534	816	-1339	-1779
1983	1	37293	1099	283	1622
	2	39678	2671	1572	1289
	3	39129	494	-2177	-3749
	4	37617	83	-411	1766
1984	1	40018	2725	2642	3053
	2	40533	855	-1870	-4512
	3	39523	394	-461	1409
	4	38235	618	224	685
1985	1	40802	784	166	-58
	2	40828	295	-489	-655
	3	41344	1821	1526	2015
	4	40983	2748	927	-599
1986	1	42905	2103	-645	-1572
	2	41942	1114	-989	-344
	3	41476	132	-982	7
	4	41372	389	257	1239

1987	1	44270	1365	976	719
	2	42322	380	-985	-1961
	3	44381	2905	2525	3510
	4	43701	2329	-576	-3101
1988	1	41308	-2962	-5291	-4715
	2	43994	1672	4634	9925
	3	41724	-2657	-4329	-8963
	4	38897	-4804	-2147	2182
1989	1	40676	-632	4172	6319
	2	41960	-2034	-1402	-5574
	3	42353	629	2663	4065
	4	36584	-2313	-2942	-5605

TREND AND SEASONAL ANALYSIS
USAFE MSC TONNAGE

	Actual	DBQ	DBQ-1	DBQ-2
Variance:	2.84E+08	1.69E+08	1.45E+08	3.56E+08
Index:	100%	60%	51%	126%

Trend:	None	None	Moderate	Strong
Seasonal:	No	Yes	Yes	Yes
			++++++++	

Year	Qtr	Actual Data	Diffs Between Same Qtr. Each Yr.	First Diffs Between Diffs	Second Diffs Between Diffs
1978	1	46778			
	2	38450			
	3	47078			
	4	37450			
1979	1	48574	1796		
	2	44156	5706	3910	
	3	44041	-3037	-8743	-12653
	4	44296	6846	9883	18626
1980	1	42971	-5603	-12449	-22332
	2	50901	6745	12348	24797
	3	56271	12230	5485	-6863
	4	49844	5548	-6682	-12167
1981	1	53377	10406	4858	11540
	2	53945	3044	-7362	-12220
	3	60785	4514	1470	8832
	4	54941	5097	583	-887
1982	1	55855	2478	-2619	-3202
	2	57543	3598	1120	3739
	3	63076	2291	-1307	-2427
	4	65851	10910	8619	9926
1983	1	91436	35581	24671	16052
	2	93263	35720	139	-24532
	3	83737	20661	-15059	-15198
	4	83617	17766	-2895	12164
1984	1	88315	-3121	-20887	-17992
	2	86968	-6295	-3174	17713
	3	101701	17964	24259	27433
	4	85521	1904	-16060	-40319
1985	1	96200	7885	5981	22041
	2	71083	-15885	-23770	-29751
	3	83702	-17999	-2114	21656
	4	77001	-8520	9479	11593
1986	1	75830	-20370	-11850	-21329
	2	79563	8480	28850	40700
	3	71583	-12119	-20599	-49449
	4	55248	-21753	-9634	10965

1987	1	57088	-18742	3011	12645
	2	63014	-16549	2193	-818
	3	68675	-2908	13641	11448
	4	69487	14239	17147	3506
1988	1	70569	13481	-758	-17905
	2	70459	7445	-6036	-5278
	3	81479	12804	5359	11395
	4	76619	7132	-5672	-11031
1989	1	73847	3278	-3854	1818
	2	63853	-6606	-9884	-6030
	3	74806	-6673	-67	9817
	4	88613	11994	18667	18734

TREND AND SEASONAL ANALYSIS
USAFE FLYING HOURS

	Actual	DBQ	DBQ-1	DBQ-2
Variance:	65769669	22106275	29956505	78090643
Index:	100%	34%	46%	119%
Trend:	None	None	Moderate	Strong
Seasonal:	No	Yes ++++++++	Yes	Yes

Year	Qtr	Actual Data	Diffs Between Same Qtr. Each Yr.	First Diffs Between Diffs	Second Diffs Between Diffs
1978	1	57480			
	2	52034			
	3	64399			
	4	69807			
1979	1	55423	-2057		
	2	56900	4866	6923	
	3	71221	6822	1956	-4967
	4	71846	2039	-4783	-6739
1980	1	60461	5038	2999	7782
	2	58808	1908	-3130	-6129
	3	69735	-1486	-3394	-264
	4	68069	-3777	-2291	1103
1981	1	63546	3085	6862	9153
	2	65588	6780	3695	-3167
	3	75263	5528	-1252	-4947
	4	75477	7408	1880	3132
1982	1	61878	-1668	-9076	-10956
	2	70025	4437	6105	15181
	3	80973	5710	1273	-4832
	4	78365	2888	-2822	-4095
1983	1	66217	4339	1451	4273
	2	70496	471	-3868	-5319
	3	77627	-3346	-3817	51
	4	77944	-421	2925	6742
1984	1	69485	3268	3689	764
	2	73093	2597	-671	-4360
	3	78651	1024	-1573	-902
	4	77702	-242	-1266	307
1985	1	69027	-458	-216	1050
	2	70119	-2974	-2516	-2300
	3	83336	4685	7659	10175
	4	79858	2156	-2529	-10188
1986	1	75552	6525	4369	6898
	2	74412	4293	-2232	-6601
	3	80673	-2663	-6956	-4724
	4	77628	-2230	433	7389

1987	1	75308	-244	1986	1553
	2	71368	-3044	-2800	-4786
	3	87519	6846	9890	12690
	4	81329	3701	-3145	-13035
1988	1	75433	125	-3576	-431
	2	76434	5066	4941	8517
	3	74331	-13188	-18254	-23195
	4	74752	-6577	6611	24865
1989	1	68136	-7297	-720	-7331
	2	74152	-2282	5015	5735
	3	85631	11300	13582	8567
	4	83012	8260	-3040	-16622

Appendix D: Time Series Analysis (PACAF and USAFE MSC)

ARIMA Procedure

Name of variable = PACAF MSC.

Mean of working series = 39621.55
Standard deviation = 6929.082
Number of observations = 42

Autocorrelations

Lag	Covariance	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
0	48012179	1.00000	¦ ¦**********
1	35649549	0.74251	¦ . ¦********** ¦
2	30163492	0.62825	¦ . ¦********** ¦
3	24096312	0.50188	¦ . ¦********** ¦
4	22964186	0.47830	¦ . ¦********** . ¦
5	16605674	0.34586	¦ . ¦******* . ¦
6	11711346	0.24392	¦ . ¦***** . ¦
7	9810391	0.20433	¦ . ¦**** . ¦
8	10977211	0.22863	¦ . ¦***** . ¦
9	12790763	0.26641	¦ . ¦***** . ¦
10	12114561	0.25232	¦ . ¦***** . ¦
11	10597599	0.22073	¦ . ¦**** . ¦
12	4469521	0.09309	¦ . ¦** . ¦
13	-1096447	-0.02284	¦ . ¦ . ¦
14	-3893742	-0.08110	¦ . **¦ . ¦
15	-6897490	-0.14366	¦ . ***¦ . ¦
16	-13755892	-0.28651	¦ . ******¦ . ¦
17	-17785147	-0.37043	¦ . *******¦ . ¦
18	-15887078	-0.33090	¦ . *******¦ . ¦
19	-13473166	-0.28062	¦ . ******¦ . ¦
20	-11830748	-0.24641	¦ . *****¦ . ¦
21	-15205155	-0.31669	¦ . ******¦ . ¦
22	-10811710	-0.22519	¦ . *****¦ . ¦
23	-10514095	-0.21899	¦ . ****¦ . ¦
24	-11463170	-0.23876	¦ . *****¦ . ¦

"." marks two standard errors

Inverse Autocorrelations

```
Lag Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  1   -0.24423  |               .*****|     .               |
  2   -0.20589  |               . ****|     .               |
  3    0.06277  |               .     |*    .               |
  4   -0.23921  |               .*****|     .               |
  5    0.02186  |               .     |     .               |
  6    0.12740  |               .     |***  .               |
  7    0.08979  |               .     |**   .               |
  8    0.01716  |               .     |     .               |
  9   -0.09046  |               .   **|     .               |
 10    0.00491  |               .     |     .               |
 11   -0.09979  |               .   **|     .               |
 12   -0.05500  |               .    *|     .               |
 13    0.17202  |               .     |***  .               |
 14   -0.01930  |               .     |     .               |
 15   -0.13686  |               .  ***|     .               |
 16    0.13080  |               .     |***  .               |
 17    0.02136  |               .     |     .               |
 18   -0.01801  |               .     |     .               |
 19    0.03337  |               .     |*    .               |
 20   -0.12660  |               .  ***|     .               |
 21    0.08992  |               .     |**   .               |
```

Partial Autocorrelations

```
Lag Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  1    0.74251  |               .     |***************      |
  2    0.17145  |               .     |***  .               |
  3   -0.02738  |               .    *|     .               |
  4    0.15907  |               .     |***  .               |
  5   -0.16773  |               .  ***|     .               |
  6   -0.09245  |               .   **|     .               |
  7    0.10488  |               .     |**   .               |
  8    0.12153  |               .     |**   .               |
  9    0.15075  |               .     |***  .               |
 10    0.00043  |               .     |     .               |
 11   -0.08434  |               .   **|     .               |
 12   -0.32117  |               ******|     .               |
 13   -0.26436  |               .*****|     .               |
 14    0.04262  |               .     |*    .               |
 15    0.03411  |               .     |*    .               |
 16   -0.16879  |               .  ***|     .               |
 17   -0.07732  |               .   **|     .               |
 18    0.06600  |               .     |*    .               |
 19   -0.04001  |               .    *|     .               |
 20    0.05647  |               .     |*    .               |
 21   -0.15255  |               .  ***|     .               |
 22    0.22454  |               .     |**** .               |
 23    0.00155  |               .     |     .               |
 24   -0.15186  |               .  ***|     .               |
```

Autocorrelation Check for White Noise

To Lag	Chi Square	DF	Prob	Autocorrelations					
6	75.17	6	0.000	0.743	0.628	0.502	0.478	0.346	0.244
12	91.31	12	0.000	0.204	0.229	0.266	0.252	0.221	0.093
18	117.60	18	0.000	-0.023	-0.081	-0.144	-0.287	-0.370	-0.331
24	153.05	24	0.000	-0.281	-0.246	-0.317	-0.225	-0.219	-0.239

Name of variable = USAFE MSC.

Mean of working series = 65243.88
Standard deviation = 17126.54
Number of observations = 42

Autocorrelations

Lag	Covariance	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
0	293318259	1.00000	¦ ¦**********¦
1	250017831	0.85238	¦ . ¦********** ¦
2	232369070	0.79221	¦ . ¦********** ¦
3	198721019	0.67749	¦ . ¦********** ¦
4	180762031	0.61627	¦ . ¦**********. ¦
5	161784695	0.55157	¦ . ¦********** . ¦
6	143576754	0.48949	¦ . ¦********** . ¦
7	109525766	0.37340	¦ . ¦******* . ¦
8	76362399	0.26034	¦ . ¦***** . ¦
9	33396978	0.11386	¦ . ¦** . ¦
10	23002501	0.07842	¦ . ¦** . ¦
11	670556	0.00229	¦ . ¦ . ¦
12	-15954055	-0.05439	¦ . * ¦ . ¦
13	-39300605	-0.13399	¦ . *** ¦ . ¦
14	-61085526	-0.20826	¦ . **** ¦ . ¦
15	-83071293	-0.28321	¦ . ****** ¦ . ¦
16	-86472888	-0.29481	¦ . ****** ¦ . ¦
17	-100132721	-0.34138	¦ . ******* ¦ . ¦
18	-105977923	-0.36131	¦ . ******* ¦ . ¦
19	-115554891	-0.39396	¦ . ******** ¦ . ¦
20	-120942731	-0.41233	¦ . ******** ¦ . ¦
21	-115437213	-0.39356	¦ . ******** ¦ . ¦
22	-108146557	-0.36870	¦ . ******* ¦ . ¦
23	-105077075	-0.35824	¦ . ******* ¦ . ¦
24	-92430423	-0.31512	¦ . ****** ¦ . ¦

"." marks two standard errors

Inverse Autocorrelations

```
Lag Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  1   -0.27127  |               .*****|     .              |
  2   -0.35492  |              *******|     .              |
  3    0.09219  |               .     |**   .              |
  4    0.03951  |               .     |*    .              |
  5    0.08011  |               .     |**   .              |
  6    0.03184  |               .     |*    .              |
  7   -0.28483  |               ******|     .              |
  8   -0.04015  |               .    *|     .              |
  9    0.41005  |               .     |********           |
 10   -0.14428  |               .  ***|     .              |
 11   -0.06951  |               .    *|     .              |
 12    0.00926  |               .     |     .              |
 13   -0.04163  |               .    *|     .              |
 14    0.05883  |               .     |*    .              |
 15    0.09758  |               .     |**   .              |
 16   -0.19514  |               . ****|     .              |
 17    0.02536  |               .     |*    .              |
 18    0.11366  |               .     |**   .              |
 19   -0.06173  |               .    *|     .              |
 20    0.04426  |               .     |*    .              |
 21   -0.02859  |               .    *|     .              |
```

Partial Autocorrelations

```
Lag Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  1    0.85238  |               .     |*****************   |
  2    0.24012  |              ..     |*****.              |
  3   -0.15638  |               .  ***|     .              |
  4    0.05903  |               .     |*    .              |
  5    0.04609  |               .     |*    .              |
  6   -0.05177  |               .    *|     .              |
  7   -0.25197  |               .*****|     .              |
  8   -0.15993  |               .  ***|     .              |
  9   -0.20564  |               . ****|     .              |
 10    0.24457  |               .     |*****.              |
 11   -0.01919  |               .     |     .              |
 12   -0.15676  |               .  ***|     .              |
 13   -0.04754  |               .    *|     .              |
 14   -0.00218  |               .     |     .              |
 15   -0.05848  |               .    *|     .              |
 16    0.06489  |               .     |*    .              |
 17   -0.16316  |               .  ***|     .              |
 18   -0.12883  |               .  ***|     .              |
 19    0.13620  |               .     |***  .              |
 20   -0.01342  |               .     |     .              |
 21    0.06167  |               .     |*    .              |
 22    0.00602  |               .     |     .              |
 23   -0.15598  |               .  ***|     .              |
 24    0.11762  |               .     |**   .              |
```

ARIMA Procedure

Autocorrelation Check for White Noise

To Lag	Chi Square	DF	Prob	Autocorrelations					
6	129.46	6	0.000	0.852	0.792	0.677	0.616	0.552	0.489
12	141.77	12	0.000	0.373	0.260	0.114	0.078	0.002	-0.054
18	176.11	18	0.000	-0.134	-0.208	-0.283	-0.295	-0.341	-0.361
24	251.72	24	0.000	-0.394	-0.412	-0.394	-0.369	-0.358	-0.315

Appendix E: Business Cycle Analysis (PACAF and USAFE MSC)

Year	Qtr	PACAF MSC	USAFE MSC	PACAF 1/4 RATIO	USAFE 1/4 RATIO	PACAF 4-QTR TOTAL	USAFE 4-QTR TOTAL
1978	1	31163	46778				
	2	30967	38450				
	3	32924	47078				
	4	32018	37450			127072	169756
1979	1	33469	48574	107	104	129378	171552
	2	30548	44156	99	115	128959	177258
	3	33046	44041	100	94	129081	174221
	4	34991	44296	109	118	132054	181067
1980	1	33145	42971	99	88	131730	175464
	2	30312	50901	99	115	131494	182209
	3	35818	56271	108	128	134266	194439
	4	32220	49844	92	113	131495	199987
1981	1	35198	53377	106	124	133548	210393
	2	30649	53945	101	106	133885	213437
	3	35193	60785	98	108	133260	217951
	4	35396	54941	110	110	136436	223048
1982	1	37343	55855	106	105	138581	225526
	2	41379	57543	135	107	149311	229124
	3	43392	63076	123	104	157510	231415
	4	42968	65851	121	120	165082	242325
1983	1	43039	91436	115	164	170778	277906
	2	49651	93263	120	162	179050	313626
	3	46352	83737	107	133	182010	334287
	4	35398	83617	82	127	174440	352053
1984	1	38462	88315	89	97	169863	348932
	2	41800	86968	84	93	162012	342637
	3	48352	101701	104	121	164012	360601
	4	49203	85521	139	102	177817	362505
1985	1	47567	96200	124	109	186922	370390
	2	49835	71083	119	82	194957	354505
	3	59435	83702	123	82	206040	336506
	4	48235	77001	98	90	205072	327986
1986	1	49040	75830	103	79	206545	307616
	2	40829	79563	82	112	197539	316096
	3	42134	71583	71	86	180238	303977
	4	34675	55248	72	72	166678	282224
1987	1	42681	57088	87	75	160319	263482
	2	39408	63014	97	79	158898	246933
	3	35796	68675	85	96	152560	244025
	4	39293	69487	113	126	157178	258264
1988	1	42387	70569	99	124	156884	271745
	2	48394	70459	123	112	165870	279190

Year	Qtr	PACAF 4/4 RATIO	USAFE 4/4 RATIO
1978	1		
	2		
	3		
	4		
1979	1		
	2		
	3		
	4	104	107
1980	1	102	102
	2	102	103
	3	104	112
	4	100	110
1981	1	101	120
	2	102	117
	3	99	112
	4	104	112
1982	1	104	107
	2	112	107
	3	118	106
	4	121	109
1983	1	123	123
	2	120	137
	3	116	144
	4	106	145
1984	1	99	126
	2	90	109
	3	90	108
	4	102	103
1985	1	110	106
	2	120	103
	3	126	93
	4	115	90
1986	1	110	83
	2	101	89
	3	87	90
	4	81	86
1987	1	78	86
	2	80	78
	3	85	80
	4	94	92
1988	1	98	103
	2	104	113

Appendix F: PACAF and USAFE Military Populations

	PACAF		USAFE	
YEAR	OFFICER	AIRMAN	OFFICER	AIRMAN
1981	5340	14635	9468	28394
1982	5537	na*	9515	na*
1983	5671	15405	9935	28432
1984	5836	15338	10025	28610
1985	5907	15417	10312	28470
1986	5995	15385	10354	28271
1987	5976	15503	10272	28090
1988	5806	15664	10004	27951
1989	5727	15579	10110	28346

* data not available.

Appendix G: PACAF Aircraft Flying Hours and Inventory by MD

FISCAL YEAR: 1985

MD	1st Qtr INV.	1st Qtr FLY.HRS.	2nd Qtr INV.	2nd Qtr FLY.HRS.	3rd Qtr INV.	3rd Qtr FLY.HRS.	4th Qtr INV.	4th Qtr FLY.HRS.
A010	26	2601.5	26	2505.3	25	2856.1	25	2405.1
A037	13	1087.9	13	977.9	13	364.2	0	2.0
B052	12	1482.6	12	1696.4	12	1586.3	12	1394.4
C009		811.2		745.7		813.8		714.3
C012	1	162.5		1398.8		1156.0		1123.9
C021		0.0		488.9		396.2		417.4
C130	40	8118.6	40	7564.5	40	8017.5	40	8597.2
C135	33	4607.9	33	4220.9	33	4227.8	33	4184.7
E003	0	0.0	0	0.0	1	420.5	1	639.1
F004	92	6087.2	89	6164.3	91	5976.0	73	5020.5
F005	12	781.8	12	879.2	12	803.7	12	766.3
F015	70	5667.7	71	5252.0	70	5617.0	70	5221.3
F016	50	4027.5	50	4224.8	51	4438.2	59	5485.9
H001	2	145.6	1	130.8	2	143.9	2	179.7
H003	13	1827.6	13	1727.4	13	1631.5	13	1536.9
T033	13	1077.5	14	1138.8	14	1142.7	14	1012.0
T039	8	963.5	4	269.3	0	0.0	0	0.0
V010	14	1352.8	15	1442.4	15	1752.5	12	2346.3
TOTAL:		40803.4		40827.4		41343.9		41047.0

1986

MD	1st Qtr INV.	1st Qtr FLY.HRS.	2nd Qtr INV.	2nd Qtr FLY.HRS.	3rd Qtr INV.	3rd Qtr FLY.HRS.	4th Qtr INV.	4th Qtr FLY.HRS.
A010	26	2535.3	26	2508.5	26	2654.5	26	2391.0
A037	0	0.0	0	0.0	0	0.0	0	0.0
B052	12	1325.1	13	1539.4	13	1452.0	12	1343.0
C009	2	627.6	3	1087.1	3	796.4	1	776.8
C012	4	1109.6	4	983.9	4	1438.0	[illegible]	1479.6
C021	2	416.5	1	276.7	1	295.8	2	341.1
C130	40	8132.9	41	7444.6	41	7402.5	39	8777.9
C135	33	4462.1	33	3786.0	33	3925.9	33	3807.9
E003	1	630.1	2	738.6	2	646.8	2	493.4
F004	91	6172.7	92	5826.7	92	5626.9	89	5313.6
F005	12	791.1	12	710.1	12	811.7	12	679.0
F015	71	5633.2	70	5586.0	70	5057.6	68	5401.1
F016	74	5759.9	75	5695.6	75	5578.6	70	5539.0
H001	2	204.1	2	164.9	2	270.4	2	196.3
H003	13	1493.3	13	1661.6	13	1526.9	10	1359.0
T033	14	1058.0	14	1101.6	14	1166.4	10	992.0
T039	0	0.0	0	0.0	0	0.0	0	0.0
V010	29	2553.7	29	2830.8	29	2826.0	29	2481.0
TOTAL:		42905.2		41942.1		41476.4		41371.7

1987

MD	1st Qtr INV.	1st Qtr FLY.HRS.	2nd Qtr INV.	2nd Qtr FLY.HRS.	3rd Qtr INV.	3rd Qtr FLY.HRS.	4th Qtr INV.	4th Qtr FLY.HRS.
A010	26	2703.5	26	2464.2	26	2810.4	25	2633.9
A037	0	0.0	0	0.0	0	0.0	0	0.0
B052	13	1608.2	13	1747.0	14	1744.7	13	1587.3
C009	3	857.9	3	820.0	3	698.0	3	733.5
C012	6	1455.4	6	1166.5	6	1463.5	6	1223.6
C021	2	464.6	2	484.1	1	426.9	2	418.8
C130	44	8639.8	42	7117.8	35	6771.9	41	7045.5
C135	26	4010.3	32	3876.9	32	3851.8	30	3793.3
E003	2	631.6	3	684.0	2	686.6	1	732.7
F004	91	5723.4	90	6053.5	91	6428.5	90	5704.6
F005	12	787.8	12	830.7	11	854.5	10	793.3
F015	66	5692.9	70	5523.1	69	5507.3	69	5427.7
F016	68	6186.7	75	5929.1	71	7343.1	87	8278.1
H001	2	190.4	2	233.0	2	259.9	2	256.7
H003	14	1590.2	14	1546.8	13	1617.4	15	1463.4
T033	14	1022.0	14	1136.6	14	1098.4	14	735.2
T039	0	0.0	0	0.0	0	0.0	0	0.0
V010	27	2705.2	26	2708.9	28	2818.0	29	2946.1
TOTAL:		44269.9		42322.2		44380.9		43773.7

1988

MD	1st Qtr INV.	1st Qtr FLY.HRS.	2nd Qtr INV.	2nd Qtr FLY.HRS.	3rd Qtr INV.	3rd Qtr FLY.HRS.	4th Qtr INV.	4th Qtr FLY.HRS.
A010	24	2530.3	23	2719.4	26	2700.5	26	2119.8
A037	0	0.0	0	0.0	0	0.0	0	0.0
B052	14	1384.5	14	1695.7	14	1782.4	14	1683.7
C009	2	755.5	2	824.2	2	729.5	2	1049.8
C012	4	1545.6	2	1399.7	2	1732.9	8	1852.9
C021	2	654.1	2	506.5	3	376.0	3	540.0
C130	34	6675.2	38	6197.0	38	6185.6	33	6669.7
C135	30	3601.6	30	3878.1	30	3887.6	30	3883.5
E003	2	726.0	1	641.8	1	629.1	2	679.0
F004	91	5815.4	86	6705.6	93	6158.5	93	4944.5
F005	11	685.4	10	484.5	6	452.1	6	253.6
F015	71	5189.0	68	5704.8	68	5140.6	67	4773.6
F016	107	7464.9	102	8905.5	110	7594.4	110	6567.9
H001	3	230.9	3	260.8	1	251.1	2	257.2
H003	13	1571.3	15	1590.9	14	1389.2	11	1359.5
T033	11	0.0	7	0.0	4	0.0	0	0.0
T039	0	0.0	0	191.2	0	249.3	0	106.6
V010	29	2478.0	28	2287.9	29	2464.5	29	2153.6
TOTAL:		41307.7		43993.6		41723.3		38894.9

1989

MD	1st Qtr INV.	1st Qtr FLY.HRS.	2nd Qtr INV.	2nd Qtr FLY.HRS.	3rd Qtr INV.	3rd Qtr FLY.HRS.	4th Qtr INV.	4th Qtr FLY.HRS.
A010	25	2146.5	24	2800.6	23	2656.6	21	2227.9
A037	0	0.0	0	0.0	0	0.0	0	0.0
B052	14	1577.5	14	1633.3	14	1652.5	12	1700.7
C009	3	753.8	3	897.1	2	910.5	3	916.0
C012	7	2214.0	9	2243.5	10	2381.7	8	2140.6
C021	1	592.7	3	676.1	3	711.5	3	489.0
C130	32	6240.3	39	5863.9	42	6408.0	41	4914.7
C135	30	3390.1	27	3253.4	27	3314.0	28	3245.9
E003	2	395.8	2	1288.0	3	665.7	1	594.2
F004	94	5837.9	99	5366.0	79	4924.4	70	4112.1
F005	6	50.0	3	0.0	0	0.0	0	0.0
F015	69	5376.7	73	6228.2	72	5565.6	73	5447.5
F016	107	9165.9	106	9904.0	117	11430.1	128	9173.9
H001	3	263.1	3	454.0	5	531.2	5	466.5
H003	9	1274.4	12	960.9	14	1201.0	12	1155.8
T033	0	0.0	0	0.0	0	0.0	0	0.0
T039	0	1.6	9	0.0	0	0.0	0	0.0
V010	27	1394.5	20	390.5	9	0.0	1	0.0
TOTAL:		40674.8		41959.5		42352.8		36584.8

1990

MD	1st Qtr INV.	1st Qtr FLY.HRS.
A010	22	1759
A037	0	0
B052	14	1382
C009	3	930
C012	7	1958
C021	2	408
C130	25	4370
C135	27	2990
E003	1	531
F004	65	4649
F005	0	0
F015	72	5677
F016	144	10525
H001	6	462
H003	12	1032
T033	0	0
T039	0	0
V010	0	0
TOTAL:		36672

Appendix H: PACAF MSC Multiple Regression Model SAS Output

First Order Regression Model

Dependent Variable: TON

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	3	364567520.74	121522506.91	6.153	0.0122
Error	10	197486140.19	19748614.019		
C Total	13	562053660.93			

Root MSE	4443.94127	R-square	0.6486
Dep Mean	44264.92857	Adj R-sq	0.5432
C.V.	10.03942		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	385572	125758.88080	3.066	0.0119
A10	1	13.819258	9.02533356	1.531	0.1567
OFF	1	-61.515186	20.19108244	-3.047	0.0123
F16	1	-1.944066	0.86970203	-2.235	0.0494

Variable	DF	Variance Inflation
INTERCEP	1	0.00000000
A10	1	1.09288403
OFF	1	1.04062524
F16	1	1.06684200

Durbin-Watson D	1.386
(For Number of Obs.)	14
1st Order Autocorrelation	0.178

Obs	Dep Var TON	Predict Value	Std Err Predict	Lower95% Mean	Upper95% Mean	Lower95% Predict	Upper95% Predict
1	47567.0	54696.7	2851.315	48343.6	61049.9	42932.0	66461.4
2	49835.0	48605.7	2095.787	43936.0	53275.4	37658.0	59553.4
3	59435.0	53042.2	3202.388	45906.8	60177.6	40837.3	65247.1
4	48235.0	44772.3	2223.852	39817.2	49727.4	33699.9	55844.7
5	49040.0	46036.1	1491.367	42713.1	49359.1	35591.6	56480.6
6	40829.0	40387.9	1729.508	36534.3	44241.5	29762.7	51013.1
7	42134.0	42633.0	1892.572	38416.0	46849.9	31870.6	53395.3
8	34675.0	39062.5	2273.832	33996.0	44128.9	27939.8	50185.1
9	42681.0	42128.1	1998.290	37675.6	46580.6	31271.3	52984.9
10	39408.0	40481.9	1724.861	36638.6	44325.1	29860.4	51103.3
11	35796.0	42514.4	2509.084	36923.8	48105.0	31143.4	53885.4
12	39293.0	38264.5	2292.814	33155.8	43373.3	27122.5	49406.5
13	42387.0	38407.8	1941.387	34082.1	42733.6	27602.4	49213.3
14	48394.0	48675.9	3874.792	40042.2	57309.5	35538.7	61813.0
15	.	50977.7	3250.200	43735.8	58219.7	38710.3	63245.2
16	.	44943.4	5716.943	32205.1	57681.6	28809.3	61077.4
17	.	40265.8	6571.074	25624.4	54907.2	22590.5	57941.1
18	.	52728.6	5615.448	40216.5	65240.6	36772.5	68684.7
19	.	47772.0	6636.996	32983.7	62560.2	29974.8	65569.1
20	.	46229.3	7086.267	30440.0	62018.6	27592.1	64866.5

Obs	Residual	Std Err Residual	Student Residual	-2-1-0 1 2	Cook's D
1	-7129.7	3408.609	-2.092	****¦ ¦	0.765
2	1229.3	3918.710	0.314	¦ ¦	0.007
3	6392.8	3081.124	2.075	¦ ¦****	1.163
4	3462.7	3847.479	0.900	¦ ¦*	0.068
5	3003.9	4186.220	0.718	¦ ¦*	0.016
6	441.1	4093.583	0.108	¦ ¦	0.001
7	-499.0	4020.794	-0.124	¦ ¦	0.001
8	-4387.5	3818.154	-1.149	**¦ ¦	0.117
9	552.9	3969.314	0.139	¦ ¦	0.001
10	-1073.9	4095.542	-0.262	¦ ¦	0.003
11	-6718.4	3667.848	-1.832	***¦ ¦	0.393
12	1028.5	3806.786	0.270	¦ ¦	0.007
13	3979.2	3997.453	0.995	¦ ¦*	0.058
14	-281.9	2175.913	-0.130	¦ ¦	0.013
15	.	.	.		.
16	.	.	.		.
17	.	.	.		.
18	.	.	.		.
19	.	.	.		.
20	.	.	.		.

Sum of Residuals 0
Sum of Squared Residuals 197486140.19
Predicted Resid SS (Press) 521860388.90

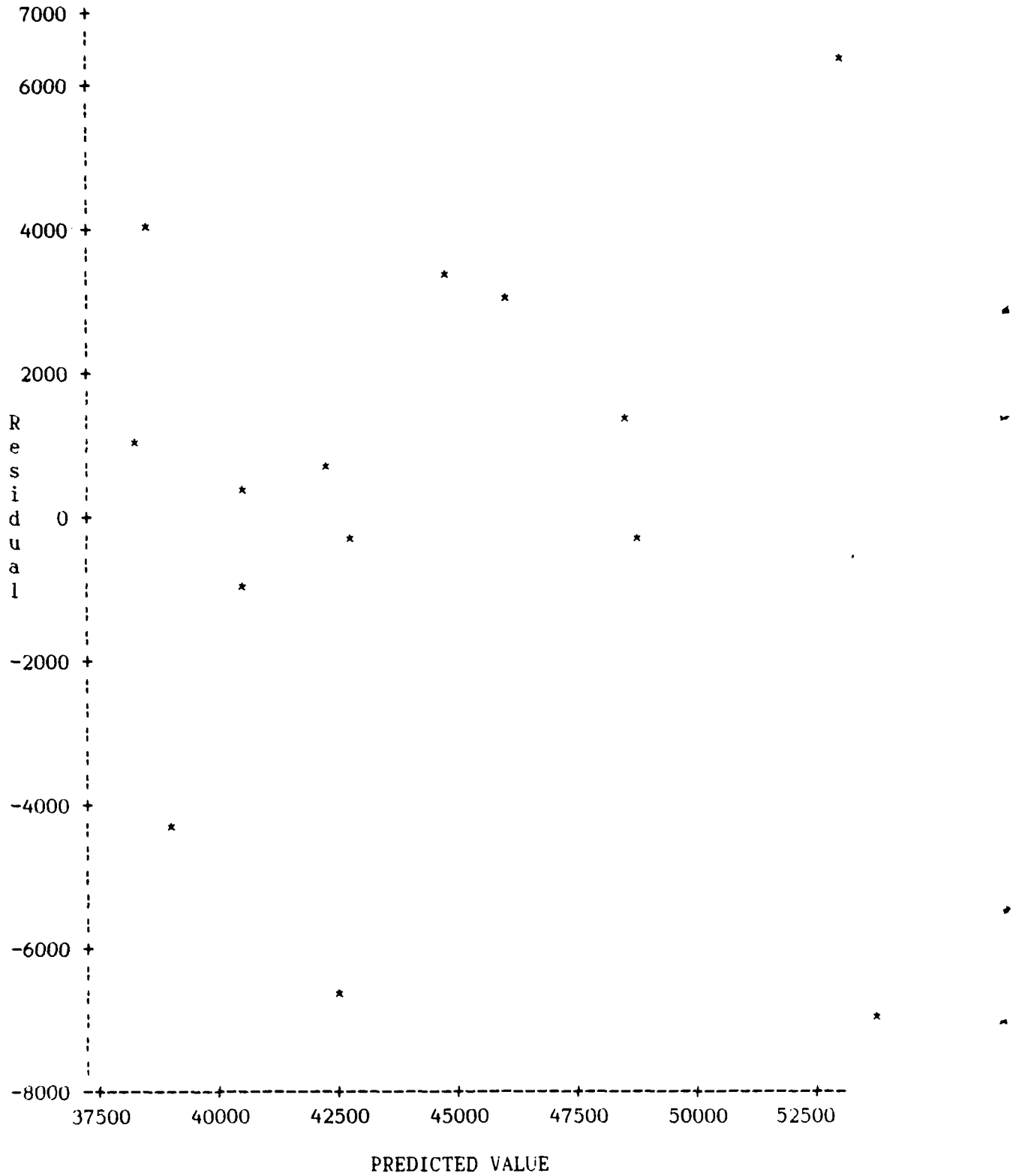

Figure 65. PACAF Multiple Regression MSC Model Residuals versus Predicted Values

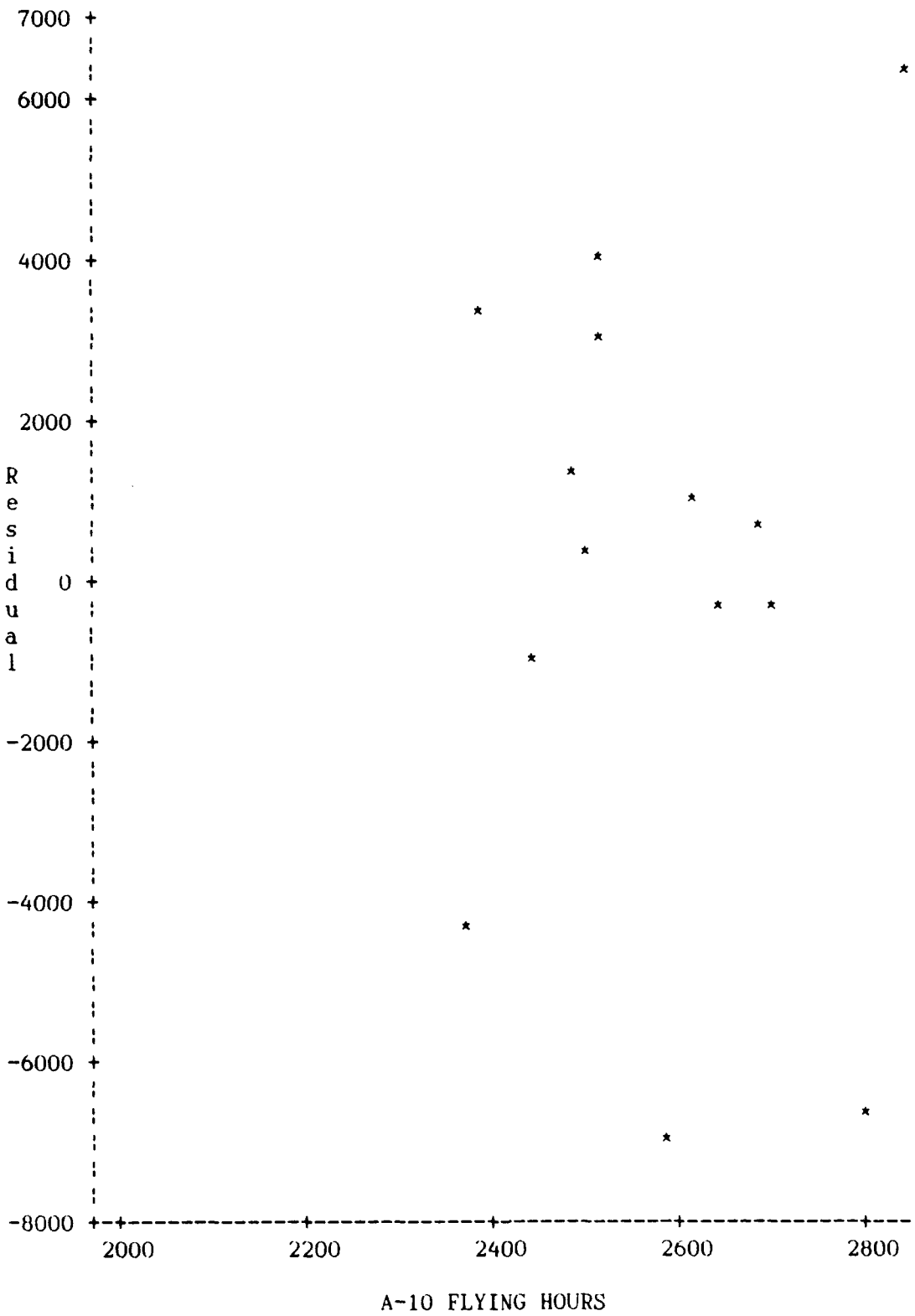

Figure 66. PACAF Multiple Regression MSC Model Residuals versus A-10 Flying Hours

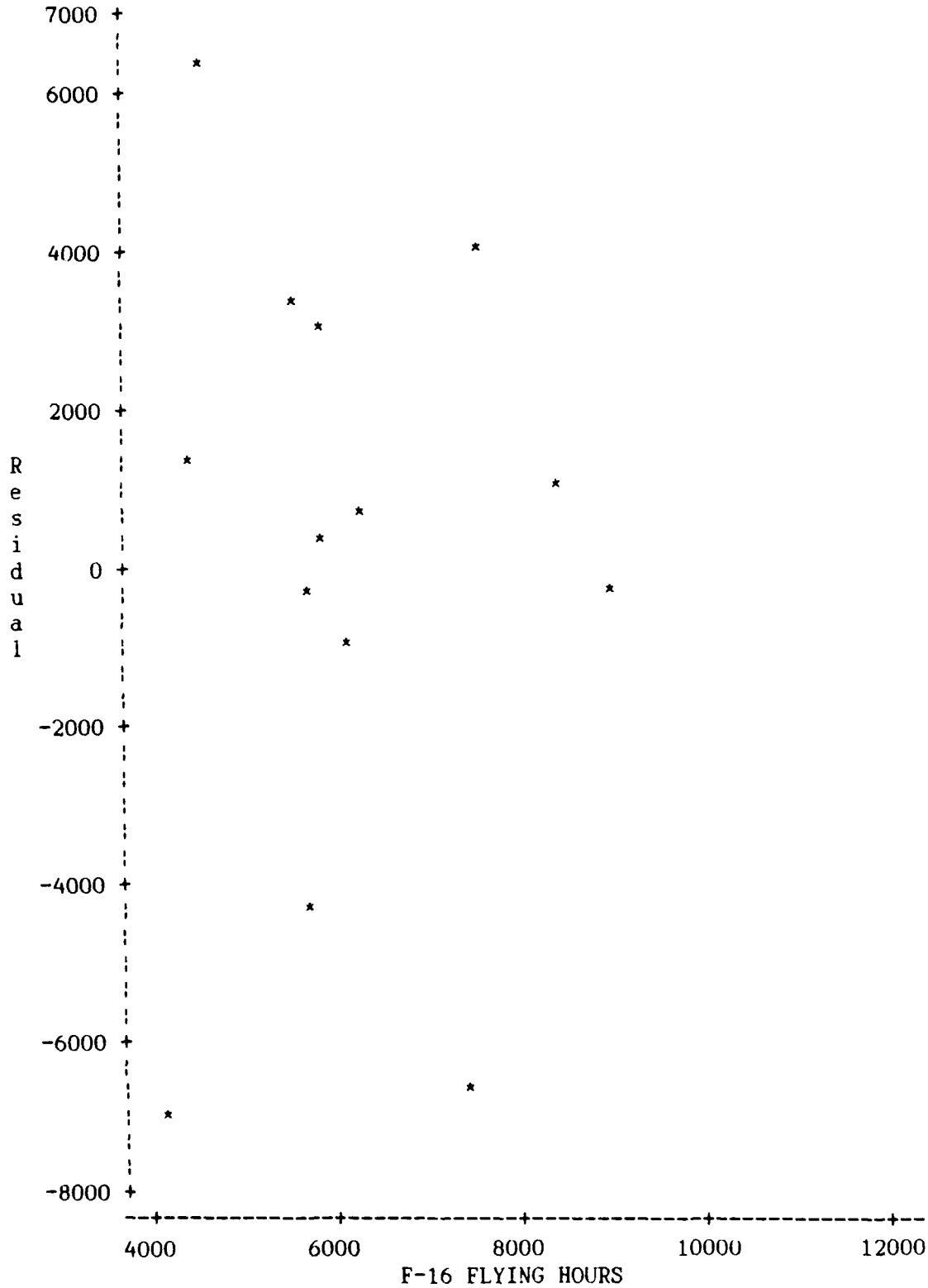

Figure 67. PACAF Multiple Regression MSC Model Residuals versus F-16 Flying Hours

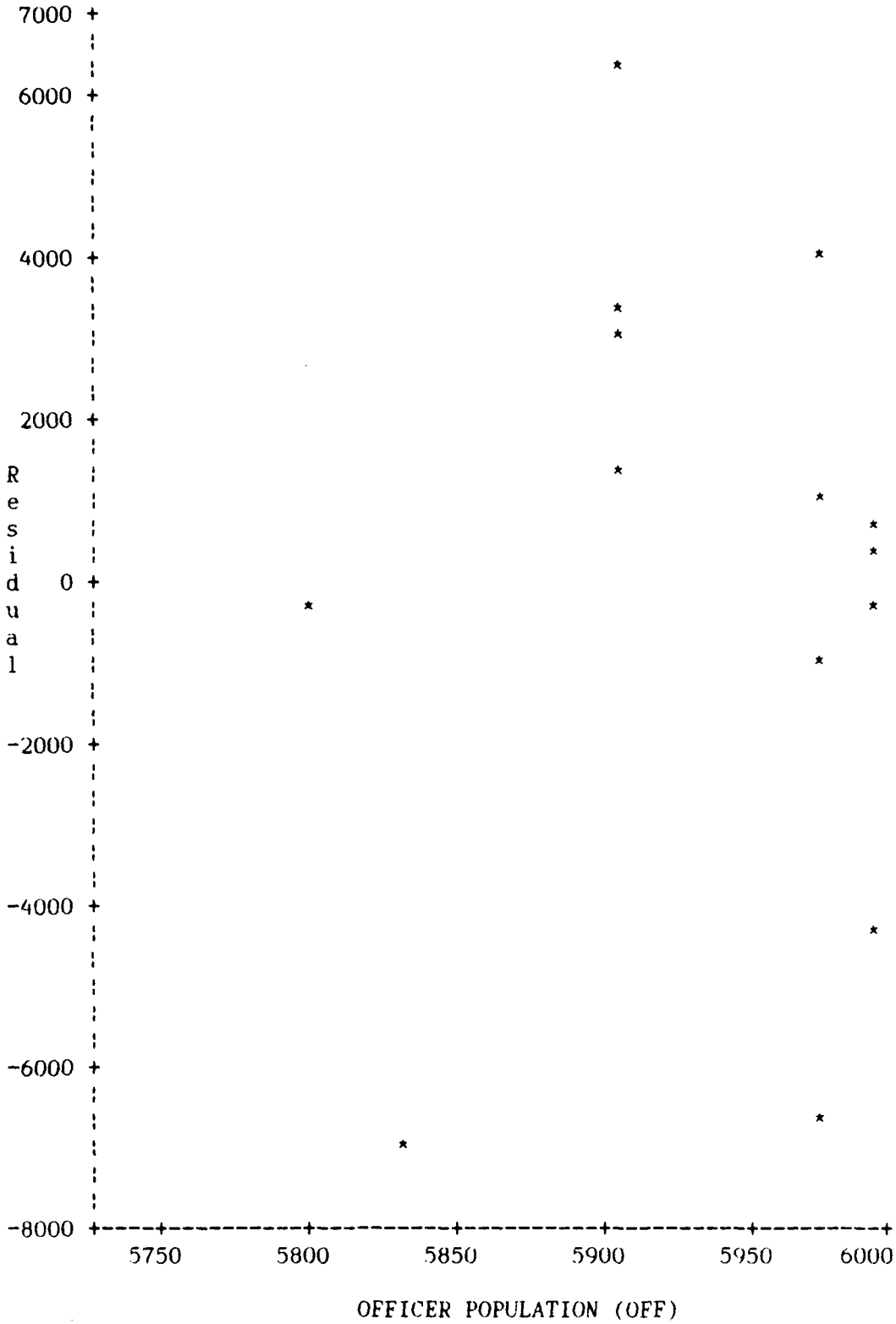

Figure 68. PACAF Multiple Regression MSC Model Residuals versus Officer Population (OFF)

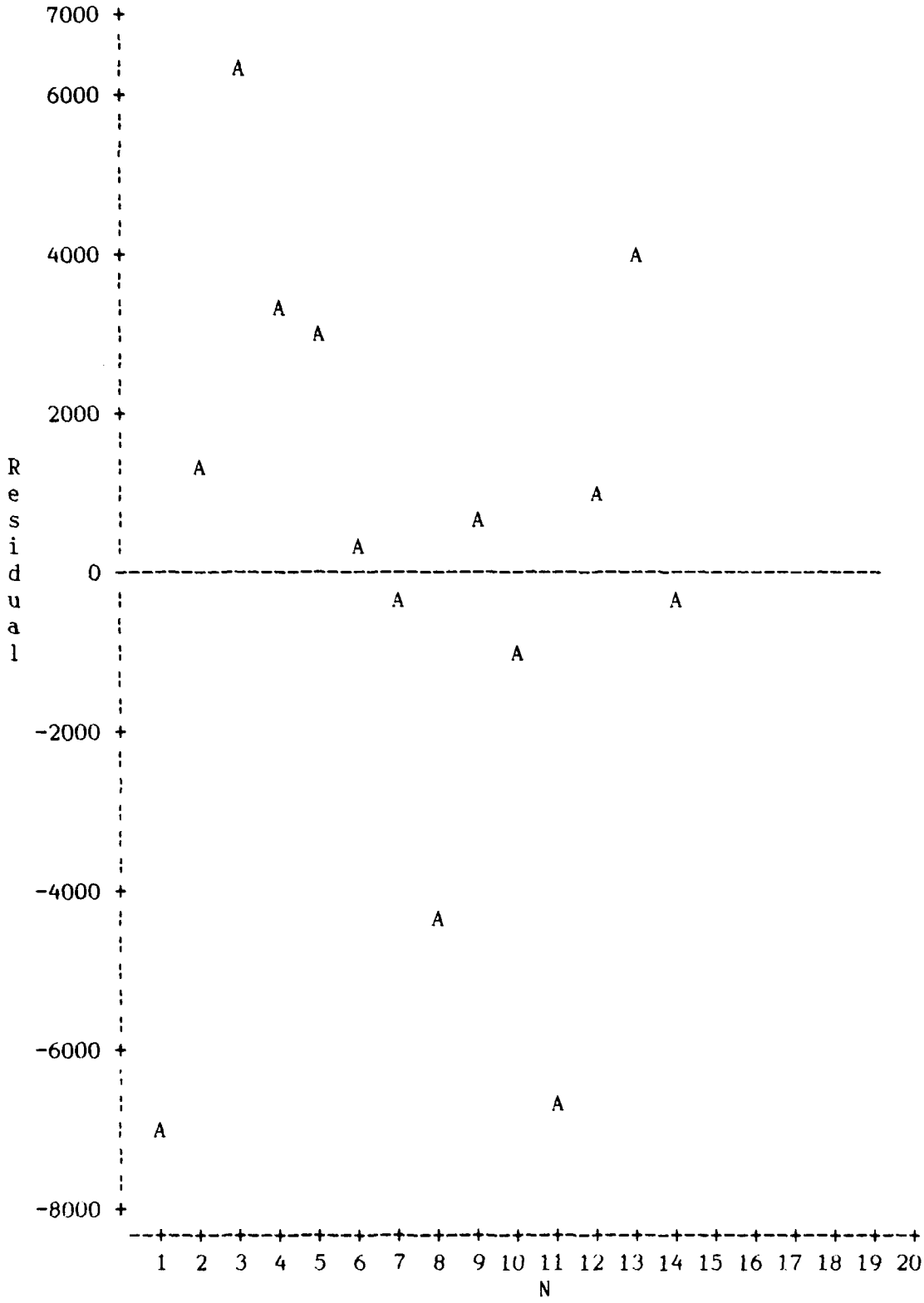

Figure 69. PACAF Multiple Regression MSC Model Residuals versus Time (N)

Univariate Procedure

Variable=RESIDUAL Residual

Moments

N	14	Sum Wgts	14
Mean	0	Sum	0
Std Dev	3897.594	Variance	15191242
Skewness	-0.53872	Kurtosis	-0.05306
USS	1.9749E8	CSS	1.9749E8
CV	.	Std Mean	1041.676
T:Mean=0	0	Prob>|T|	1.0000
Num ^= 0	14	Num > 0	8
M(Sign)	1	Prob>|M|	0.7905
Sgn Rank	4.5	Prob>|S|	0.8077
W:Normal	0.940673	Prob<W	0.4089

Quantiles(Def=5)

100% Max	6392.836	99%	6392.836
75% Q3	3003.873	95%	6392.836
50% Med	496.9806	90%	3979.15
25% Q1	-1073.86	10%	-6718.42
0% Min	-7129.72	5%	-7129.72
		1%	-7129.72
Range	13522.55		
Q3-Q1	4077.738		
Mode	-7129.72		

Extremes

Lowest	Obs	Highest	Obs
-7129.72(	1)	1229.309(	2)
-6718.42(	11)	3003.873(	5)
-4387.46(	8)	3462.703(	4)
-1073.86(	10)	3979.15(	13)
-498.977(	7)	6392.836(	3)

Missing Value	.
Count	6
% Count/Nobs	30.00

Second Order Model

Dependent Variable: TON

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	3	395623696.61	131874565.54	7.924	0.0053
Error	10	166429964.32	16642996.432		
C Total	13	562053660.93			

Root MSE	4079.58287	R-square	0.7039
Dep Mean	44264.92857	Adj R-sq	0.6151
C.V.	9.21629		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	-32409859	11757503.096	-2.757	0.0203
A10	1	12.182102	8.12234698	1.500	0.1646
OFF	1	11035	3976.2335756	2.775	0.0196
SOFF	1	-0.938747	0.33623126	-2.792	0.0191

Variable	DF	Variance Inflation
INTERCEP	1	0.00000000
A10	1	1.05030573
OFF	1	47887.705728
SOFF	1	47876.954297

Collinearity Diagnostics(intercept adjusted)

Number	Eigenvalue	Condition Number	Var Prop A10	Var Prop OFF	Var Prop SOFF
1	2.05281	1.00000	0.0222	0.0000	0.0000
2	0.94718	1.47217	0.9572	0.0000	0.0000
3	0.0000104	443.37959	0.0207	1.0000	1.0000

Durbin-Watson D	1.517
(For Number of Obs.)	14
1st Order Autocorrelation	0.227

Obs	Dep Var TON	Predict Value	Std Err Predict	Lower95% Mean	Upper95% Mean	Lower95% Predict	Upper95% Predict
1	47567.0	49731.3	2252.771	44711.8	54750.8	39347.6	60115.1
2	49835.0	49355.4	1997.138	44905.5	53805.3	39234.7	59476.1
3	59435.0	53631.3	2931.689	47099.1	60163.6	42437.7	64825.0
4	48235.0	48137.2	2377.309	42840.2	53434.2	37616.5	58657.9
5	49040.0	49720.9	1935.701	45407.8	54033.9	39659.6	59782.2
6	40829.0	37267.9	1818.660	33215.6	41320.1	27315.6	47220.2
7	42134.0	39046.5	1787.437	35063.8	43029.2	29122.3	48970.6
8	34675.0	35830.4	2335.316	30627.0	41033.8	25356.5	46304.3
9	42681.0	39643.4	1946.152	35307.1	43979.7	29572.1	49714.7
10	39408.0	40570.7	1581.768	37046.3	44095.1	30821.4	50320.0
11	35796.0	44785.7	2273.404	39720.2	49851.2	34379.7	55191.8
12	39293.0	42641.7	1355.788	39620.8	45662.6	33062.9	52220.4
13	42387.0	41374.7	1338.444	38392.5	44357.0	31808.1	50941.4
14	48394.0	47971.8	3531.500	40103.0	55840.5	35949.1	59994.4
15	.	47752.5	3523.598	39901.4	55603.6	35741.4	59763.6
16	.	40674.7	5794.483	27763.7	53585.7	24884.8	56464.6
17	.	41003.6	5621.909	28477.1	53530.1	25526.5	56480.7
18	.	32500.3	10754.34	8538.0	56462.6	6871.8	58128.8
19	.	30746.1	10863.71	6540.1	54952.1	4889.6	56602.6
20	.	25520.0	11886.17	-964.2	52004.2	-2480.7	53520.7

Obs	Residual	Std Err Residual	Student Residual	-2-1-0 1 2	Cook's D
1	-2164.3	3401.179	-0.636	| * |	0.044
2	479.6	3557.308	0.135	| |	0.001
3	5803.7	2836.934	2.046	| |****	1.117
4	97.7893	3315.328	0.029	| |	0.000
5	-680.9	3591.108	-0.190	| |	0.003
6	3561.1	3651.776	0.975	| |*	0.059
7	3087.5	3667.161	0.842	| |*	0.042
8	-1155.4	3345.040	-0.345	| |	0.015
9	3037.6	3585.455	0.847	| |*	0.053
10	-1162.7	3760.453	-0.309	| |	0.004
11	-8989.7	3387.422	-2.654	|*****|	0.793
12	-3348.7	3847.705	-0.870	| * |	0.024
13	1012.3	3853.773	0.263	| |	0.002
14	422.2	2042.426	0.207	| |	0.032
15	.	.	.		.
16	.	.	.		.
17	.	.	.		.
18	.	.	.		.
19	.	.	.		.
20	.	.	.		.

Sum of Residuals -1.5E-8
Sum of Squared Residuals 166429964.42
Predicted Resid SS (Press) 397874609.63

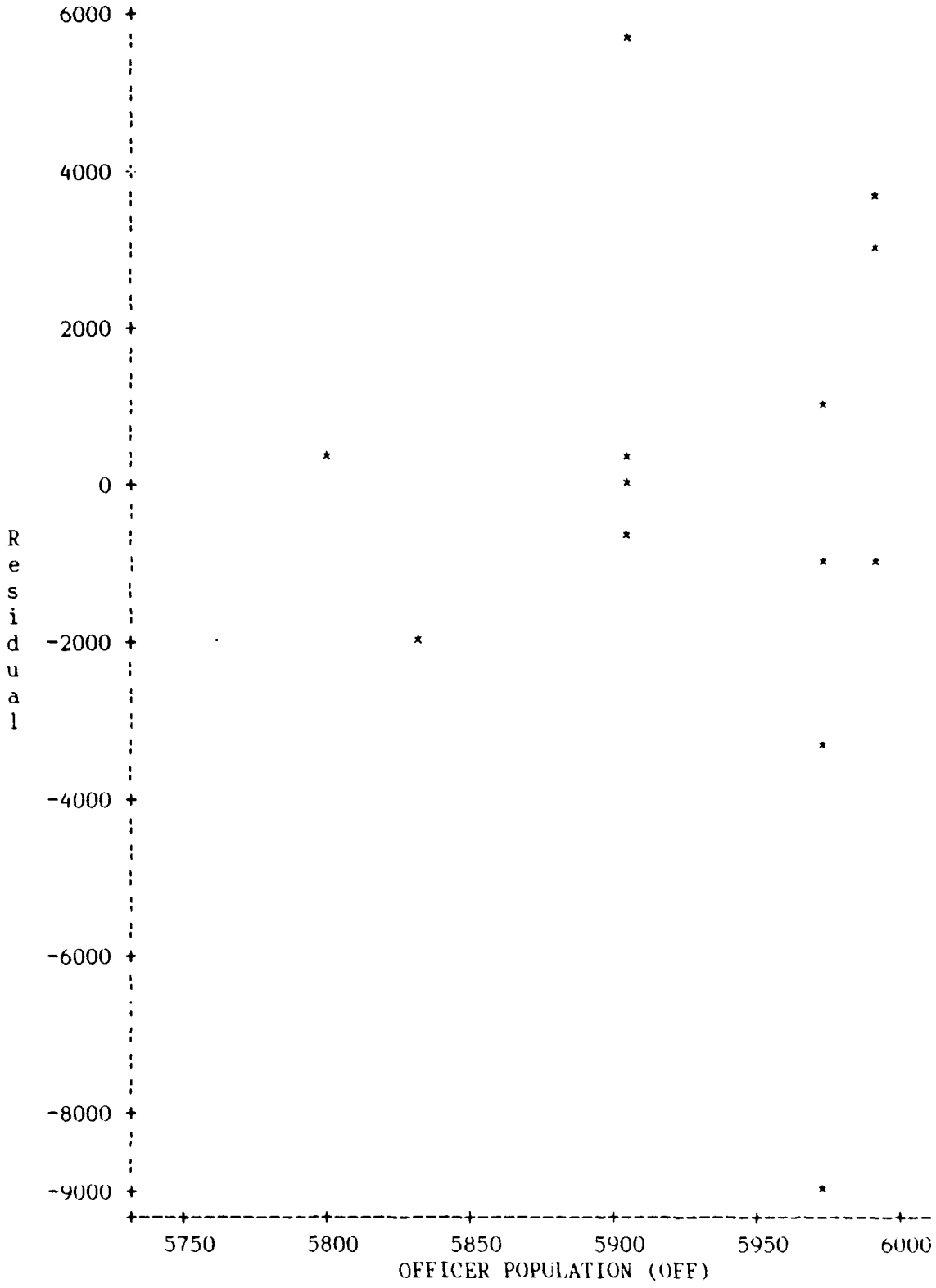

Figure 70. PACAF Multiple Regression MSC Model (Second Order) Residuals versus Officer Population (OFF)

Univariate Procedure

Variable=RESIDUAL

Residual

Moments

N	14	Sum Wgts	14
Mean	-1.1E-9	Sum	-1.54E-8
Std Dev	3578.031	Variance	12802305
Skewness	-0.93707	Kurtosis	2.248215
USS	1.6643E8	CSS	1.6643E8
CV	-3.26E14	Std Mean	956.269
T:Mean=0	-115E-14	Prob>|T|	1.0000
Num ^= 0	14	Num > 0	8
M(Sign)	1	Prob>|M|	0.7905
Sgn Rank	2.5	Prob>|S|	0.9032
W:Normal	0.932305	Prob<W	0.3178

Quantiles(Def=5)

100% Max	5803.661	99%	5803.661
75% Q3	3037.597	95%	5803.661
50% Med	260.0139	90%	3561.107
25% Q1	-1162.72	10%	-3348.68
0% Min	-8989.73	5%	-8989.73
		1%	-8989.73
Range	14793.39		
Q3-Q1	4200.321		
Mode	-8989.73		

Extremes

Lowest	Obs	Highest	Obs
-8989.73(	11)	1012.257(	13)
-3348.68(	12)	3037.597(	9)
-2164.32(	1)	3087.52(	7)
-1162.72(	10)	3561.107(	6)
-1155.41(	8)	5803.661(	3)

Missing Value	.
Count	6
% Count/Nobs	30.00

Appendix I: PACAF MSC Independent Variable Correlation Matrix

Correlation Analysis

10 'VAR' Variables:	A10	B52	C130	C135		
	F4	F15	F16	OFF	AMN	FH

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
A10	20	2549	210.3304	50973	2120	2856
B52	20	1581	141.1555	31617	1325	1782
C130	20	7139	1037	142789	4915	8778
C135	20	3861	374.5776	77211	3246	4608
F4	20	5698	607.5924	113965	4112	6706
F15	20	5451	303.7384	109015	4774	6228
F16	20	6935	2028	138695	4028	11430
OFF	20	5888	98.8291	117753	5727	5995
AMN	20	15498	110.7208	309951	15338	15664
FH	20	41691	1822	833822	36584	44381

Correlation Analysis

Pearson Correlation Coefficients / Prob > |R| under Ho: Rho=0 / Number of Observations

	A10	B52	C130	C135
A10	1.00000 0.0 20	0.13118 0.5815 20	0.13198 0.5791 20	0.19355 0.4136 20
B52	0.13118 0.5815 20	1.00000 0.0 20	-0.60094 0.0051 20	-0.31367 0.1781 20
C130	0.13198 0.5791 20	-0.60094 0.0051 20	1.00000 0.0 20	0.73894 0.0002 20
C135	0.19355 0.4136 20	-0.31367 0.1781 20	0.73894 0.0002 20	1.00000 0.0 20
F4	0.50949 0.0218 20	0.08656 0.7167 20	0.23080 0.3276 20	0.49756 0.0256 20
F15	0.52486 0.0175 20	0.04250 0.8588 20	-0.03142 0.8954 20	-0.11977 0.6150 20
F16	-0.00728 0.9757 20	0.35618 0.1232 20	-0.76353 0.0001 20	-0.84326 0.0001 20
OFF	0.15884 0.5036 20	-0.38857 0.0904 20	0.65899 0.0016 20	0.41469 0.0691 20
AMN	-0.23557 0.3174 20	0.58774 0.0064 20	-0.81130 0.0001 20	-0.60291 0.0049 20
FH	0.68616 0.0008 20	0.00800 0.9733 20	0.32168 0.1666 20	0.21473 0.3633 20

Correlation Analysis

Pearson Correlation Coefficients / Prob > |R| under Ho: Rho=0 / Number of Observations

	F4	F15	F16	OFF	AMN	FH
A10	0.50949	0.52486	-0.00728	0.15884	-0.23557	0.68616
	0.0218	0.0175	0.9757	0.5036	0.3174	0.0008
	20	20	20	20	20	20
B52	0.08656	0.04250	0.35618	-0.38857	0.58774	0.00800
	0.7167	0.8588	0.1232	0.0904	0.0064	0.9733
	20	20	20	20	20	20
C130	0.23080	-0.03142	-0.76353	0.65899	-0.81130	0.32168
	0.3276	0.8954	0.0001	0.0016	0.0001	0.1666
	20	20	20	20	20	20
C135	0.49756	-0.11977	-0.84326	0.41469	-0.60291	0.21473
	0.0256	0.6150	0.0001	0.0691	0.0049	0.3633
	20	20	20	20	20	20
F4	1.00000	0.18727	-0.31341	0.37484	-0.11112	0.68490
	0.0	0.4292	0.1784	0.1034	0.6409	0.0009
	20	20	20	20	20	20
F15	0.18727	1.00000	0.24315	-0.19402	-0.13059	0.39529
	0.4292	0.0	0.3016	0.4124	0.5832	0.0845
	20	20	20	20	20	20
F16	-0.31341	0.24315	1.00000	-0.61016	0.74391	0.03572
	0.1784	0.3016	0.0	0.0043	0.0002	0.8812
	20	20	20	20	20	20
OFF	0.37484	-0.19402	-0.61016	1.00000	-0.66383	0.44059
	0.1034	0.4124	0.0043	0.0	0.0014	0.0519
	20	20	20	20	20	20
AMN	-0.11112	-0.13059	0.74391	-0.66383	1.00000	-0.16132
	0.6409	0.5832	0.0002	0.0014	0.0	0.4968
	20	20	20	20	20	20
FH	0.68490	0.39529	0.03572	0.44059	-0.16132	1.00000
	0.0009	0.0845	0.8812	0.0519	0.4968	0.0
	20	20	20	20	20	20

Appendix J: USAFE Aircraft Flying Hours and Inventory by MD

FISCAL YEAR: 1985

MD	1st Qtr INV.	1st Qtr FLY.HRS.	2nd Qtr INV.	2nd Qtr FLY.HRS.	3rd Qtr INV.	3rd Qtr FLY.HRS.	4th Qtr INV.	4th Qtr FLY.HRS.
A010	113	11591.3	113	11292.3	111	14426.2	111	13062.0
C009		1597.2		1535.6		1645.3		1402.2
C012		773.7		782.0		885.7		815.3
C021		925.7		821.1		942.5		925.9
C023		389.5		1320.5		2631.2		3881.9
C130	37	8686.5	37	7495.3	37	8567.7	37	8566.4
C135	36	5029.5	36	4522.5	36	4697.4	36	5099.8
C140		417.0		487.0		559.4		438.8
F004	148	9413.3	151	11082.8	153	13074.9	153	12479.4
F005	16	1163.2	15	1209.9	17	1295.9	15	1217.0
F015	99	6767.8	101	6864.5	102	8562.9	101	7971.8
F016	154	11087.1	152	10944.3	154	12165.6	156	11449.1
F111	144	9491.5	147	9841.9	144	11378.3	145	10260.9
H001		594.5		531.0		674.9		578.0
H053	7	848.2	7	579.5	8	1154.6	8	996.7
R001		201.9		808.9		672.8		801.5
T039		49.6		0.0		0.0		0.0
TOTAL:		69027.5		70119.1		83335.3		79946.7

1986

MD	1st Qtr INV.	1st Qtr FLY.HRS.	2nd Qtr INV.	2nd Qtr FLY.HRS.	3rd Qtr INV.	3rd Qtr FLY.HRS.	4th Qtr INV.	4th Qtr FLY.HRS.
A010	111	12746.8	112	10217.6	110	13284.3	113	12175.3
C009	5	1562.2	5	2493.6	5	885.9	4	1506.4
C012	6	795.4	6	1132.3	6	1040.8	6	1153.0
C021	4	966.0	6	1162.8	5	1089.9	6	734.1
C023	0	4054.3	17	2774.0	17	4402.5	17	4852.4
C130	37	8820.2	37	10724.4	36	7125.5	38	8482.9
C135	36	4829.3	43	4156.0	37	4666.9	31	4604.8
C140	4	425.2	4	399.9	4	331.9	4	332.6
F004	141	9237.9	129	9120.3	118	9877.8	103	8650.0
F005	15	1207.4	16	1177.8	16	1287.7	15	1335.1
F015	96	7284.6	92	6570.3	90	8023.4	91	7387.7
F016	145	11755.0	154	11883.1	149	14014.4	174	12841.7
F111	150	10123.9	151	10409.8	99	12009.3	141	11019.9
H001	7	463.6	7	606.7	3	671.5	6	640.1
H053	9	536.8	11	729.3	9	1189.3	11	914.8
R001	5	743.3	6	854.5	8	771.7	9	996.8
TOTAL		75551.9		74412.4		80672.8		77627.6

1987

MD	1st Qtr INV.	1st Qtr FLY.HRS.	2nd Qtr INV.	2nd Qtr FLY.HRS.	3rd Qtr INV.	3rd Qtr FLY.HRS.	4th Qtr INV.	4th Qtr FLY.HRS.
A010	116	11215.7	91	10566.3	113	13777.1	39	12720.9
C009	3	1491.7	4	1391.8	4	1588.1	5	1552.5
C012	6	1002.1	2	1139.8	6	1224.1	6	1256.6
C020	0	0.0	0	0.0	0	235.1	1	545.6
C021	6	1083.8	2	1132.0	6	1235.2	4	1213.2
C023	17	4158.5	18	3187.0	18	4105.4	17	4550.3
C130	33	8201.6	41	7079.1	42	9082.8	41	8982.4
C135	34	4529.0	30	4124.3	36	4612.4	31	4900.9
C140	4	291.8	1	385.6	3	252.0	2	159.9
F004	43	7060.8	100	6452.4	96	6750.5	74	5516.4
F005	10	1139.6	11	1163.1	16	1271.4	15	1243.8
F015	2	6999.0	80	6780.1	92	7811.0	42	7430.1
F016	158	15209.6	193	15380.0	209	19900.8	172	16697.8
F111	112	10939.6	127	9966.8	143	12716.8	121	11866.6
H001	7	578.3	4	560.3	6	643.2	7	655.7
H053	6	671.1	8	566.6	7	904.4	7	1025.6
L109	2	0.0	13	510.9	36	160.2	31	0.0
R001	3010	735.6	6	982.3	10	1248.8	9	1175.2
TOTAL:		75307.8		71368.4		87519.3		81493.5

1988

MD	1st Qtr INV.	1st Qtr FLY.HRS.	2nd Qtr INV.	2nd Qtr FLY.HRS.	3rd Qtr INV.	3rd Qtr FLY.HRS.	4th Qtr INV.	4th Qtr FLY.HRS.
A010	131	11924.4	111	10946.8	111	9874.8	110	11695.2
C009	1	1512.2	1	1473.3	5	1876.1	5	1429.8
C012	6	1023.9	6	1401.5	6	981.0	6	784.7
C020	2	393.4	3	589.6	3	516.8	3	581.5
C021	6	1098.2	6	1055.5	6	1090.2	6	1005.5
C023	0	3998.4	11	4266.9	11	2142.1	16	3023.6
C130	29	9592.8	34	8583.7	47	9088.5	46	8035.1
C135	25	3984.4	34	3764.5	37	4674.7	36	4771.3
F004	62	3732.5	46	3693.9	54	3913.5	54	4522.4
F005	15	1104.3	13	989.1	11	709.6	3	5.4
F015	112	6611.7	96	6950.4	94	7308.5	94	7329.5
F016	235	17330.2	233	19349.4	230	18275.9	216	18468.5
F111	144	11045.1	145	10956.7	142	11164.4	137	10955.9
H001	7	508.4	6	571.9	5	673.5	4	466.8
H053	10	582.0	10	696.0	4	489.4	4	298.4
R001	6	917.7	0	987.0	9	1332.8	10	1167.9
T039	0	73.8	0	157.7	0	218.7	1	210.0
TOTAL:		75433.4		76433.9		74330.5		74751.5

1989

MD	1st Qtr INV.	1st Qtr FLY.HRS.	2nd Qtr INV.	2nd Qtr FLY.HRS.	3rd Qtr INV.	3rd Qtr FLY.HRS.	4th Qtr INV.	4th Qtr FLY.HRS.
A010	111	10358.5	136	11447.5	85	13193.0	107	13590.9
C009	5	1461.2	4	1333.4	5	1286.8	5	1430.9
C012	6	1076.0	6	1037.1	6	1063.6	6	1040.0
C020	3	531.2	2	577.8	2	543.8	2	653.0
C021	6	1334.2	6	1314.0	5	1376.2	6	1260.5
C023	11	3265.1	17	3362.4	17	3495.6	0	3394.9
C130	47	8879.5	46	8015.0	46	8825.9	40	8897.5
C135	33	4536.5	30	3953.3	31	4430.5	35	4536.9
F004	52	2677.5	51	3541.5	53	4970.2	52	4309.1
F015	94	5390.1	95	7019.9	79	9055.7	95	7212.7
F016	238	16595.3	234	18843.6	210	23118.7	242	22608.5
F111	140	10073.5	143	11594.1	116	11863.9	138	11441.5
H001	5	341.6	4	408.2	4	461.0	4	442.0
H053	4	374.1	4	345.7	4	559.1	4	497.7
R001	10	1100.2	11	1001.4	11	1070.5	12	1414.6
T039	1	113.8	1	103.5	1	123.5	1	61.6
T043	0	28.7	1	253.4	1	192.3	1	218.6
TOTAL:		68137.0		74151.8		85630.3		83010.9

1990

MD	1st Qtr INV.	1st Qtr FLY.HRS.
A010	107	10639
C009	5	1516
C012	5	798
C020	2	526
C021	6	926
C023	16	2864
C130	43	7855
C135	29	3595
F004	54	3487
F015	94	6194
F016	242	18831
F111	138	10188
H001	4	309
H053	4	439
R001	4	0
T039	1	89
T043	1	159
TOTAL:		68417.6

Appendix K: USAFE MSC Multiple Regression Model SAS Output

First Order Model

Dependent Variable: TON

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	3	915845924.65	305281974.88	7.136	0.0076
Error	10	427834146.85	42783414.685		
C Total	13	1343680071.5			

Root MSE	6540.90320	R-square	0.6816
Dep Mean	73417.50000	Adj R-sq	0.5861
C.V.	8.90919		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	630601	175849.60077	3.586	0.0050
C130	1	4.464226	1.93861574	2.303	0.0440
F4	1	2.478257	0.68166009	3.636	0.0046
OFF	1	-59.980187	17.29779539	-3.468	0.0060

Variable	DF	Variance Inflation
INTERCEP	1	0.00000000
C130	1	1.03680183
F4	1	1.20108991
OFF	1	1.16148972

Durbin-Watson D	2.123
(For Number of Obs.)	14
1st Order Autocorrelation	-0.155

Obs	Dep Var TON	Predict Value	Std Err Predict	Lower95% Mean	Upper95% Mean	Lower95% Predict	Upper95% Predict
1	96200.0	91403.3	4919.119	80442.8	102364	73167.7	109639
2	71083.0	73008.3	3026.709	66264.3	79752.3	56949.5	89067.2
3	83702.0	82730.7	3525.498	74875.3	90586.0	66174.3	99287.0
4	77001.0	81251.6	3188.892	74146.3	88357.0	65037.7	97465.5
5	78732.0	74351.0	1992.410	69911.6	78790.4	59115.8	89586.3
6	80227.0	80041.8	4763.798	69427.3	90656.3	62012.0	98071.6
7	73784.0	65851.1	3495.006	58063.7	73638.5	49326.9	82375.3
8	56587.0	68868.2	2234.929	63888.5	73848.0	53466.8	84269.6
9	57977.0	63673.3	2756.989	57530.3	69816.3	47857.5	79489.2
10	64120.0	62076.1	3792.434	53626.0	70526.2	45229.4	78922.7
11	72994.0	71756.4	2187.709	66881.9	76631.0	56388.7	87124.2
12	74368.0	68251.8	2601.266	62455.8	74047.9	52567.5	83936.2
13	70603.0	66553.8	3799.897	58087.1	75020.6	49698.8	83408.8
14	70467.0	78027.5	4829.900	67265.7	88789.2	59910.6	96144.3
15	.	80827.1	4817.011	70094.1	91560.1	62727.3	98926.9
16	.	77635.5	4872.527	66778.8	88492.3	59462.1	95809.0
17	.	76831.0	5052.622	65572.9	88089.0	58415.0	95246.9
18	.	68757.2	4095.383	59632.0	77882.3	51562.0	85952.3
19	.	75914.6	3336.487	68480.4	83348.8	59553.9	92275.3
20	.	74597.9	3523.494	66747.0	82448.8	58043.7	91152.1

Obs	Residual	Std Err Residual	Student Residual	-2-1-0 1 2	Cook's D
1	4796.7	4311.112	1.113	| |** |	0.403
2	-1925.3	5798.487	-0.332	| | |	0.008
3	971.3	5509.472	0.176	| | |	0.003
4	-4250.6	5710.900	-0.744	| *| |	0.043
5	4381.0	6230.065	0.703	| |* |	0.013
6	185.2	4482.147	0.041	| | |	0.000
7	7932.9	5528.865	1.435	| |** |	0.206
8	-12281.2	6147.236	-1.998	| ***| |	0.132
9	-5696.3	5931.477	-0.960	| *| |	0.050
10	2043.9	5329.246	0.384	| | |	0.019
11	1237.6	6164.199	0.201	| | |	0.001
12	6116.2	6001.402	1.019	| |** |	0.049
13	4049.2	5323.927	0.761	| |* |	0.074
14	-7560.5	4410.836	-1.714	| ***| |	0.881
15	.	.	.		.
16	.	.	.		.
17	.	.	.		.
18	.	.	.		.
19	.	.	.		.
20	.	.	.		.

Sum of Residuals 0
Sum of Squared Residuals 427834146.85
Predicted Resid SS (Press) 926934269.27

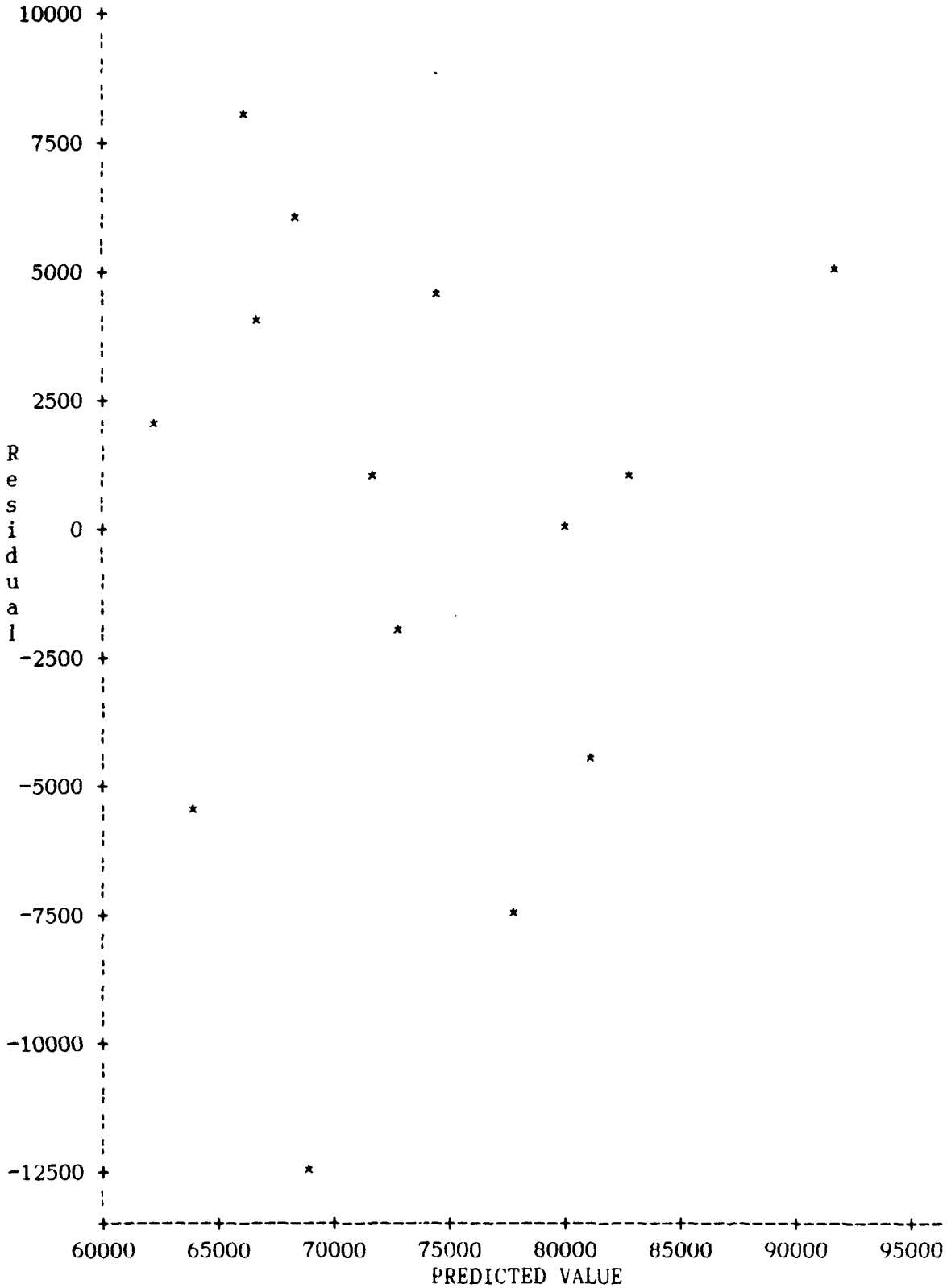

Figure 71. USAFE Multiple Regression MSC Model Residuals versus Predicted Values

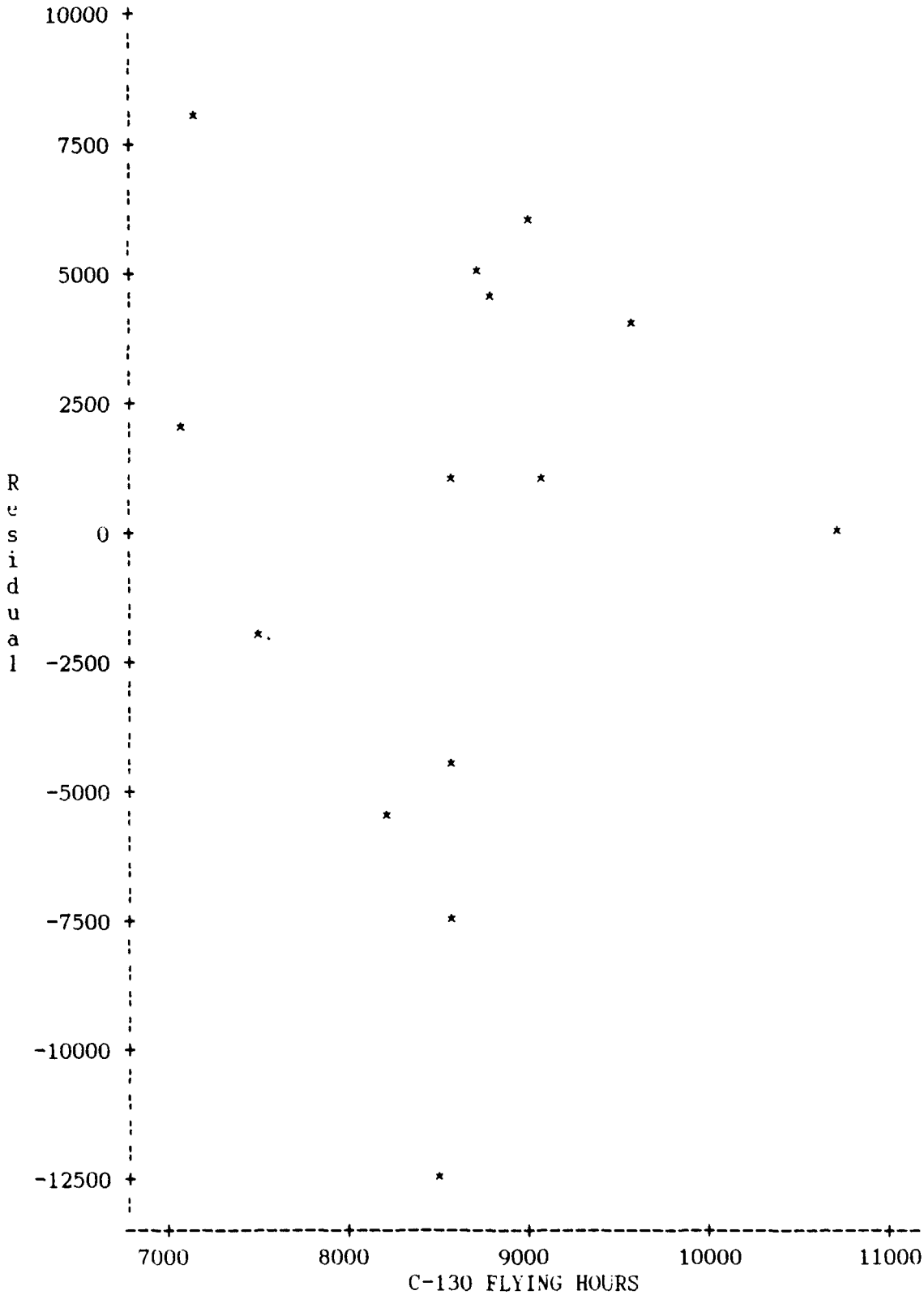

Figure 72. USAFE Multiple Regression MSC Model Residuals versus C-130 Flying Hours

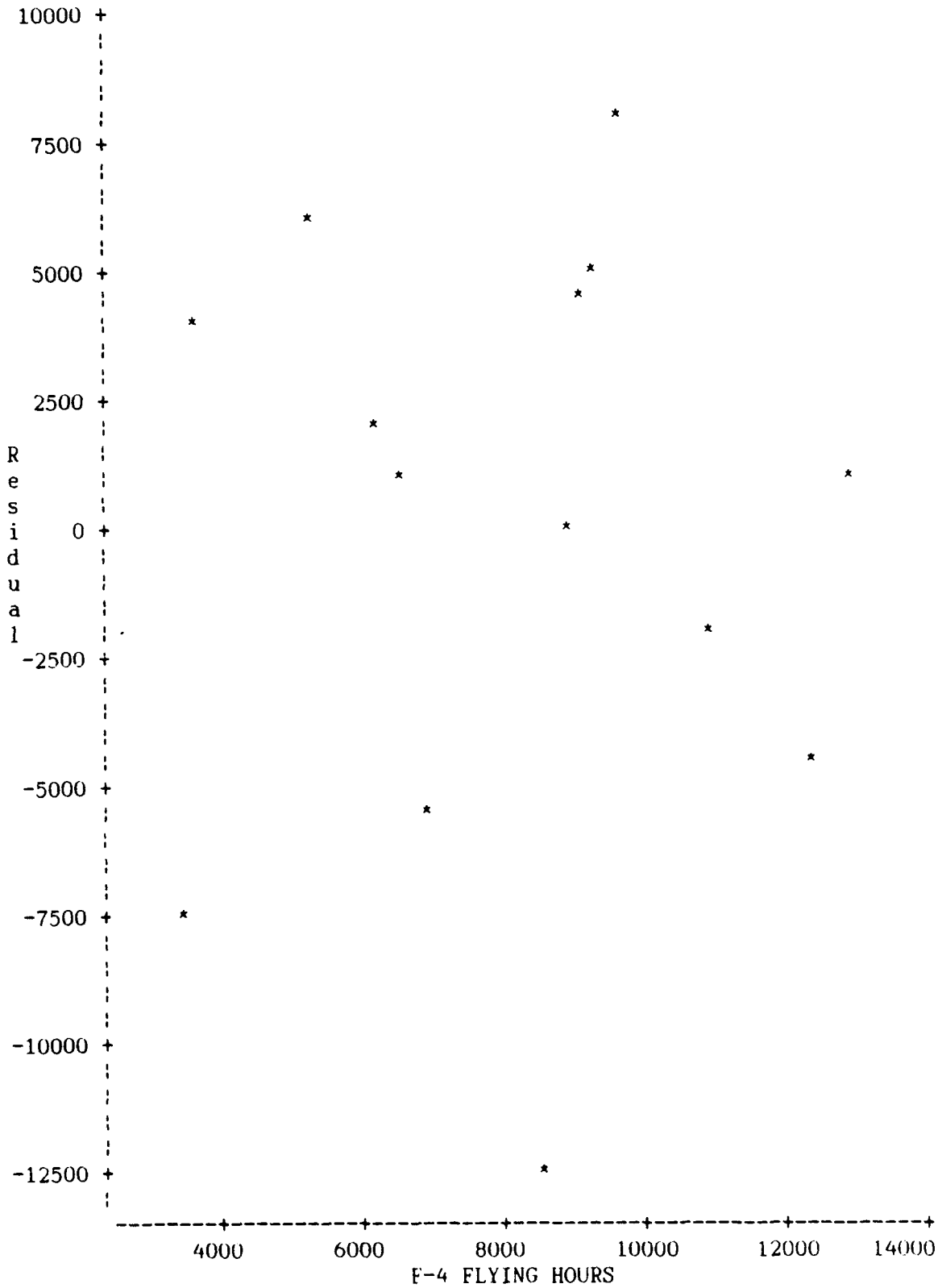

Figure 73. USAFE Multiple Regression MSC Model Residuals versus F-4 Flying Hours

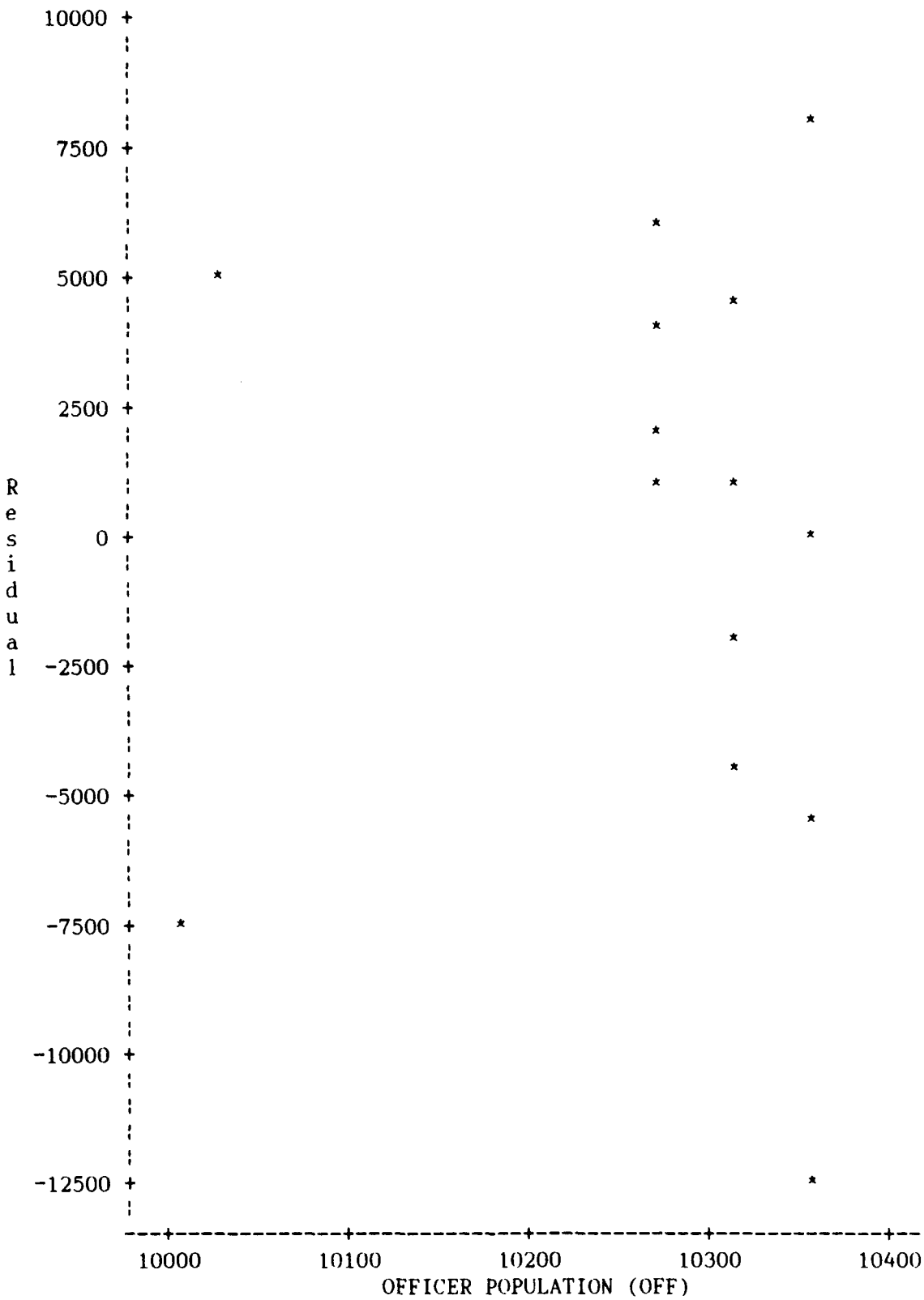

Figure 74. USAFE Multiple Regression MSC Model Residuals versus Officer Population (OFF)

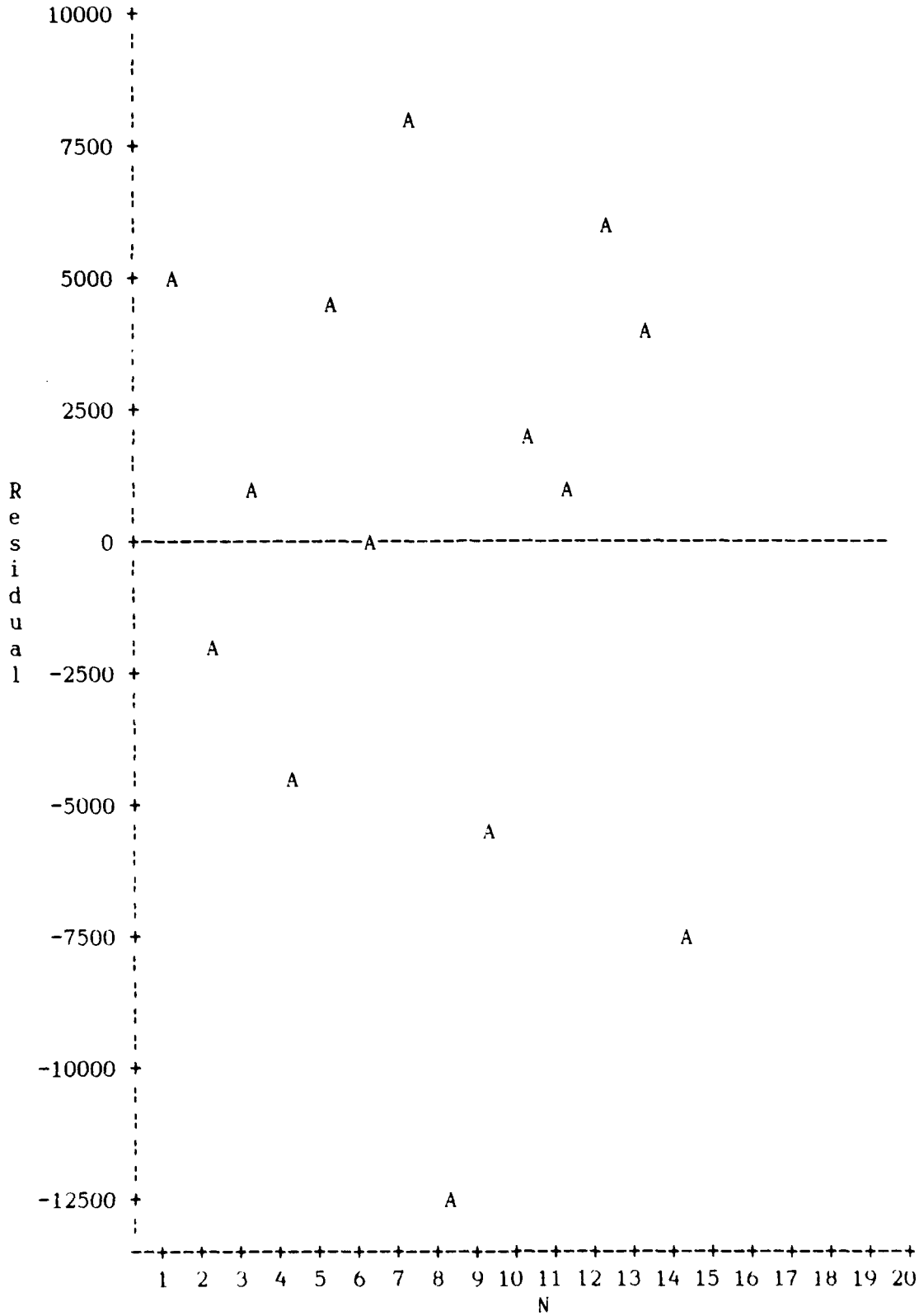

Figure 75. USAFE Multiple Regression MSC Model Residuals versus Time (N)

Univariate Procedure

Variable=RESIDUAL Residual

Moments

N	14	Sum Wgts	14
Mean	0	Sum	0
Std Dev	5736.752	Variance	32910319
Skewness	-0.73824	Kurtosis	-0.00419
USS	4.2783E8	CSS	4.2783E8
CV	.	Std Mean	1533.211
T:Mean=0	0	Prob>|T|	1.0000
Num ^= 0	14	Num > 0	9
M(Sign)	2	Prob>|M|	0.4240
Sgn Rank	5.5	Prob>|S|	0.7609
W:Normal	0.951712	Prob<W	0.5565

Quantiles(Def=5)

100% Max	7932.917	99%	7932.917
75% Q3	4380.97	95%	7932.917
50% Med	1104.454	90%	6116.151
25% Q1	-4250.63	10%	-7560.45
0% Min	-12281.2	5%	-12281.2
		1%	-12281.2
Range	20214.13		
Q3-Q1	8631.595		
Mode	-12281.2		

Extremes

Lowest	Obs	Highest	Obs
-12281.2(	8)	4049.183(	13)
-7560.45(	14)	4380.97(	5)
-5696.34(	9)	4796.689(	1)
-4250.63(	4)	6116.151(	12)
-1925.31(	2)	7932.917(	7)

Missing Value	.
Count	6
% Count/Nobs	30.00

Second Order Model

Dependent Variable: TON

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	4	1011343900.9	252835975.22	6.847	0.0082
Error	9	332336170.6	36926241.178		
C Total	13	1343680071.5			

Root MSE	6076.69657	R-square	0.7527
Dep Mean	73417.50000	Adj R-sq	0.6427
C.V.	8.27690		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	-30448299	19326429.051	-1.575	0.1496
C130	1	4.580591	1.80248554	2.541	0.0316
F4	1	2.706706	0.64902005	4.170	0.0024
OFF	1	6050.498256	3799.7022610	1.592	0.1458
SOFF	1	-0.300314	0.18674404	-1.608	0.1423

Variable	DF	Variance Inflation
INTERCEP	1	0.00000000
C130	1	1.03847538
F4	1	1.26152643
OFF	1	64934.176488
SOFF	1	64979.893180

Collinearity Diagnostics(intercept adjusted)

Number	Eigenvalue	Condition Number	Var Prop C130	Var Prop F4	Var Prop OFF	Var Prop SOFF
1	2.23588	1.00000	0.0037	0.0571	0.0000	0.0000
2	1.06728	1.44739	0.7059	0.1112	0.0000	0.0000
3	0.69684	1.79126	0.2888	0.7842	0.0000	0.0000
4	7.69743E-6	538.95360	0.0016	0.0476	1.0000	1.0000

Durbin-Watson D	2.292
(For Number of Obs.)	14
1st Order Autocorrelation	-0.208

Obs	Dep Var TON	Predict Value	Std Err Predict	Lower95% Mean	Upper95% Mean	Lower95% Predict	Upper95% Predict
1	96200.0	91417.8	4570.019	81079.6	101756	74217.6	108618
2	71083.0	74121.5	2895.851	67570.6	80672.4	58893.8	89349.2
3	83702.0	84423.7	3440.326	76641.0	92206.3	68626.9	100220
4	77001.0	82808.6	3116.760	75757.9	89859.2	67359.3	98257.8
5	78732.0	75196.9	1924.293	70843.8	79550.0	60777.6	89616.2
6	80227.0	77058.0	4798.894	66202.1	87914.0	59541.8	94574.3
7	73784.0	62621.5	3817.845	53984.8	71258.1	46386.9	78856.0
8	56587.0	65516.2	2942.064	58860.7	72171.7	50243.2	80789.2
9	57977.0	59925.4	3462.949	52091.6	67759.2	44103.4	75747.4
10	64120.0	64930.9	3945.248	56006.1	73855.8	48541.3	81320.6
11	72994.0	74912.5	2825.286	68521.2	81303.7	59752.8	90072.2
12	74368.0	71114.3	3001.415	64324.6	77904.0	55782.4	86446.3
13	70603.0	69079.7	3863.867	60339.0	77820.5	52789.6	85369.9
14	70467.0	74718.1	4936.495	63550.9	85885.3	57007.2	92429.0
15	.	77626.8	4897.678	66547.4	88706.2	59971.2	95282.4
16	.	74451.8	4940.705	63275.1	85628.6	56734.9	92168.7
17	.	73324.0	5175.874	61615.2	85032.7	55266.8	91381.2
18	.	72762.0	4547.277	62475.3	83048.8	55592.7	89931.3
19	.	80340.2	4145.030	70963.4	89717.0	63700.1	96980.3
20	.	78880.9	4219.981	69334.5	88427.2	62144.6	95617.1

Obs	Residual	Std Err Residual	Student Residual	-2-1-0 1 2	Cook's D
1	4782.2	4005.143	1.194	| |** |	0.371
2	-3038.5	5342.311	-0.569	| *| |	0.019
3	-721.7	5009.032	-0.144	| | |	0.002
4	-5807.6	5216.517	-1.113	| **| |	0.088
5	3535.1	5763.969	0.613	| |* |	0.008
6	3169.0	3727.848	0.850	| |* |	0.240
7	11162.5	4727.611	2.361	| |**** |	0.727
8	-8929.2	5317.001	-1.679	| ***| |	0.173
9	-1948.4	4993.418	-0.390	| | |	0.015
10	-810.9	4621.824	-0.175	| | |	0.004
11	-1918.5	5379.963	-0.357	| | |	0.007
12	3253.7	5283.725	0.616	| |* |	0.024
13	1523.3	4690.072	0.325	| | |	0.014
14	-4251.1	3543.622	-1.200	| **| |	0.559
15	.	.	.		.
16	.	.	.		.
17	.	.	.		.
18	.	.	.		.
19	.	.	.		.
20	.	.	.		.

Sum of Residuals 1.036E-9
Sum of Squared Residuals 332336170.37
Predicted Resid SS (Press) 959980694.73

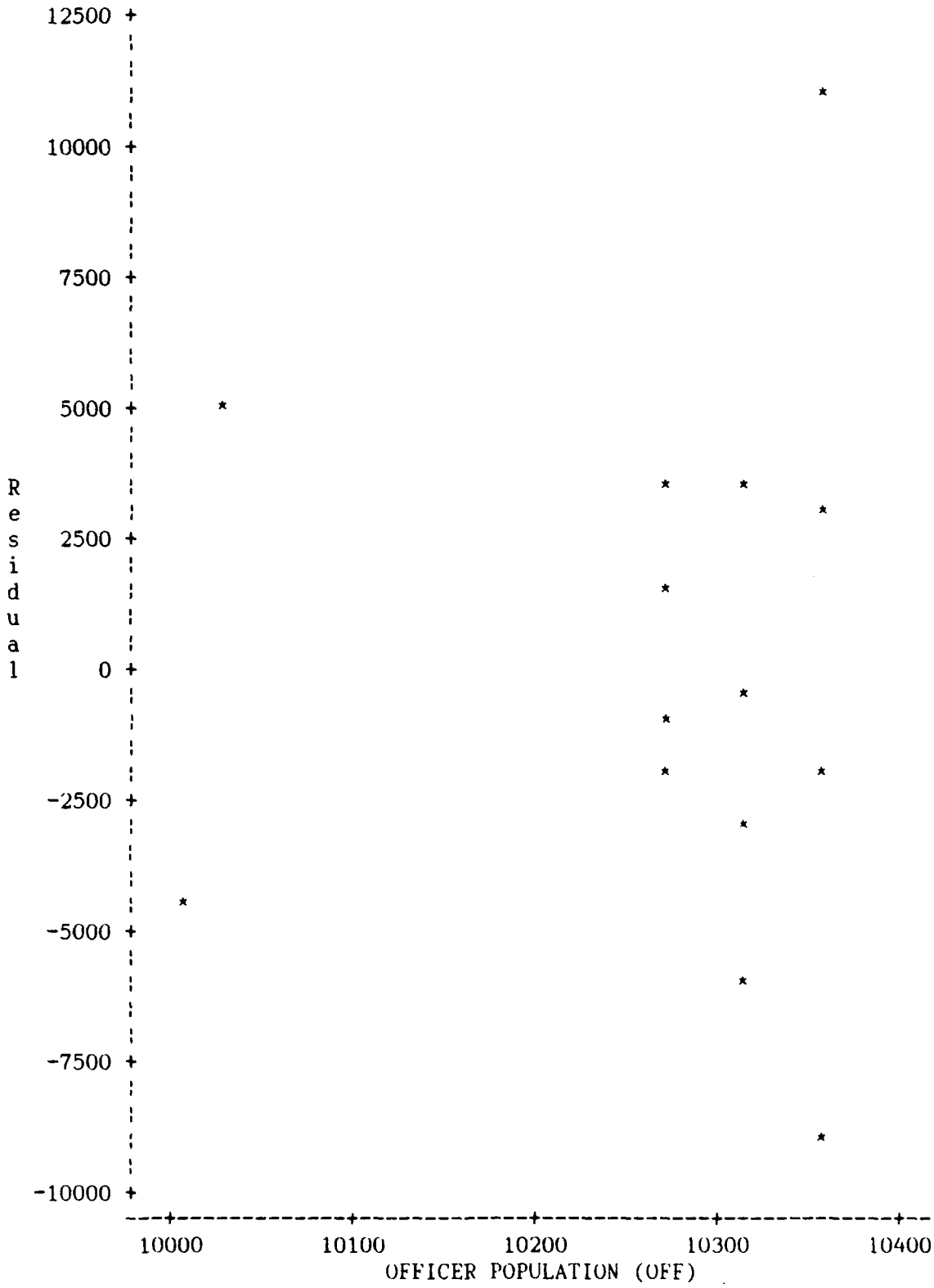

Figure 76. USAFE Multiple Regression MSC Model (Second Order) Residuals versus Officer Population (OFF)

Univariate Procedure

Variable=RESIDUAL Residual

Moments

N	14	Sum Wgts	14
Mean	9.98E-11	Sum	1.397E-9
Std Dev	5056.117	Variance	25564321
Skewness	0.406986	Kurtosis	0.720298
USS	3.3234E8	CSS	3.3234E8
CV	5.067E15	Std Mean	1351.304
T:Mean=0	7.38E-14	Prob>|T|	1.0000
Num ^= 0	14	Num > 0	6
M(Sign)	-1	Prob>|M|	0.7905
Sgn Rank	-0.5	Prob>|S|	1.0000
W:Normal	0.973954	Prob<W	0.8865

Quantiles(Def=5)

100% Max	11162.54	99%	11162.54
75% Q3	3253.681	95%	11162.54
50% Med	-766.29	90%	4782.243
25% Q1	-3038.5	10%	-5807.58
0% Min	-8929.2	5%	-8929.2
		1%	-8929.2
Range	20091.73		
Q3-Q1	6292.18		
Mode	-8929.2		

Extremes

Lowest	Obs	Highest	Obs
-8929.2(	8)	3168.967(	6)
-5807.58(	4)	3253.681(	12)
-4251.11(	14)	3535.09(	5)
-3038.5(	2)	4782.243(	1)
-1948.39(	9)	11162.54(	7)

Missing Value	.
Count	6
% Count/Nobs	30.00

Appendix L: USAFE MSC Independent Variable Correlation Matrix

Correlation Analysis

10 'VAR' Variables: A10 C130 C135 F4 F15 F16 F111 OFF AMN FH

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
A10	20	12005	1300	240098	9874	14426
C130	20	8586	824.8657	171723	7079	10724
C135	20	4521	360.3656	90415	3764	5099
F4	20	7003	3177	140067	2677	13074
F15	20	7266	781.6823	145322	5390	9055
F16	20	15896	3822	317910	10944	23118
F111	20	10955	838.5357	219108	9491	12716
OFF	20	10206	140.2270	204123	10004	10354
AMN	20	28239	206.4878	564776	27951	28610
FH	20	76913	5428	1538251	68137	87519

Correlation Analysis

Pearson Correlation Coefficients / Prob > |R| under Ho: Rho=0 / Number of Observations

	A10	C130	C135	F4
A10	1.00000 0.0 20	-0.07831 0.7428 20	0.39907 0.0813 20	0.39670 0.0833 20
C130	-0.07831 0.7428 20	1.00000 0.0 20	-0.06920 0.7719 20	-0.15546 0.5128 20
C135	0.39907 0.0813 20	-0.06920 0.7719 20	1.00000 0.0 20	0.49552 0.0263 20
F4	0.39670 0.0833 20	-0.15546 0.5128 20	0.49552 0.0263 20	1.00000 0.0 20
F15	0.73328 0.0002 20	-0.14296 0.5476 20	0.30368 0.1930 20	0.38319 0.0954 20
F16	0.09218 0.6991 20	0.09166 0.7007 20	-0.37228 0.1060 20	-0.80098 0.0001 20
F111	0.53614 0.0148 20	0.06372 0.7896 20	-0.09432 0.6925 20	-0.25568 0.2766 20
OFF	0.33781 0.1452 20	-0.04828 0.8398 20	0.10552 0.6579 20	0.65388 0.0018 20
AMN	0.43033 0.0582 20	-0.09078 0.7035 20	0.37328 0.1050 20	0.72129 0.0003 20
FH	0.81559 0.0001 20	0.14869 0.5315 20	0.15646 0.5101 20	0.11947 0.6159 20

Correlation Analysis

Pearson Correlation Coefficients / Prob > |R| under Ho: Rho=0
/ Number of Observations

	F15	F16	F111	OFF	AMN	FH
A10	0.73328	0.09218	0.53614	0.33781	0.43033	0.81559
	0.0002	0.6991	0.0148	0.1452	0.0582	0.0001
	20	20	20	20	20	20
C130	-0.14296	0.09166	0.06372	-0.04828	-0.09078	0.14869
	0.5476	0.7007	0.7896	0.8398	0.7035	0.5315
	20	20	20	20	20	20
C135	0.30368	-0.37228	-0.09432	0.10552	0.37328	0.15646
	0.1930	0.1060	0.6925	0.6579	0.1050	0.5101
	20	20	20	20	20	20
F4	0.38319	-0.80098	-0.25568	0.65388	0.72129	0.11947
	0.0954	0.0001	0.2766	0.0018	0.0003	0.6159
	20	20	20	20	20	20
F15	1.00000	0.15577	0.56575	0.23977	0.32944	0.80272
	0.0	0.5120	0.0093	0.3086	0.1561	0.0001
	20	20	20	20	20	20
F16	0.15577	1.00000	0.61660	-0.53702	-0.51689	0.41797
	0.5120	0.0	0.0038	0.0146	0.0196	0.0667
	20	20	20	20	20	20
F111	0.56575	0.61660	1.00000	0.07598	-0.27078	0.81824
	0.0093	0.0038	0.0	0.7502	0.2482	0.0001
	20	20	20	20	20	20
OFF	0.23977	-0.53702	0.07598	1.00000	0.38588	0.26475
	0.3086	0.0146	0.7502	0.0	0.0929	0.2593
	20	20	20	20	20	20
AMN	0.32944	-0.51689	-0.27078	0.38588	1.00000	0.07005
	0.1561	0.0196	0.2482	0.0929	0.0	0.7692
	20	20	20	20	20	20
FH	0.80272	0.41797	0.81824	0.26475	0.07005	1.00000
	0.0001	0.0667	0.0001	0.2593	0.7692	0.0
	20	20	20	20	20	20

Appendix M: PACAF Transformed Network Data

FY/QTR	MSC	MAC	A10	C130	C135	F4	F15
85/1	0.5341	0.6045	0.6020	0.7603	0.8509	0.7114	0.5840
85/2	0.6067	0.7101	0.5050	0.6467	0.6445	0.7351	0.3760
85/3	0.9139	0.6435	0.8560	0.7396	0.6483	0.6772	0.5585
85/4	0.5555	0.6824	0.4050	0.8584	0.6253	0.3834	0.3605
86/1	0.5813	0.6144	0.5350	0.7632	0.7731	0.7378	0.5665
86/2	0.3185	0.5451	0.5090	0.6221	0.4125	0.6314	0.5430
86/3	0.3603	0.8685	0.6550	0.6134	0.4872	0.5698	0.2790
86/4	0.1216	0.7360	0.3910	0.8955	0.4243	0.4735	0.4505
87/1	0.3778	0.6403	0.7040	0.8672	0.5320	0.5994	0.5965
87/2	0.2731	0.7243	0.4640	0.5550	0.4611	0.7012	0.5115
87/3	0.1575	0.8888	0.8100	0.4840	0.4477	0.8166	0.5035
87/4	0.2694	0.7251	0.6340	0.5402	0.4163	0.5938	0.4640
88/1	0.3684	0.5904	0.5300	0.4641	0.3144	0.6277	0.3445
88/2	0.5606	0.5349	0.7190	0.3661	0.4616	0.9018	0.6025
88/3	0.7756	0.3208	0.7010	0.3638	0.4669	0.7335	0.3205
88/4	0.3640	0.1016	0.1200	0.4631	0.4648	0.3600	0.1370
89/1	0.3905	0.0843	0.1470	0.3749	0.2013	0.6348	0.4385
89/2	0.4548	0.1597	0.8010	0.2977	0.1283	0.4895	0.8640
89/3	0.4203	0.1416	0.6570	0.4093	0.1608	0.3535	0.5330
89/4	0.3312	0.0843	0.2280	0.1031	0.1245	0.1037	0.4740
90/1	0.1629	0.2080	-0.2410	-0.0087	-0.0125	0.2686	0.5880

FY/QTR	F16	B52	OFF	AMN
85/1	0.1028	0.3928	0.4627	0.1760
85/2	0.1225	0.7336	0.6520	0.3340
85/3	0.1438	0.5576	0.6520	0.3340
85/4	0.2486	0.2504	0.6520	0.3340
86/1	0.2760	0.1400	0.6520	0.3340
86/2	0.2696	0.4824	0.8867	0.2700
86/3	0.2579	0.3432	0.8867	0.2700
86/4	0.2539	0.1688	0.8867	0.2700
87/1	0.3187	0.5928	0.8867	0.2700
87/2	0.2929	0.8152	0.8360	0.5060
87/3	0.4343	0.8120	0.8360	0.5060
87/4	0.5278	0.5592	0.8360	0.5060
88/1	0.4465	0.2360	0.8360	0.5060
88/2	0.5906	0.7336	0.3827	0.8280
88/3	0.4594	0.8712	0.3827	0.8280
88/4	0.3568	0.7144	0.3827	0.8280
89/1	0.6166	0.5448	0.3827	0.8280
89/2	0.6904	0.6328	0.1720	0.6580
89/3	0.8430	0.6648	0.1720	0.6580
89/4	0.6174	0.7416	0.1720	0.6580
90/1	0.7525	0.2312	0.1720	0.6580

PACAF TRANSFORMATION EQUATIONS

Var.	Transformation Equation
MSC	Transformed MSC = (MSC - 34000)(.4/12500) + .1
MAC	Transformed MAC = (MAC - 3900)(.4/1500) + .1
A10	Transformed A10 = (A10 - 2100)(.4/400) + .1
C130	Transformed C130 = (C130 - 4900)(.4/1950) + .1
C135	Transformed C135 = (C135 - 3200)(.4/750) + .1
F4	Transformed F4 = (F4 - 4100)(.4/1300) + .1
F15	Transformed F15 = (F15 - 4700)(.4/800) + .1
F16	Transformed F16 = (F16 - 4000)(.4/4000) + .1
B52	Transformed B52 = (B52 - 1300)(.4/250) + .1
OFF	Transformed OFF = (OFF - 5700)(.4/150) + .1
AMN	Transformed AMN = (AMN - 15300)(.4/200) + .1

Appendix N: USAFE Transformed Network Data

FY/QTR	MSC	MAC	A10	C130	C135	F4	F15
85/1	0.9040	0.6262	0.4049	0.4372	0.8594	0.6191	0.4088
85/2	0.4017	0.6502	0.3540	0.1990	0.5697	0.7462	0.4293
85/3	0.6540	0.7814	0.8874	0.4134	0.6697	0.8980	0.7867
85/4	0.5200	0.6403	0.6552	0.4132	0.8994	0.8527	0.6623
86/1	0.5546	0.7219	0.6014	0.4640	0.7451	0.6057	0.5177
86/2	0.5845	0.6413	0.1710	0.8448	0.3606	0.5968	0.3674
86/3	0.4557	0.8213	0.6930	0.1250	0.6520	0.6544	0.6733
86/4	0.1117	0.8520	0.5043	0.3964	0.6166	0.5610	0.5394
87/1	0.1395	0.5691	0.3409	0.3402	0.5737	0.4398	0.4577
87/2	0.2624	0.5859	0.2304	0.1158	0.3423	0.3935	0.4116
87/3	0.4399	0.9693	0.7769	0.5164	0.6211	0.4162	0.6286
87/4	0.4674	0.9267	0.5970	0.4964	0.7857	0.3222	0.5484
88/1	0.3921	0.7330	0.4615	0.6184	0.2623	0.1862	0.3760
88/2	0.3893	0.4667	0.2951	0.4166	0.1366	0.1833	0.4474
88/3	0.6097	0.2592	0.1126	0.5176	0.6566	0.2000	0.5227
88/4	0.5138	0.2483	0.4226	0.3070	0.7120	0.2464	0.5272
89/1	0.4569	0.3510	0.1950	0.4758	0.5777	0.1059	0.1189
89/2	0.2571	0.1008	0.3803	0.3030	0.2446	0.1717	0.4619
89/3	0.4761	0.1416	0.6775	0.4650	0.5171	0.2806	0.8905
89/4	0.7523	0.1691	0.7451	0.4794	0.5777	0.2302	0.5025
90/1	0.2731	0.0554	0.2428	0.2710	0.0400	0.1676	0.2882

FY/QTR	F16	F111	OFF	AMN
85/1	0.1122	0.1099	0.1500	0.8573
85/2	0.1029	0.1948	0.7240	0.7080
85/3	0.1823	0.5674	0.7240	0.7080
85/4	0.1357	0.2964	0.7240	0.7080
86/1	0.1556	0.2632	0.7240	0.7080
86/2	0.1639	0.3325	0.8080	0.4957
86/3	0.3025	0.7204	0.8080	0.4957
86/4	0.2262	0.4804	0.8080	0.4957
87/1	0.3803	0.4610	0.8080	0.4957
87/2	0.3914	0.2251	0.6440	0.3027
87/3	0.6854	0.8918	0.6440	0.3027
87/4	0.4770	0.6857	0.6440	0.3027
88/1	0.5182	0.4867	0.6440	0.3027
88/2	0.6495	0.4651	0.1080	0.1544
88/3	0.5797	0.5155	0.1080	0.1544
88/4	0.5922	0.4648	0.1080	0.1544
89/1	0.4704	0.2510	0.1080	0.1544
89/2	0.6166	0.6198	0.3200	0.5757
89/3	0.8947	0.6850	0.3200	0.5757
89/4	0.8615	0.5827	0.3200	0.5757
90/1	0.6158	0.2789	0.3200	0.5757

USAFE TRANSFORMATION EQUATIONS

Var.	Transformation Equation
MSC	Transformed MSC = (MSC - 56000)(.4/20000) + .1
MAC	Transformed MAC = (MAC - 6000)(.4/2500) + .1
A10	Transformed A10 = (A10 - 9800)(.4/2350) + .1
C130	Transformed C130 = (C130 - 7000)(.4/2000) + .1
C135	Transformed C135 = (C135 - 3700)(.4/700) + .1
F4	Transformed F4 = (F4 - 2600)(.4/5250) + .1
F15	Transformed F15 = (F15 - 5300)(.4/1900) + .1
F16	Transformed F16 = (F16 - 10900)(.4/6150) + .1
F111	Transformed F111 = (F111 - 9400)(.4/1650) + .1
OFF	Transformed OFF = (OFF - 10000)(.4/200) + .1
AMN	Transformed AMN = (AMN - 27900)(.4/375) + .1

Appendix O: PACAF and USAFE MSC Multivariable Network Output

PACAF MSC FULL MULTIVARIABLE NETWORK OUTPUT

FY/Qtr	Actual Tons	Predicted Tons	SE	SY	e_t	$(e_t-e_{(t-1)})^2$	e_t^2
85/1	47567	55460	62293529	10903676	-7893		62293529
85/2	49835	48718	1248178	31025696	1117	81177284	1248178
85/3	59435	53681	33114270	2.30E+08	5755	21504377	33114270
85/4	48235	46999	1528778	15761467	1236	20412889	1528778
86/1	49040	49496	207993	22801307	-456	2864556	207993
86/2	40629	38417	5818950	11805605	2412	8227217	5818950
86/3	42134	42374	57360	4540857	-240	7031778	57360
86/4	34675	39022	18894779	91966730	-4347	16870016	18894779
87/1	42681	42668	181	2508830	13	19011780	181
87/2	39408	40015	368715	23589755	-607	385214.2	368715
87/3	35796	42323	42602953	71722751	-6527	35044920	42602953
87/4	39293	39359	4373	24720074	-66	41744117	4373
88/1	42387	38149	17959055	3526616	4238	18523878	17959055
88/2	48394	49503	1230644	17049231	-1109	28592080	1230644
					Total:	3.01E+08	1.85E+08

FY/Qtr	Actual Tons	Predicted Tons	AE
88/3	55113	52321	2792
88/4	42250	46438	4188
89/1	43079	38324	4755
89/2	45086	45486	400
89/3	44009	43804	205
89/4	41224	39175	2049

MAE:	2398
MIN:	205
MAX:	4755
SSE:	1.85E+08
SSY:	5.62E+08
RSQUARE:	0.67
DW:	1.63
Y BAR:	44264.93

PACAF MSC REDUCED MULTIVARIABLE NETWORK OUTPUT

FY/Qtr	Actual Tons	Predicted Tons	SE	SY	e_t	$(e_t-e_{(t-1)})^2$	e_t^2
85/1	47567	56579	81223466	10903676	-9012		81223466
85/2	49835	49583	63457	31025696	252	85827486	63457
85/3	59435	55296	17130545	2.30E+08	4139	15108769	17130545
85/4	48235	44857	11407928	15761467	3378	579644.3	11407928
86/1	49040	46498	6461764	22801307	2542	698164.7	6461764
86/2	40829	40504	105321	11805605	325	4917168	105321
86/3	42134	42894	578265	4540857	-760	1177157	578265
86/4	34675	39240	20841793	91966730	-4565	14476836	20841793
87/1	42681	42403	77249	2508830	278	23456768	77249
87/2	39408	40523	1244271	23589755	-1115	1941581	1244271
87/3	35796	42877	50137906	71722751	-7081	35585326	50137906
87/4	39293	38754	290689	24720074	539	58063924	290689
88/1	42387	38790	12937510	3526616	3597	9349644	12937510
					Total:	2.76E+08	2.04E+08

FY/Qtr	Actual Tons	Predicted Tons	AE
88/3	55113	52512	2601
88/4	42250	44522	2272
89/1	43079	39791	3288
89/2	45086	54536	9450
89/3	44009	48379	4370
89/4	41224	45988	4764

MAE:	4457
MIN:	2272
MAX:	9450
SSE:	2.04E+08
SSY:	5.62E+08
RSQUARE:	0.64
DW:	1.35
Y BAR:	44264.93

USAFE MSC FULL MULTIVARIABLE NETWORK OUTPUT

FY/Qtr	Actual Tons	Predicted Tons	SE	SY	e_t	$(e_t-e_{(t-1)})^2$	e_t^2
85/1	96200	95829	137641	5.19E+08	371		137641
85/2	71083	69285	3234422	5449890	1798	2037614	3234422
85/3	83702	83487	46311	1.06E+08	215	2506681	46311
85/4	77001	80320	11016425	12841472	-3319	12491276	11016425
86/1	78732	78085	418156	28243910	647	15727173	418156
86/2	80227	77007	10370654	46369290	3220	6623932	10370654
86/3	73784	65964	61153964	134322.3	7820	21157700	61153964
86/4	56587	68991	1.54E+08	2.83E+08	-12404	4.09E+08	1.54E+08
87/1	57977	64534	42997528	2.38E+08	-6557	34181562	42997528
87/2	64120	59904	17770862	86443506	4216	1.16E+08	17770862
87/3	72994	74109	1242445	179352.3	-1115	28411032	1242445
87/4	74368	71528	8063896	903450.3	2840	15636884	8063896
88/1	70603	69884	517249	7921410	719	4496520	517249
88/2	70467	74158	13626803	8705450	-3691	19453833	13626803
					Total:	6.88E+08	3.24E+08

FY/Qtr	Actual Tons	Predicted Tons	AE
88/3	81486	76908	4578
88/4	76688	75211	1477
89/1	73847	76125	2278
89/2	63853	73171	9318
89/3	74806	81517	6711
89/4	88613	83334	5279

MAE:	4940
MIN:	1474
MAX:	9318
SSE:	3.24E+08
SSY:	1.34E+09
RSQUARE:	0.76
DW:	2.12
Y BAR:	73417.5

USAFE MSC REDUCED MULTIVARIABLE NETWORK OUTPUT

FY/Qtr	Actual Tons	Predicted Tons	SE	SY	e_t	$(e_t-e_{(t-1)})^2$	e_t^2
85/1	96200	90872	2.84E+07	5.19E+08	5328.35		28391314
85/2	71083	70667	1.73E+05	5449890	415.9	24132165	172972.8
85/3	83702	81030	7.14E+06	1.06E+08	2672.3	5091341	7141187
85/4	77001	79436	5.93E+06	12841472	-2434.95	26084003	5928982
86/1	78732	72448	3.95E+07	28243910	6283.55	76012242	39483001
86/2	80227	78999	1.51E+06	46369290	1228.35	25555047	1508844
86/3	73784	64652	8.34E+07	134322.3	9132.2	62470845	83397077
86/4	56587	67375	1.16E+08	2.83E+08	-10787.6	3.97E+08	1.16E+08
87/1	57977	63465	3.01E+07	2.38E+08	-5487.75	28088410	30115400
87/2	64120	62198	3.69E+06	86443506	1921.7	54899949	3692931
87/3	72994	70164	8.01E+06	179352.3	2830.2	825372.3	8010032
87/4	74368	67110	5.27E+07	9^3450.3	7257.75	19603199	52674935
88/1	70603	65964	2.15E+07	7921410	4639.05	6857590	21520785
88/2	70467	76468	3.60E+07	8705450	-6000.9	1.13E+08	36010801
						Total: 8.4E+08	4.34E+08

FY/Qtr	Actual Tons	Predicted Tons	AE
88/3	81486	79684	1803
88/4	76688	75821	867
89/1	73847	75344	1497
89/2	63853	67320	3467
89/3	74806	74269	537
89/4	88613	72977	15636

MAE:	3968
MIN:	537
MAX:	15636
SSE:	4.34E+08
SSY	1.34E+09
RSQUARE:	0.68
DW:	1.93
Y BAR:	73417.5

Appendix P: Trend and Seasonal Analysis (PACAF and USAFE MAC Data)

TREND AND SEASONAL ANALYSIS: PACAF MAC

	Actual	DBQ	DBQ-1	DBQ-2
Variance:	189169.5	372812.9	417384.8	989592.1
Index:	100%	197%	221%	523%
Trend:	None	None	Moderate	Strong
Seasonal:	No	Yes	Yes	Yes

++++++++

Year	Qtr	Actual Data	Diffs Between Same Qtr. Ea.Year	First Diffs Between Diffs	Second Diffs Between Diffs
1978	1	5421			
	2	5801			
	3	6021			
	4	5754			
1979	1	5427	6		
	2	5456	-345	-351	
	3	5692	-329	16	367
	4	5940	186	515	499
1980	1	6495	1068	882	367
	2	6647	1191	123	-759
	3	6951	1259	68	-55
	4	6406	466	-793	-861
1981	1	5902	-593	-1059	-266
	2	6016	-631	-38	1021
	3	6166	-785	-154	-116
	4	6818	412	1197	1351
1982	1	6363	461	49	-1148
	2	5860	-156	-617	-666
	3	5934	-232	-76	541
	4	5368	-1450	-1218	-1142
1983	1	5729	-634	816	2034
	2	6768	908	1542	726
	3	6386	452	-456	-1998
	4	6203	835	383	839
1984	1	6020	291	-544	-927
	2	6916	148	-143	401
	3	6050	-336	-484	-341
	4	5829	-374	-38	446
1985	1	5792	-228	146	184
	2	6188	-728	-500	-646
	3	5938	-112	616	1116
	4	6084	255	367	-249

1986	1	5829	37	-218	-585
	2	5569	-619	-656	-438
	3	6782	844	1463	2119
	4	6285	201	-643	-2106
1987	1	5926	97	-104	539
	2	6241	672	575	679
	3	6858	76	-596	-1171
	4	6244	-41	-117	479
1988	1	5739	-187	-146	-29
	2	5531	-710	-523	-377

TREND AND SEASONAL ANALYSIS: USAFE MAC

	Actual	DBQ	DBQ-1	DBQ-2
Variance:	555795.1	318387.1	569278.2	1494079
Index:	100%	168%	301%	790%

Trend:	None	None	Moderate	Strong
Seasonal:	No	Yes	Yes	Yes

++++++++

Year	Qtr	Actual Data	Diffs Between Same Qtr. Ea.Year	First Diffs Between Diffs	Second Diffs Between Diffs
1978	1	9828			
	2	8942			
	3	9793			
	4	9430			
1979	1	8821	-1007		
	2	8831	-111	896	
	3	8841	-952	-841	-1737
	4	9322	-108	844	1685
1980	1	8496	-325	-217	-1061
	2	8573	-258	67	284
	3	8867	26	284	217
	4	9174	-148	-174	-458
1981	1	9244	748	896	1070
	2	8372	-201	-949	-1845
	3	9170	303	504	1453
	4	8747	-427	-730	-1234
1982	1	8610	-634	-207	523
	2	8138	-234	400	607
	3	9270	100	334	-66
	4	9153	406	306	-28
1983	1	8637	27	-379	-685
	2	8242	104	77	456
	3	9178	-92	-196	-273
	4	9327	174	266	462
1984	1	8704	67	-107	-373
	2	9249	1007	940	1047
	3	10058	880	-127	-1067
	4	9672	345	-535	-408
1985	1	9289	585	240	775
	2	9439	190	-395	-635
	3	10259	201	11	406
	4	9377	-295	-496	-507
1986	1	9887	598	893	1389
	2	9383	-56	-654	-1547
	3	10508	249	305	959
	4	10700	1323	1074	769

1987	1	8932	-955	-2278	-3352
	2	9037	-346	609	2887
	3	11433	925	1271	662
	4	11167	467	-458	-1729
1988	1	9956	1024	557	1015
	2	8292	-745	-1769	-2326

Appendix Q: Time Series Analysis (PACAF and USAFE MAC Data)

ARIMA Procedure

Name of variable = PACAF MAC Tonnage.

Mean of working series = 6079.643
Standard deviation = 429.7271
Number of observations = 42

Autocorrelations

```
Lag Covariance Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  0     184665    1.00000  |                    |********************|
  1  67446.713    0.36524  |              .     |*******            |
  2 -13699.566   -0.07419  |             .     *|      .            |
  3 -20158.961   -0.10916  |             .    **|      .            |
  4 -32.950923   -0.00018  |             .      |      .            |
  5 -14004.342   -0.07584  |             .    **|      .            |
  6 -19872.086   -0.10761  |             .    **|      .            |
  7 -15073.557   -0.08163  |             .    **|      .            |
  8 -20288.315   -0.10987  |             .    **|      .            |
  9 -20924.716   -0.11331  |             .    **|      .            |
 10 -32539.597   -0.17621  |             .   ****|      .            |
 11 -24268.040   -0.13142  |             .   ***|      .            |
 12   7772.725    0.04209  |            .       |*      .           |
 13  57953.341    0.31383  |            .       |******  .          |
 14  31688.784    0.17160  |            .       |***    .           |
 15  -2554.411   -0.01383  |            .       |       .           |
 16  -1581.958   -0.00857  |            .       |       .           |
 17  -5107.834   -0.02766  |            .      *|       .           |
 18 -30806.527   -0.16682  |            .    ***|       .           |
 19 -14576.216   -0.07893  |            .     **|       .           |
 20 -21747.100   -0.11776  |            .     **|       .           |
 21 -18905.326   -0.10238  |            .     **|       .           |
 22  -1948.932   -0.01055  |            .       |       .           |
 23  14853.041    0.08043  |            .       |**     .           |
 24  13781.543    0.07463  |            .       |*      .           |
```

." marks two standard errors

Inverse Autocorrelations

```
Lag Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  1   -0.43377  |            *********|     .              |
  2    0.16623  |               .    |***  .              |
  3    0.12504  |               .    |***  .              |
  4   -0.10991  |               .   **|     .              |
  5   -0.00509  |               .    |     .              |
  6    0.15576  |               .    |***  .              |
  7   -0.12543  |               .  ***|     .              |
  8    0.13354  |               .    |***  .              |
  9    0.00064  |               .    |     .              |
 10    0.14782  |               .    |***  .              |
 11   -0.09365  |               .   **|     .              |
 12    0.21459  |               .    |**** .              |
 13   -0.18885  |               . ****|     .              |
 14   -0.00343  |               .    |     .              |
 15    0.12089  |               .    |**   .              |
 16   -0.05776  |               .    *|     .              |
 17   -0.04419  |               .    *|     .              |
 18    0.24284  |               .    |*****.              |
 19   -0.19687  |               . ****|     .              |
 20    0.17587  |               .    |**** .              |
 21   -0.01964  |               .    |     .              |
```

Partial Autocorrelations

```
Lag Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  1    0.36524  |               .    |*******             |
  2   -0.23954  |               .*****|     .              |
  3    0.01458  |               .    |     .              |
  4    0.03188  |               .    |*    .              |
  5   -0.13901  |               .  ***|     .              |
  6   -0.02316  |               .    |     .              |
  7   -0.05206  |               .    *|     .              |
  8   -0.12198  |               .   **|     .              |
  9   -0.05044  |               .    *|     .              |
 10   -0.18778  |               . ****|     .              |
 11   -0.06176  |               .    *|     .              |
 12    0.06619  |               .    |*    .              |
 13    0.24683  |               .    |*****.              |
 14   -0.08525  |               .   **|     .              |
 15   -0.00620  |               .    |     .              |
 16    0.02446  |               .    |     .              |
 17   -0.13463  |               .  ***|     .              |
 18   -0.16076  |               .  ***|     .              |
 19    0.07290  |               .    |*    .              |
 20   -0.26233  |               .*****|     .              |
 21    0.03255  |               .    |*    .              |
 22    0.07362  |               .    |*    .              |
 23    0.07721  |               .    |**   .              |
 24    0.08937  |               .    |**   .              |
```

Autocorrelation Check for White Noise

To Lag	Chi Square	DF	Prob	Autocorrelations					
6	7.71	6	0.260	0.365	-0.074	-0.109	-0.000	-0.076	-0.108
12	12.37	12	0.416	-0.082	-0.110	-0.113	-0.176	-0.131	0.042
18	22.81	18	0.198	0.314	0.172	-0.014	-0.009	-0.028	-0.167
24	26.61	24	0.323	-0.079	-0.118	-0.102	-0.011	0.080	0.075

ARIMA Procedure

Name of variable = USAFE MAC Tonnage.

Mean of working series = 9294
Standard deviation = 736.588
Number of observations = 42

Autocorrelations

```
Lag Covariance Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  0     542562    1.00000  ¦                    ¦********************¦
  1     221749    0.40871  ¦              .     ¦********            ¦
  2  28547.429    0.05262  ¦             .      ¦*     .             ¦
  3     107850    0.19878  ¦             .      ¦****  .             ¦
  4     278311    0.51296  ¦             .      ¦**********          ¦
  5     175977    0.32435  ¦           .        ¦******  .           ¦
  6  22780.476    0.04199  ¦           .        ¦*       .           ¦
  7  76289.381    0.14061  ¦           .       .¦***     .           ¦
  8     147937    0.27266  ¦           .        ¦*****   .           ¦
  9  57770.810    0.10648  ¦           .        ¦**      .           ¦
 10 -29755.071   -0.05484  ¦          .        *¦          .         ¦
 11  -8761.167   -0.01615  ¦          .         ¦          .         ¦
 12  89693.214    0.16531  ¦          .         ¦***       .         ¦
 13 -24931.643   -0.04595  ¦          .        *¦          .         ¦
 14    -105047   -0.19361  ¦          .      ****¦          .         ¦
 15 -78078.024   -0.14391  ¦          .       ***¦          .         ¦
 16  -1535.619   -0.00283  ¦          .         ¦          .         ¦
 17    -125921   -0.23209  ¦          .     *****¦          .         ¦
 18    -170698   -0.31462  ¦          .    ******¦          .         ¦
 19    -105633   -0.19469  ¦          .      ****¦          .         ¦
 20 -18452.095   -0.03401  ¦         .         *¦           .        ¦
 21    -116275   -0.21431  ¦         .      ****¦           .        ¦
 22    -158866   -0.29281  ¦         .    ******¦           .        ¦
 23 -82926.667   -0.15284  ¦         .       ***¦           .        ¦
 24 -45710.881   -0.08425  ¦         .        **¦           .        ¦
                            "." marks two standard errors
```

Inverse Autocorrelations

```
Lag Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  1   -0.38692  |             ********|      .              |
  2    0.14777  |               .     |***   .              |
  3    0.04984  |               .     |*     .              |
  4   -0.17805  |               . ****|      .              |
  5   -0.13524  |               .  ***|      .              |
  6    0.14497  |               .     |***   .              |
  7   -0.17260  |               .  ***|      .              |
  8    0.08650  |               .     |**    .              |
  9    0.05418  |               .     |*     .              |
 10   -0.15073  |               .  ***|      .              |
 11    0.17520  |               .     |**** .               |
 12   -0.18141  |               . ****|      .              |
 13    0.02496  |               .     |      .              |
 14   -0.02037  |               .     |      .              |
 15    0.08750  |               .     |**    .              |
 16   -0.10443  |               .   **|      .              |
 17    0.20913  |               .     |**** .               |
 18   -0.08627  |               .   **|      .              |
 19    0.08367  |               .     |**    .              |
 20   -0.06586  |               .    *|      .              |
 21    0.02162  |               .     |      .              |
```

Partial Autocorrelations

```
Lag Correlation -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  1    0.40871  |               .     |********            |
  2   -0.13737  |               .  ***|      .              |
  3    0.28200  |               .     |******              |
  4    0.41439  |               .     |********            |
  5   -0.02568  |               .    *|      .              |
  6   -0.05271  |               .    *|      .              |
  7    0.11075  |               .     |**    .              |
  8   -0.06273  |               .    *|      .              |
  9   -0.15383  |               .  ***|      .              |
 10   -0.03165  |               .    *|      .              |
 11   -0.10071  |               .   **|      .              |
 12    0.09457  |               .     |**    .              |
 13   -0.19490  |               . ****|      .              |
 14   -0.04429  |               .    *|      .              |
 15   -0.05405  |               .    *|      .              |
 16   -0.04166  |               .    *|      .              |
 17   -0.27164  |               .*****|      .              |
 18    0.04371  |               .     |*     .              |
 19   -0.05067  |               .    *|      .              |
 20    0.07602  |               .     |**    .              |
 21   -0.03425  |               .    *|      .              |
 22    0.07611  |               .     |**    .              |
 23    0.05113  |               .     |*     .              |
 24   -0.10441  |               .   **|      .              |
```

ARIMA Procedure

Autocorrelation Check for White Noise

To Lag	Chi Square	DF	Prob	Autocorrelations					
6	27.67	6	0.000	0.409	0.053	0.199	0.513	0.324	0.042
12	35.26	12	0.000	0.141	0.273	0.106	-0.055	-0.016	0.165
18	50.89	18	0.000	-0.046	-0.194	-0.144	-0.003	-0.232	-0.315
24	69.00	24	0.000	-0.195	-0.034	-0.214	-0.293	-0.153	-0.084

Appendix R: DSXR USAFE MAC Model SAS Regression Output

Dependent Variable: MAC

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	1	10782499.865	10782499.865	6.808	0.0402
Error	6	9503160.0104	1583860.0017		
C Total	7	20285659.875			

Root MSE	1258.51500	R-square	0.5315
Dep Mean	9092.37500	Adj R-sq	0.4535
C.V.	13.84143		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	-9755.237091	7237.3126493	-1.348	0.2263
FH	1	0.244523	0.09371681	2.609	0.0402

Durbin-Watson D	0.986
(For Number of Obs.)	8
1st Order Autocorrelation	0.369

Obs	Dep Var MAC	Predict Value	Std Err Predict	Lower95% Mean	Upper95% Mean	Lower95% Predict	Upper95% Predict
1	8932.0	8659.0	474.940	7496.9	9821.2	5367.6	11950.5
2	9037.0	7695.8	696.035	5992.7	9399.0	4176.8	11214.9
3	11433.0	11645.1	1074.807	9015.2	14275.1	7595.5	15694.8
4	11167.0	10171.6	607.522	8685.1	11658.2	6752.1	13591.1
5	9956.0	8689.8	470.941	7537.5	9842.2	5401.8	11977.9
6	8292.0	8934.4	449.055	7835.6	10033.2	5664.7	12204.0
7	6995.0	8420.1	514.166	7162.0	9678.2	5093.6	11746.7
8	6927.0	8523.1	495.572	7310.4	9735.7	5213.4	11832.7
9	.	6905.8	948.837	4584.1	9227.5	3049.2	10762.4
10	.	8376.4	522.774	7097.2	9655.5	5041.8	11710.9
11	.	11183.2	916.593	8940.4	13426.0	7373.6	14992.9
12	.	10542.6	711.975	8800.4	12284.7	7004.5	14080.7
13	.	6974.5	925.660	4709.5	9239.5	3151.7	10797.3

Obs	Residual	Std Err Residual	Student Residual	-2-1-0 1 2	Cook's D
1	273.0	1165.458	0.234		0.005
2	1341.2	1048.521	1.279	**	0.360
3	-212.1	654.714	-0.324		0.141
4	995.4	1102.169	0.903	*	0.124
5	1266.2	1167.080	1.085	**	0.096
6	-642.4	1175.674	-0.546	*	0.022
7	-1425.1	1148.692	-1.241	**	0.154
8	-1596.1	1156.835	-1.380	**	0.175
9	.	.	.		.
10	.	.	.		.
11	.	.	.		.
12	.	.	.		.
13	.	.	.		.

Sum of Residuals 0
Sum of Squared Residuals 9503160.0104
Predicted Resid SS (Press) 15337245.361

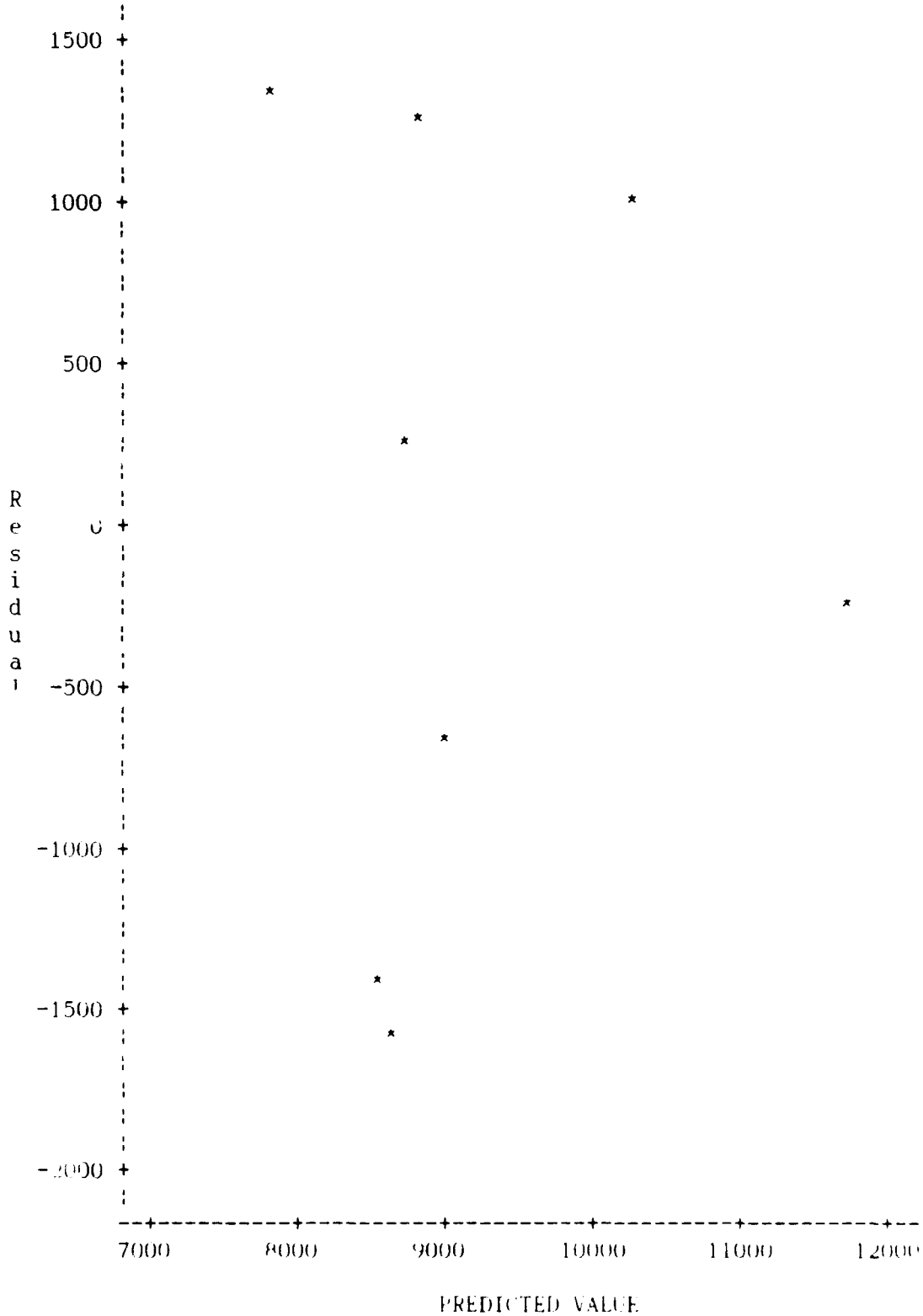

Figure 77. DSXR USAFE MAC Model Residuals versus Predicted Values

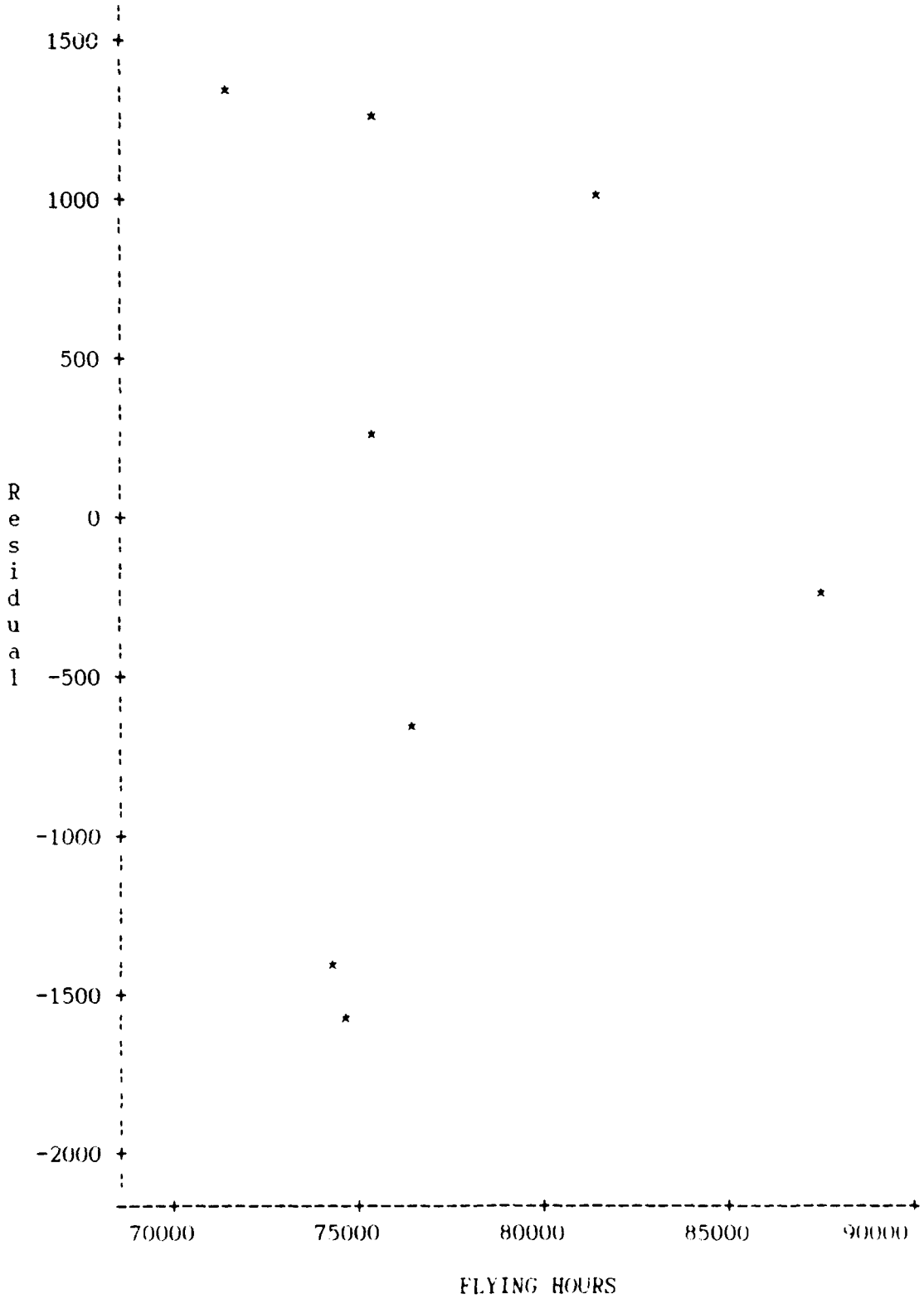

Figure 78. DSXR USAFE MAC Model Residuals versus Flying Hours

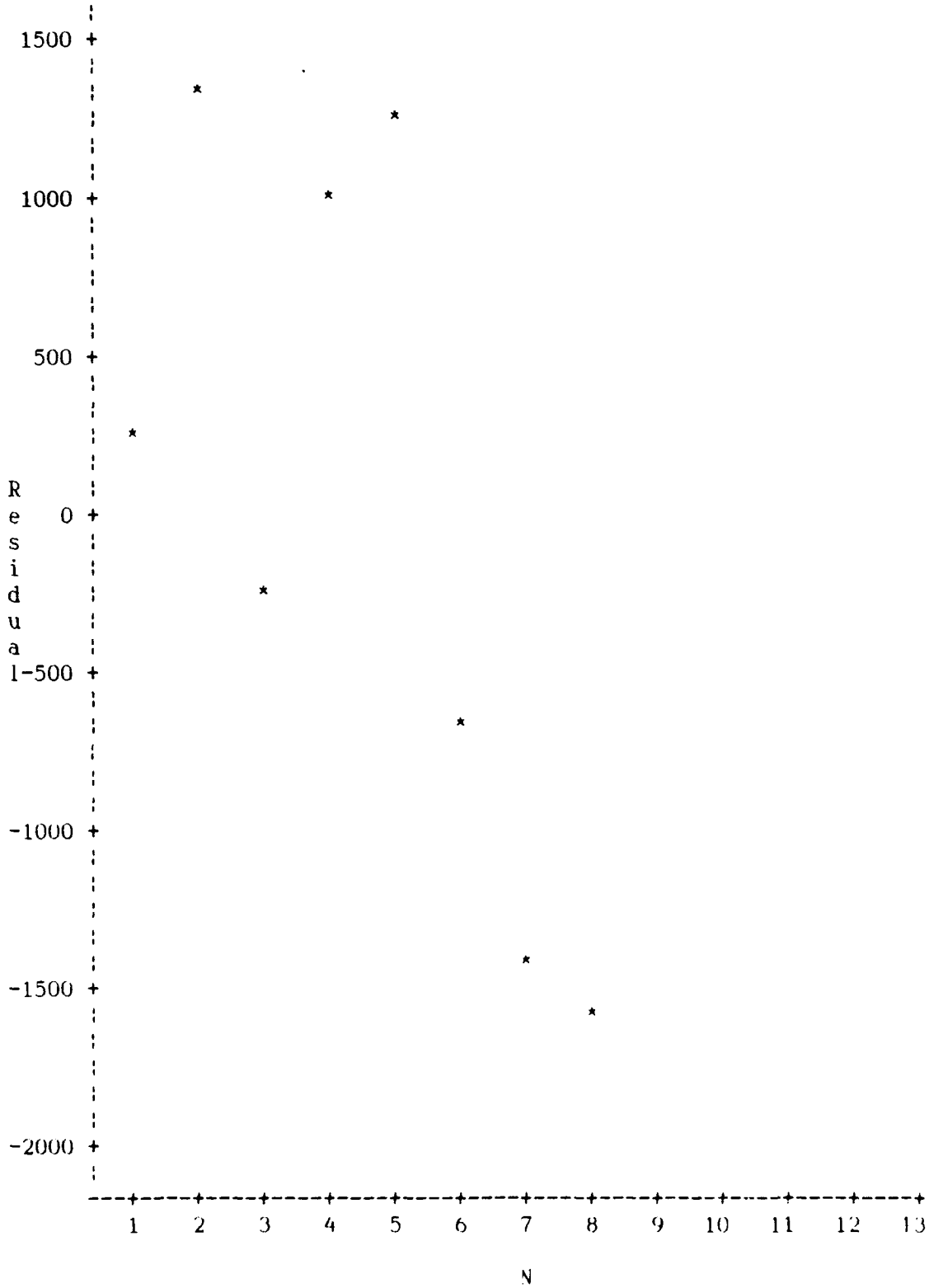

Figure 79. DSXR USAFE MAC Model Residuals versus Time (N)

Univariate Procedure

Variable=RESIDUAL

Residual

Moments

N	8	Sum Wgts	8
Mean	0	Sum	0
Std Dev	1165.158	Variance	1357594
Skewness	-0.21508	Kurtosis	-1.62983
USS	9503160	CSS	9503160
CV	.	Std Mean	411.9457
T:Mean=0	0	Prob>|T|	1.0000
Num ^= 0	8	Num > 0	4
M(Sign)	0	Prob>|M|	1.0000
Sgn Rank	-1	Prob>|S|	0.9453
W:Normal	0.912811	Prob<W	0.3791

Quantiles(Def=5)

100% Max	1341.154	99%	1341.154
75% Q3	1130.767	95%	1341.154
50% Med	30.42555	90%	1341.154
25% Q1	-1033.74	10%	-1596.07
0% Min	-1596.07	5%	-1596.07
		1%	-1596.07
Range	2937.22		
Q3-Q1	2164.504		
Mode	-1596.07		

Extremes

Lowest	Obs	Highest	Obs
-1596.07(	8)	-212.129(	3)
-1425.12(	7)	272.98(	1)
-642.352(	6)	995.3637(	4)
-212.129(	3)	1266.17(	5)
272.98(	1)	1341.154(	2)

Missing Value	.
Count	5
% Count/Nobs	38.46

Appendix S: PACAF MAC Multiple Regression Model SAS Output

Dependent Variable: LMAC

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	4	0.24888	0.06222	21.417	0.0001
Error	9	0.02615	0.00291		
C Total	13	0.27503			

Root MSE	0.05390	R-square	0.9049
Dep Mean	8.66209	Adj R-sq	0.8627
C.V.	0.62225		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	14.700077	3.45308085	4.257	0.0021
AMN	1	-0.000375	0.00022556	-1.664	0.1306
B52	1	0.000298	0.00013712	2.175	0.0576
F15	1	-0.000224	0.00008736	-2.563	0.0305
TP2	1	0.575525	0.09754821	5.900	0.0002

Variable	DF	Variance Inflation
INTERCEP	1	0.00000000
AMN	1	2.64958340
B52	1	2.12662552
F15	1	2.61012673
TP2	1	3.56787427

Collinearity Diagnostics(intercept adjusted)

Number	Eigenvalue	Condition Number	Var Prop AMN	Var Prop B52	Var Prop F15	Var Prop TP2
1	2.42328	1.00000	0.0491	0.0324	0.0282	0.0393
2	1.14843	1.45261	0.0238	0.1782	0.1443	0.0146
3	0.25379	3.09004	0.8755	0.6235	0.0588	0.0396
4	0.17450	3.72648	0.0516	0.1660	0.7687	0.9065

Durbin-Watson D	2.409
(For Number of Obs.)	14
1st Order Autocorrelation	-0.214

Obs	Dep Var LMAC	Predict Value	Std Err Predict	Lower95% Mean	Upper95% Mean	Lower95% Predict	Upper95% Predict
1	8.6891	8.7059	0.023	8.6539	8.7579	8.5733	8.8384
2	8.7134	8.7373	0.025	8.6802	8.7944	8.6026	8.8719
3	8.6706	8.6245	0.038	8.5380	8.7109	8.4750	8.7739
4	8.6250	8.7108	0.025	8.6549	8.7668	8.5767	8.8450
5	8.8220	8.8031	0.037	8.7183	8.8879	8.6546	8.9516
6	8.7459	8.6938	0.026	8.6352	8.7524	8.5585	8.8291
7	8.6871	8.7074	0.031	8.6365	8.7784	8.5664	8.8485
8	8.7389	8.7427	0.027	8.6808	8.8046	8.6059	8.8794
9	8.8332	8.7457	0.028	8.6834	8.8079	8.6088	8.8826
10	8.7394	8.7162	0.018	8.6749	8.7575	8.5875	8.8450
11	8.6550	8.7095	0.034	8.6331	8.7859	8.5656	8.8534
12	8.6181	8.6263	0.044	8.5271	8.7255	8.4691	8.7835
13	8.4613	8.4905	0.030	8.4220	8.5590	8.3506	8.6303
14	8.2703	8.2557	0.049	8.1446	8.3667	8.0908	8.4206
15	.	8.0890	0.080	7.9076	8.2705	7.8704	8.3077
16	.	7.9468	0.142	7.6251	8.2685	7.6028	8.2908
17	.	8.1010	0.090	7.8964	8.3055	7.8628	8.3391
18	.	8.1417	0.081	7.9589	8.3245	7.9220	8.3615
19	.	7.9952	0.116	7.7337	8.2567	7.7067	8.2837

Obs	Residual	Std Err Residual	Student Residual	-2-1-0 1 2	Cook's D
1	-0.0168	0.049	-0.344		0.005
2	-0.0239	0.048	-0.501	*	0.014
3	0.0461	0.038	1.213	**	0.297
4	-0.0858	0.048	-1.793	***	0.171
5	0.0189	0.039	0.489		0.045
6	0.0521	0.047	1.103	**	0.073
7	-0.0203	0.044	-0.464		0.022
8	-0.00379	0.046	-0.082		0.000
9	0.0875	0.046	1.888	***	0.251
10	0.0231	0.051	0.456		0.003
11	-0.0545	0.042	-1.297	**	0.218
12	-0.00818	0.031	-0.261		0.027
13	-0.0292	0.045	-0.655	*	0.040
14	0.0146	0.022	0.655	*	0.416
15	.	.	.		.
16	.	.	.		.
17	.	.	.		.
18	.	.	.		.
19	.	.	.		.

Sum of Residuals	0
Sum of Squared Residuals	0.0261
Predicted Resid SS (Press)	0.0611

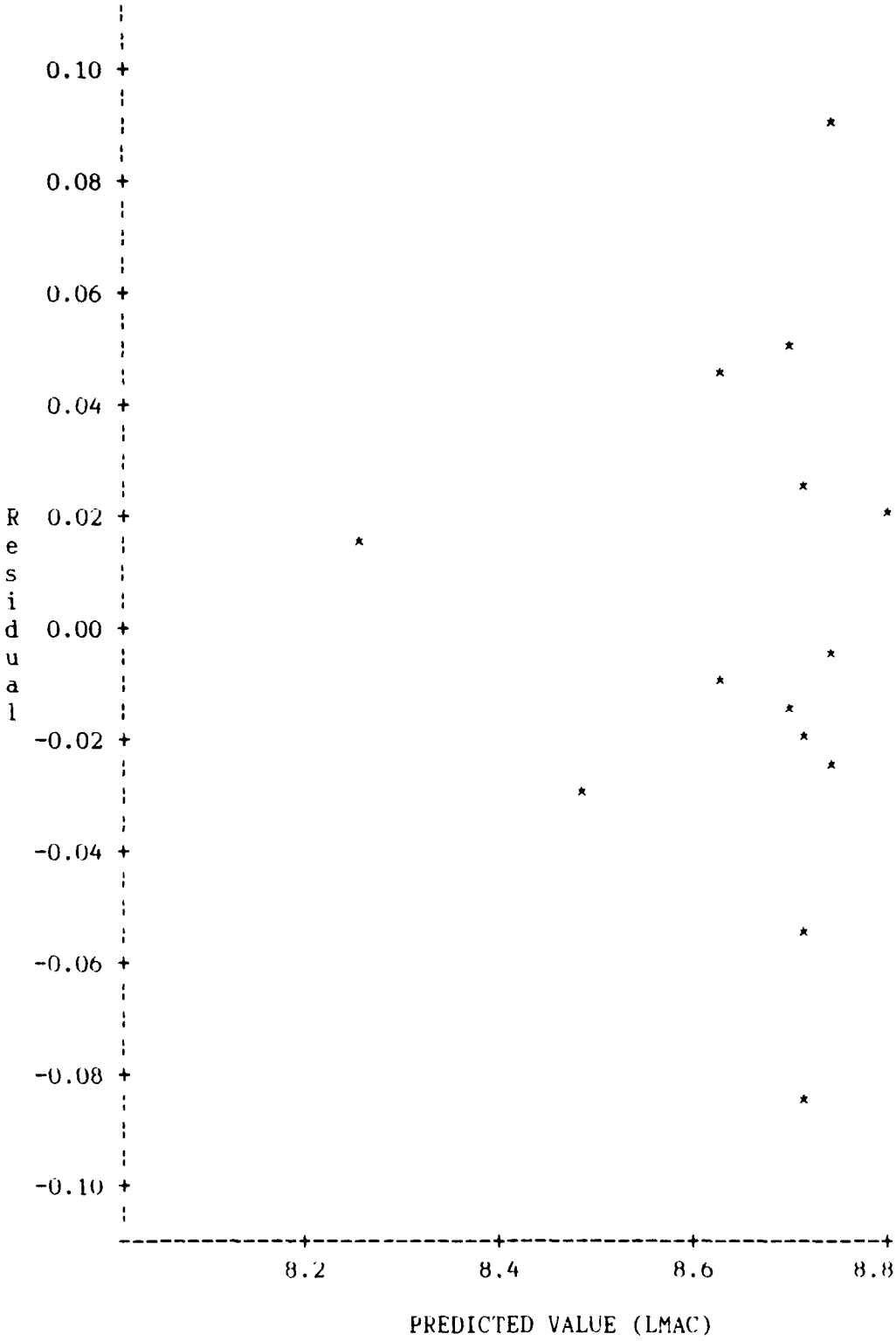

Figure 80. PACAF Multiple Regression MAC Model Residuals versus Predicted Values (LMAC)

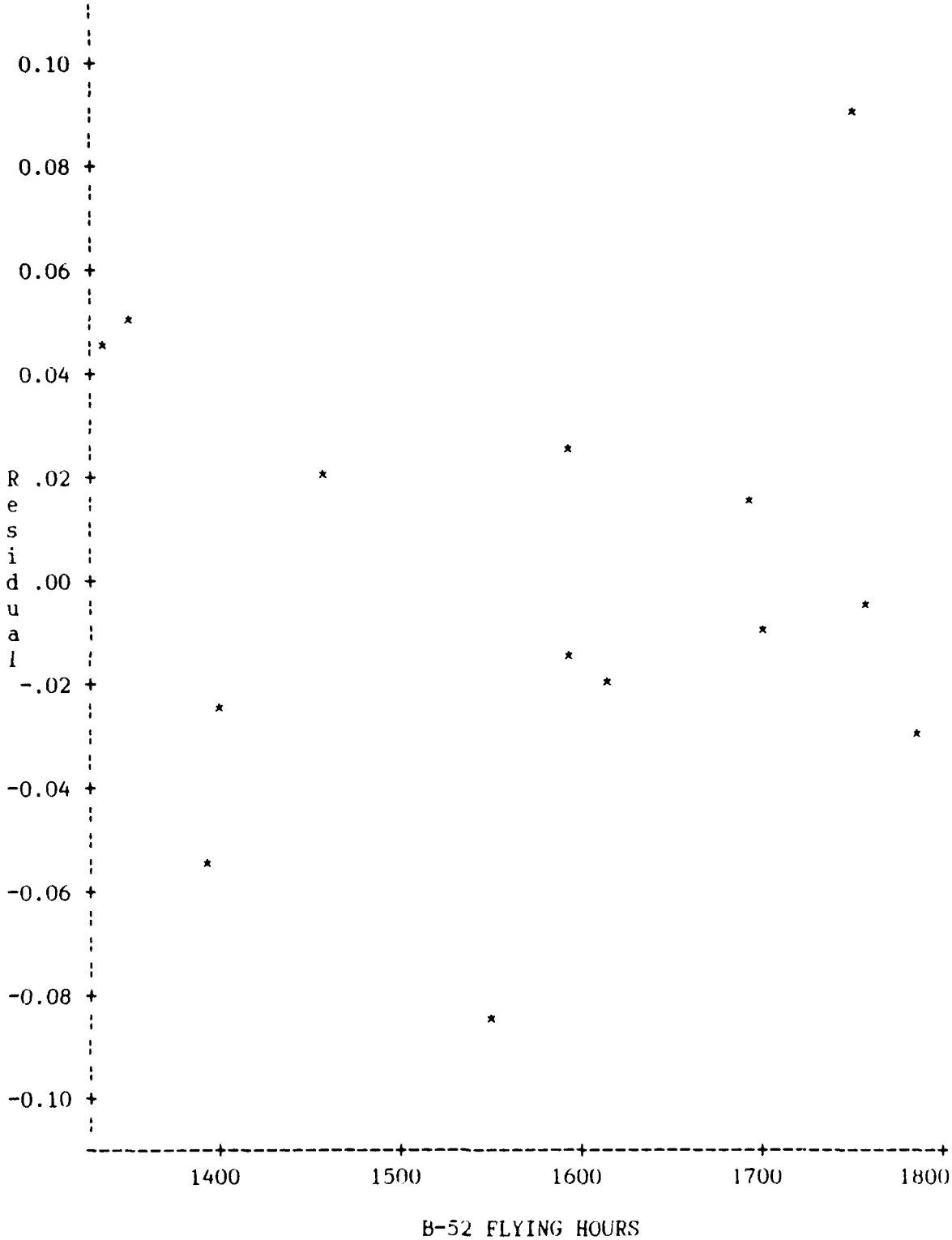

Figure 81. PACAF Multiple Regression MAC Model Residuals versus B-52 Flying Hours

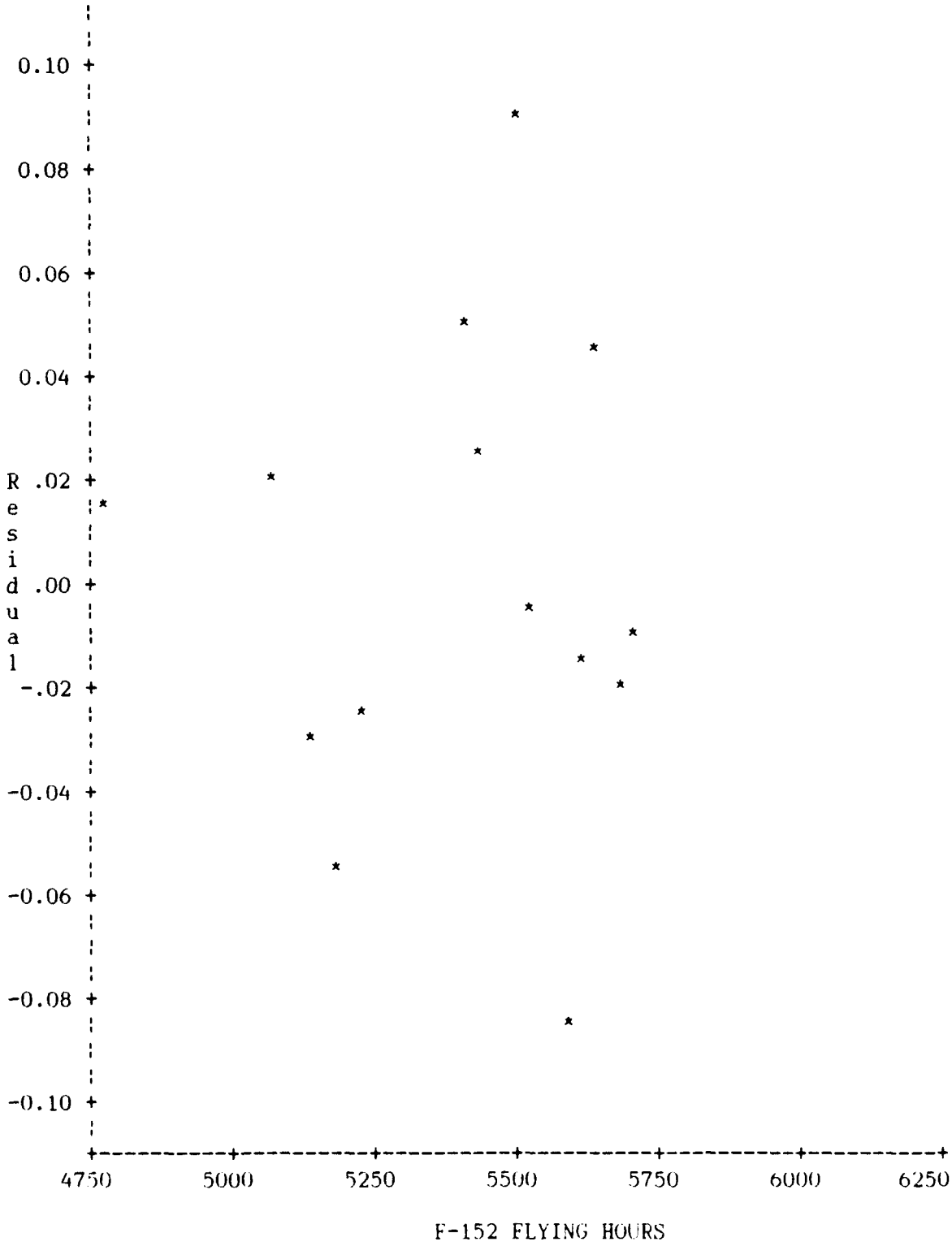

Figure 82. PACAF Multiple Regression MAC Model Residuals versus F-15 Flying Hours

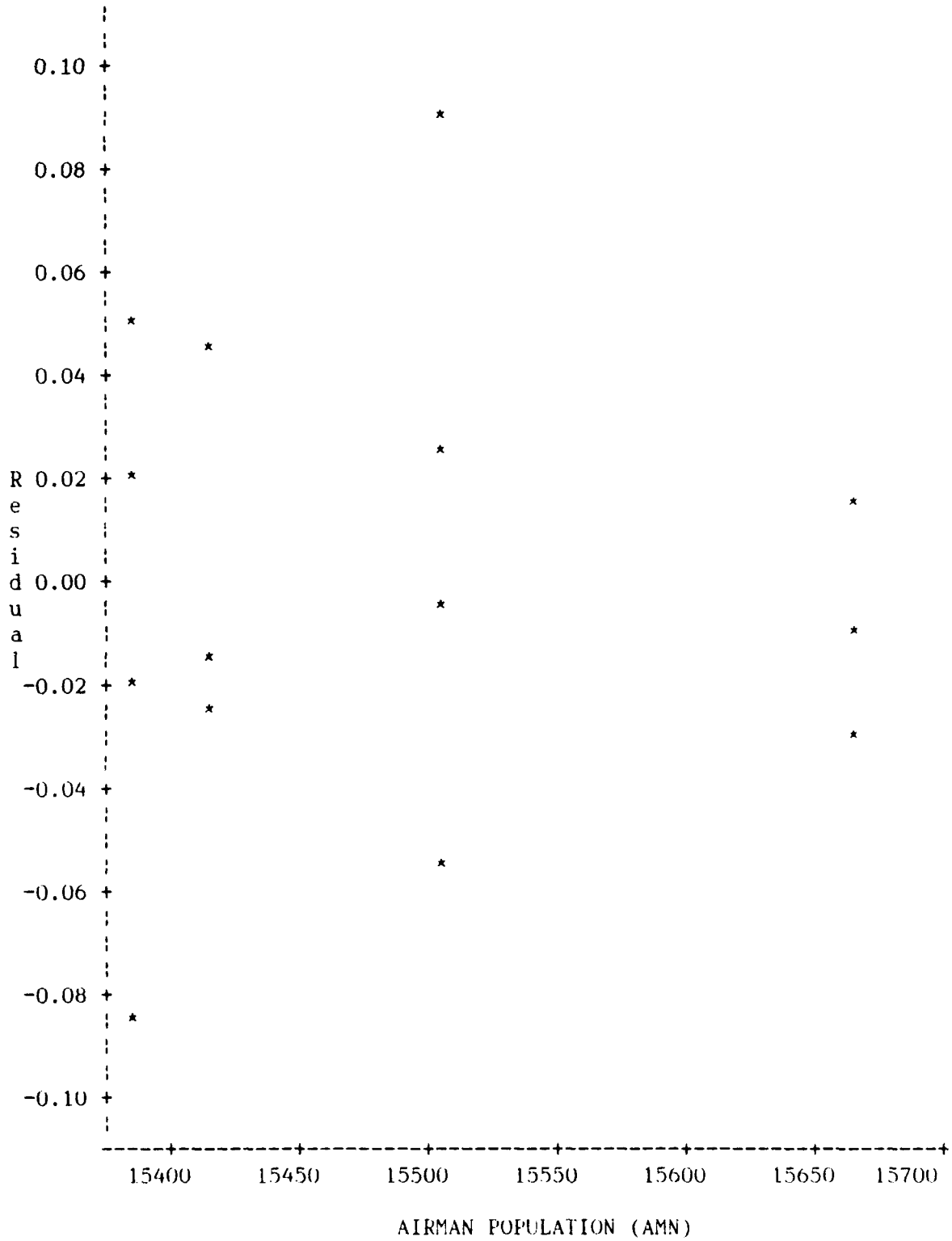

Figure 83. PACAF Multiple Regression MAC Model Residuals versus Airman Population (AMN)

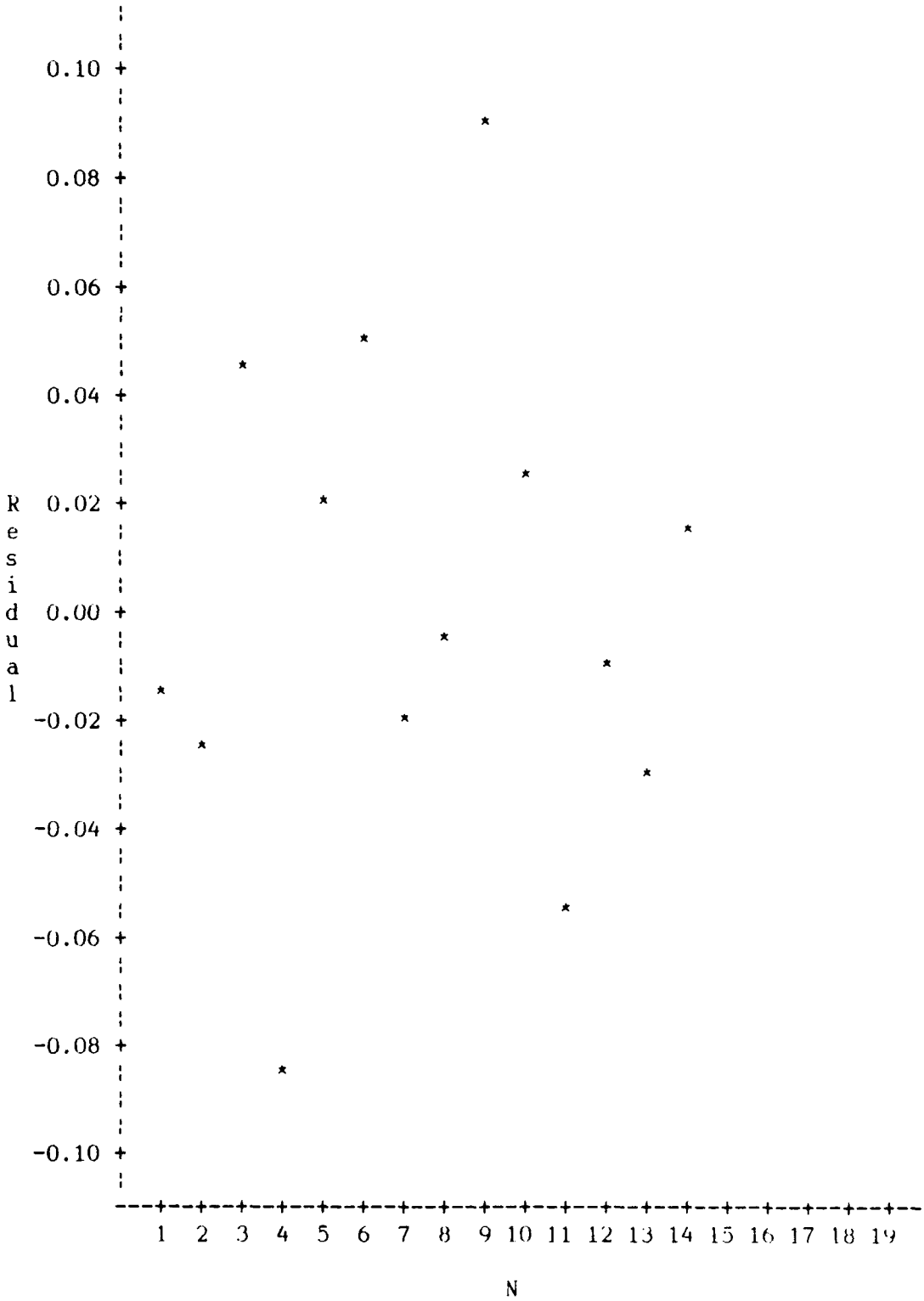

Figure 84. PACAF Multiple Regression MAC Model Residuals versus Time (N)

Univariate Procedure

Variable=RESIDUAL

Residual

Moments

N	14	Sum Wgts	14
Mean	0	Sum	0
Std Dev	0.044847	Variance	0.002011
Skewness	0.085316	Kurtosis	0.256101
USS	0.026146	CSS	0.026146
CV	.	Std Mean	0.011986
T:Mean=0	0	Prob>|T|	1.0000
Num ^= 0	14	Num > 0	6
M(Sign)	-1	Prob>|M|	0.7905
Sgn Rank	-2.5	Prob>|S|	0.9032
W:Normal	0.986504	Prob<W	0.9882

Quantiles(Def=5)

100% Max	0.087503	99%	0.087503
75% Q3	0.023143	95%	0.087503
50% Med	-0.00598	90%	0.052137
25% Q1	-0.02387	10%	-0.05446
0% Min	-0.08585	5%	-0.08585
		1%	-0.08585
Range	0.173351		
Q3-Q1	0.047016		
Mode	-0.08585		

Extremes

Lowest	Obs	Highest	Obs
-0.08585(	4)	0.018934(	5)
-0.05446(	11)	0.023143(	10)
-0.02922(	13)	0.046139(	3)
-0.02387(	2)	0.052137(	6)
-0.02034(	7)	0.087503(	9)

Missing Value	.
Count	5
% Count/Nobs	26.32

Appendix T: PACAF MAC Independent Variable Correlation Matrix

Correlation Analysis

10 'VAR' Variables:	A10	B52	C130	C135	F4
	F15	F16	OFF	AMN	TP2

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
A10	19	2507	281.4353	47625	1759	2856
B52	19	1569	147.7980	29819	1325	1782
C130	19	6920	1203	131475	4370	8778
C135	19	3756	374.9541	71371	2989	4462
F4	19	5598	647.9980	106362	4112	6706
F15	19	5462	308.6415	103771	4774	6228
F16	19	7419	1983	140967	4438	11430
OFF	19	5882	105.1704	111767	5727	5995
AMN	19	15514	105.9798	294775	15385	15664
TP2	19	0.6579	0.4730	12.5000	0	1.0000

Correlation Analysis

Pearson Correlation Coefficients / Prob > |R| under Ho: Rho=0
/ Number of Observations

	A10	B52	C130	C135	F4
A10	1.00000 0.0 19	0.31109 0.1948 19	0.41437 0.0778 19	0.45830 0.0484 19	0.60496 0.0061 19
B52	0.31109 0.1948 19	1.00000 0.0 19	-0.34951 0.1424 19	-0.13225 0.5894 19	0.18190 0.4561 19
C130	0.41437 0.0778 19	-0.34951 0.1424 19	1.00000 0.0 19	0.80641 0.0001 19	0.33316 0.1634 19
C135	0.45830 0.0484 19	-0.13225 0.5894 19	0.80641 0.0001 19	1.00000 0.0 19	0.55106 0.0145 19
F4	0.60496 0.0061 19	0.18190 0.4561 19	0.33316 0.1634 19	0.55106 0.0145 19	1.00000 0.0 19
F15	0.28631 0.2347 19	0.04398 0.8581 19	-0.13391 0.5847 19	-0.25175 0.2985 19	0.12470 0.6110 19
F16	-0.24592 0.3102 19	0.25638 0.2894 19	-0.80188 0.0001 19	-0.83621 0.0001 19	-0.33694 0.1584 19
OFF	0.31038 0.1959 19	-0.30483 0.2045 19	0.73154 0.0004 19	0.60147 0.0064 19	0.46146 0.0467 19
AMN	-0.27944 0.2466 19	0.55184 0.0143 19	-0.75573 0.0002 19	-0.50929 0.0259 19	-0.07749 .7525 19
TP2	0.51349 0.0245 19	-0.25414 0.2938 19	0.72460 0.0004 19	0.75074 0.0002 19	0.62984 0.0039 19

Correlation Analysis

Pearson Correlation Coefficients / Prob > |R| under Ho: Rho=0
/ Number of Observations

	F15	F16	OFF	AMN	TP2
A10	0.28631 0.2347 19	-0.24592 0.3102 19	0.31038 0.1959 19	-0.27944 0.2466 19	0.51349 0.0245 19
B52	0.04398 0.8581 19	0.25638 0.2894 19	-0.30483 0.2045 19	0.55184 0.0143 19	-0.25414 0.2938 19
C130	-0.13391 0.5847 19	-0.80188 0.0001 19	0.73154 0.0004 19	-0.75573 0.0002 19	0.72460 0.0004 19
C135	-0.25175 0.2985 19	-0.83621 0.0001 19	0.60147 0.0064 19	-0.50929 0.0259 19	0.75074 0.0002 19
F4	0.12470 0.6110 19	-0.33694 0.1584 19	0.46146 0.0467 19	-0.07749 0.7525 19	0.62984 0.0039 19
F15	1.00000 0.0 19	0.33108 0.1662 19	-0.21214 0.3833 19	-0.08482 0.7299 19	-0.05286 0.8298 19
F16	0.33108 0.1662 19	1.00000 0.0 19	-0.75728 0.0002 19	0.69479 0.0010 19	-0.73101 0.0004 19
OFF	-0.21214 0.3833 19	-0.75728 0.0002 19	1.00000 0.0 19	-0.76597 0.0001 19	0.87462 0.0001 19
AMN	-0.08482 0.7299 19	0.69479 0.0010 19	-0.76597 0.0001 19	1.00000 0.0 19	-0.70041 0.0008 19
TP2	-0.05286 0.8298 19	-0.73101 0.0004 19	0.87462 0.0001 19	-0.70041 0.0008 19	1.00000 0.0 19

Appendix U: USAFE MAC Multiple Regression Model SAS Output

Dependent Variable: LMAC

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	4	0.29379	0.07345	38.162	0.0001
Error	9	0.01732	0.00192		
C Total	13	0.31111			

Root MSE	0.04387	R-square	0.9443
Dep Mean	9.14729	Adj R-sq	0.9196
C.V.	0.47960		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	12.070129	2.30016919	5.247	0.0005
OFF	1	0.000712	0.00016972	4.193	0.0023
AMN	1	-0.000396	0.00010479	-3.776	0.0044
A10	1	0.000061234	0.00001052	5.820	0.0003
TP2	1	0.227948	0.06024109	3.784	0.0043

Variable	DF	Variance Inflation
INTERCEP	1	0.00000000
OFF	1	3.56382732
AMN	1	2.69698262
A10	1	1.40819433
TP2	1	2.05395829

Collinearity Diagnostics(intercept adjusted)

Number	Eigenvalue	Condition Number	Var Prop OFF	Var Prop AMN	Var Prop A10	Var Prop TP2
1	2.64072	1.00000	0.0336	0.0408	0.0439	0.0424
2	0.77644	1.84419	0.0203	0.0097	0.5708	0.1878
3	0.40358	2.55797	0.0293	0.3800	0.3657	0.4058
4	0.17926	3.83814	0.9168	0.5695	0.0196	0.3640

Durbin-Watson D	2.449
(For Number of Obs.)	14
1st Order Autocorrelation	-0.235

Obs	Dep Var LMAC	Predict Value	Std Err Predict	Lower95% Mean	Upper95% Mean	Lower95% Predict	Upper95% Predict
1	9.2359	9.2546	0.028	9.1912	9.3180	9.1369	9.3724
2	9.1460	9.1711	0.024	9.1165	9.2257	9.0578	9.2844
3	9.1990	9.1517	0.024	9.0966	9.2069	9.0382	9.2653
4	9.1467	9.1055	0.028	9.0425	9.1685	8.9880	9.2231
5	9.2599	9.2933	0.019	9.2512	9.3355	9.1855	9.4011
6	9.2780	9.2254	0.016	9.1888	9.2620	9.1196	9.3312
7	9.0974	9.1666	0.020	9.1209	9.2123	9.0574	9.2759
8	9.1091	9.1402	0.021	9.0921	9.1882	9.0299	9.2504
9	9.3443	9.3368	0.028	9.2727	9.4008	9.2186	9.4549
10	9.3207	9.2720	0.021	9.2249	9.3192	9.1622	9.3819
11	9.2059	9.2233	0.018	9.1828	9.2638	9.1161	9.3305
12	9.0230	9.0277	0.038	8.9415	9.1139	8.8963	9.1591
13	8.8530	8.8481	0.028	8.7852	8.9109	8.7306	8.9655
14	8.8432	8.8456	0.040	8.7541	8.9371	8.7106	8.9806
15	.	8.7637	0.041	8.6716	8.8559	8.6283	8.8992
16	.	8.7496	0.053	8.6297	8.8694	8.5940	8.9052
17	.	8.8565	0.052	8.7389	8.9740	8.7026	9.0103
18	.	8.8808	0.053	8.7617	8.9999	8.7258	9.0358
19	.	8.7001	0.056	8.5745	8.8257	8.5400	8.8602

Obs	Residual	Std Err Residual	Student Residual	-2-1-0 1 2	Cook's D
1	-0.0187	0.034	-0.554	*| |	0.042
2	-0.0251	0.037	-0.685	*| |	0.041
3	0.0472	0.036	1.295	|** |	0.150
4	0.0411	0.034	1.214	|** |	0.199
5	-0.0334	0.040	-0.842	*| |	0.031
6	0.0526	0.041	1.290	|** |	0.052
7	-0.0692	0.039	-1.778	***| |	0.170
8	-0.0311	0.038	-0.809	*| |	0.040
9	0.00748	0.034	0.223	| |	0.007
10	0.0487	0.039	1.261	|** |	0.093
11	-0.0174	0.040	-0.434	| |	0.008
12	-0.00465	0.022	-0.214	| |	0.028
13	0.00487	0.034	0.143	| |	0.003
14	-0.00243	0.017	-0.143	| |	0.023
15	.	.	.		.
16	.	.	.		.
17	.	.	.		.
18	.	.	.		.
19	.	.	.		.

Sum of Residuals	-374E-17
Sum of Squared Residuals	0.0173
Predicted Resid SS (Press)	0.0317

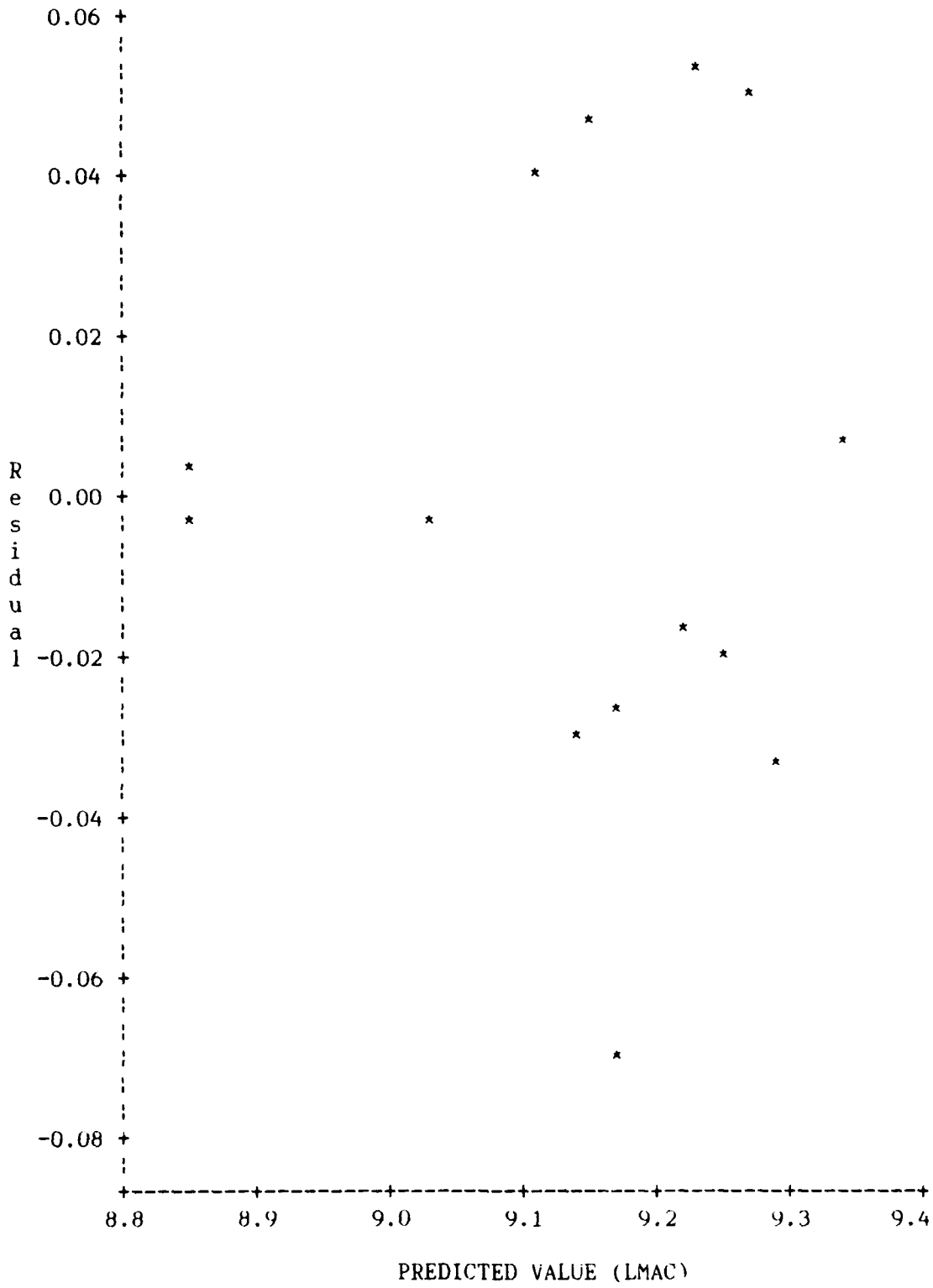

Figure 85. USAFE Multiple Regression MAC Model Residuals versus Predicted Values (LMAC)

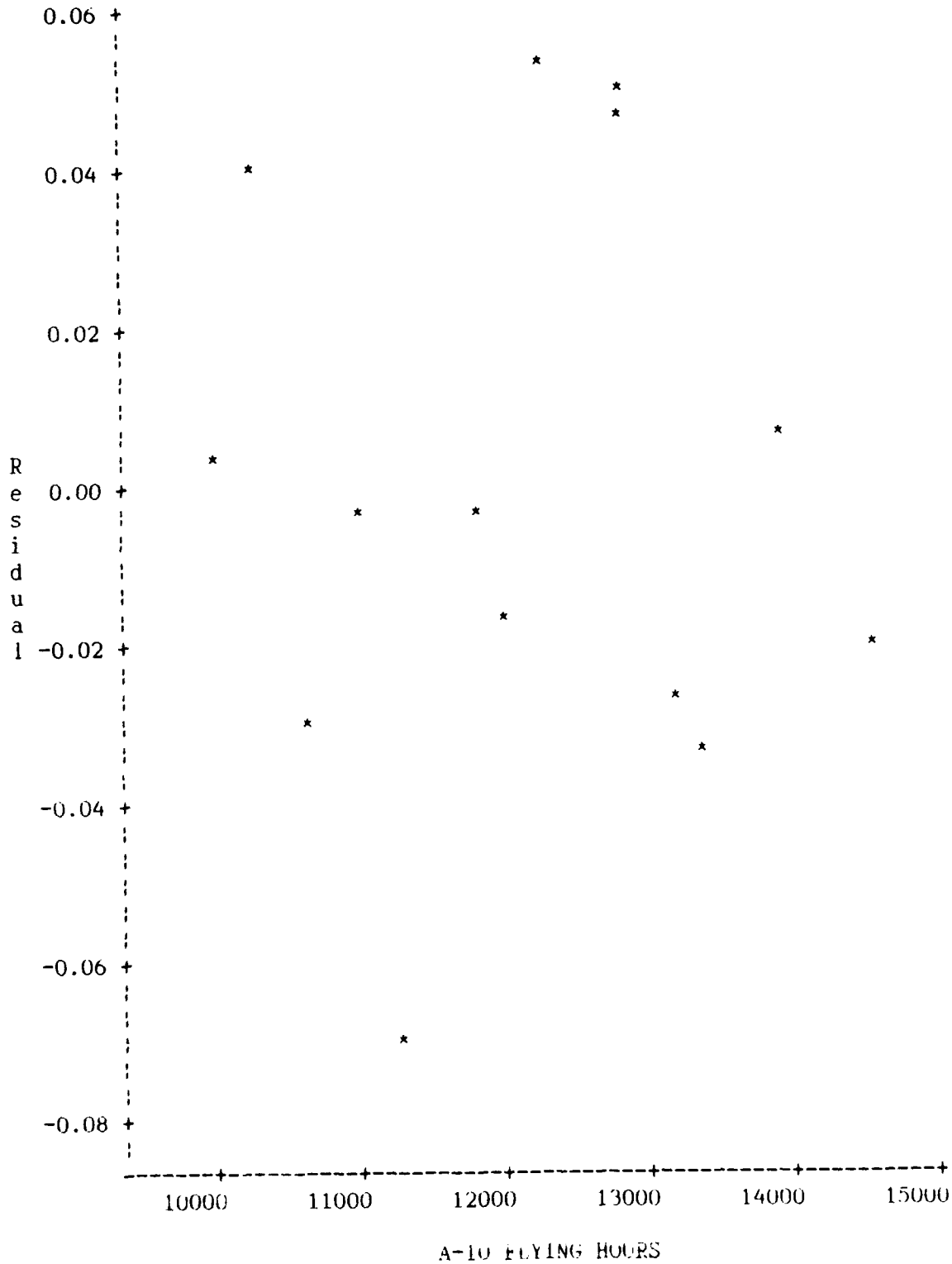

Figure 86. USAFE Multiple Regression MAC Model Residuals versus A-10 Flying Hours

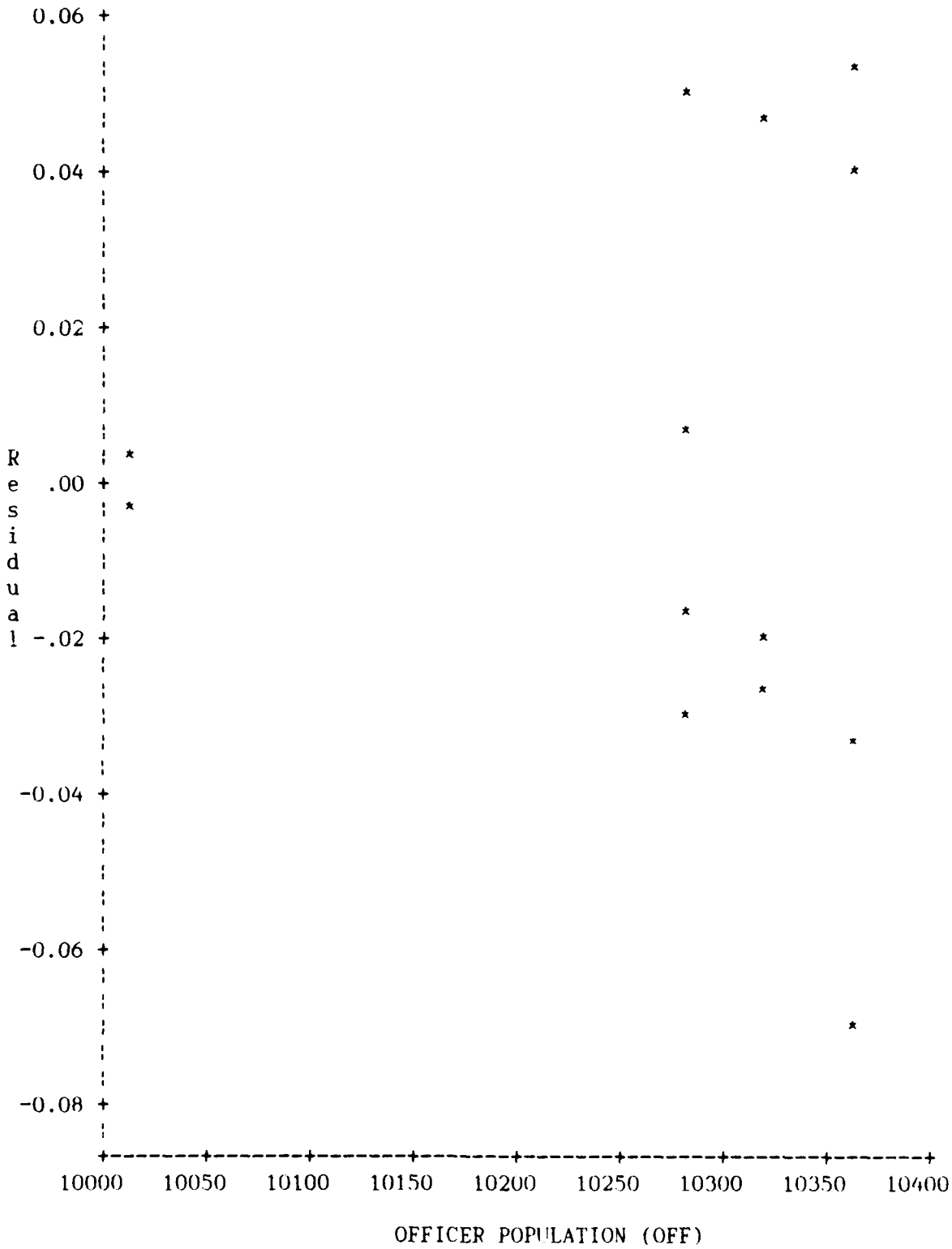

Figure 87. USAFE Multiple Regression MAC Model Residuals versus Officer Population (OFF)

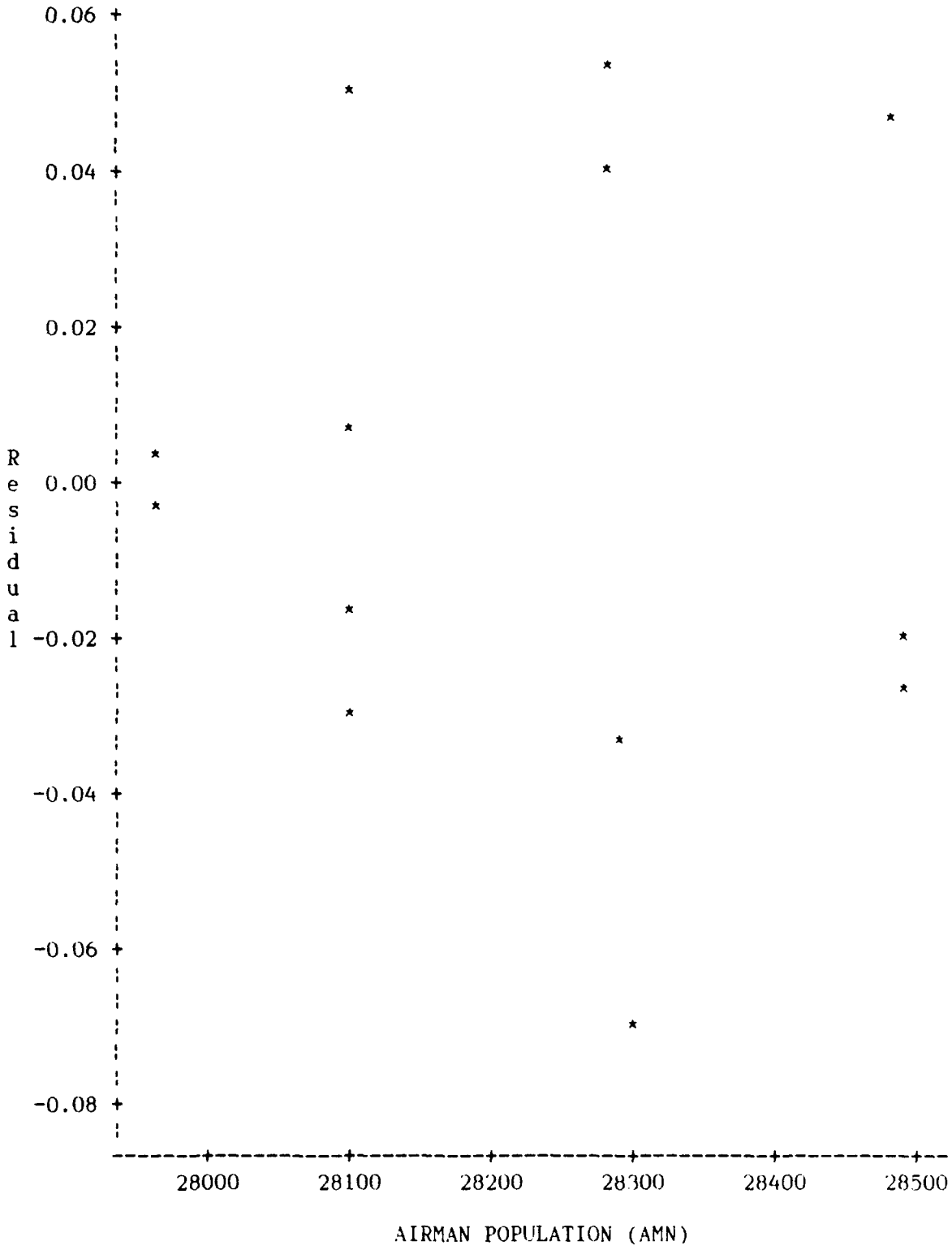

Figure 88. USAFE Multiple Regression MAC Model Residuals versus Airman Population (AMN)

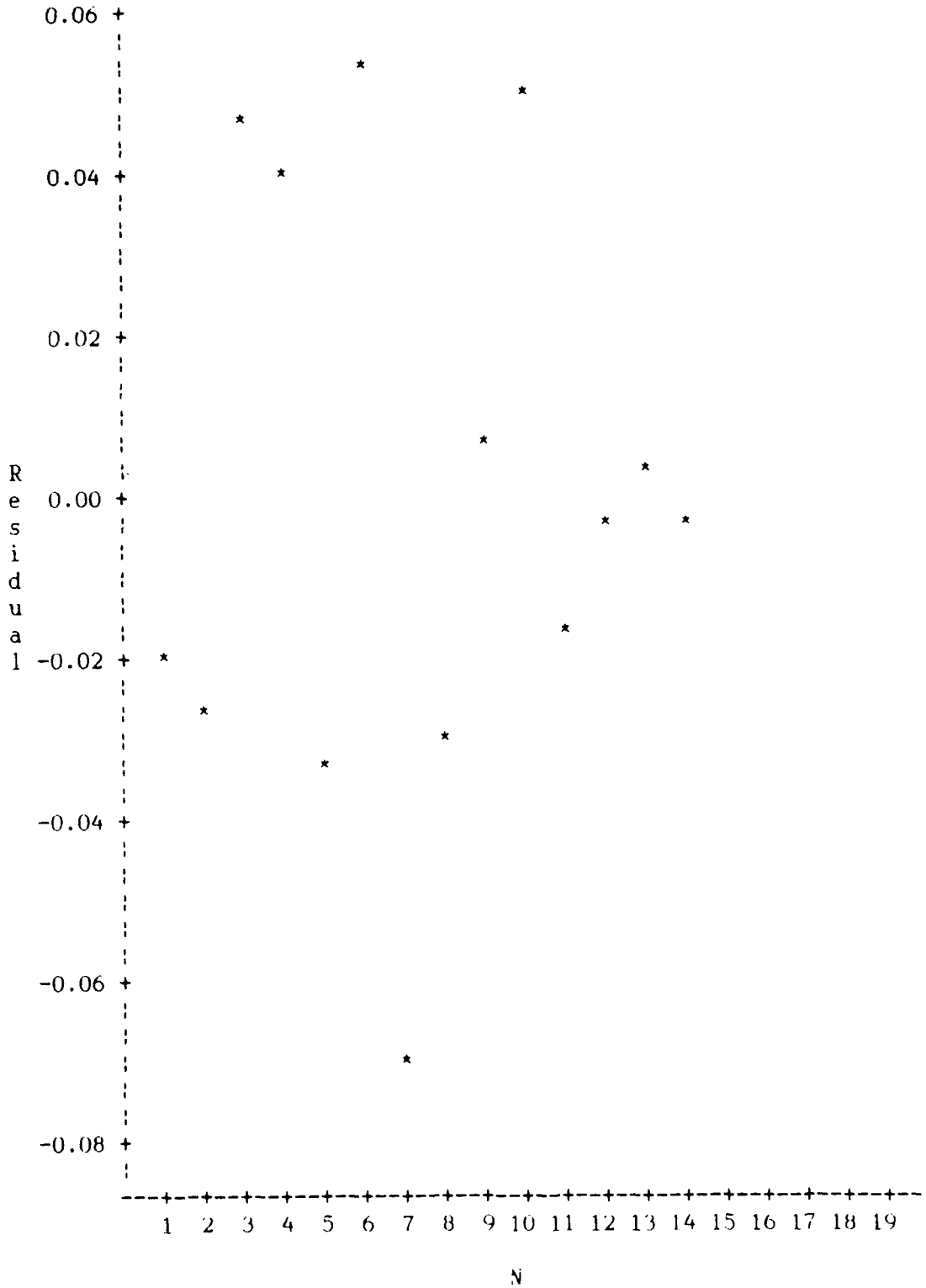

Figure 89. USAFE Multiple Regression MAC Model Residuals versus Time (N)

Univariate Procedure

Variable=RESIDUAL Residual

Moments

N	14	Sum Wgts	14
Mean	0	Sum	0
Std Dev	0.036502	Variance	0.001332
Skewness	0.017305	Kurtosis	-0.61031
USS	0.017321	CSS	0.017321
CV	.	Std Mean	0.009756
T:Mean=0	0	Prob>|T|	1.0000
Num ^= 0	14	Num > 0	6
M(Sign)	-1	Prob>|M|	0.7905
Sgn Rank	0.5	Prob>|S|	1.0000
W:Normal	0.937476	Prob<W	0.3719

Quantiles(Def=5)

100% Max	0.052587	99%	0.052587
75% Q3	0.04114	95%	0.052587
50% Med	-0.00354	90%	0.048669
25% Q1	-0.02508	10%	-0.03343
0% Min	-0.06923	5%	-0.06923
		1%	-0.06923
Range	0.121818		
Q3-Q1	0.06622		
Mode	-0.06923		

Extremes

Lowest	Obs	Highest	Obs
-0.06923(	7)	0.007485(	9)
-0.03343(	5)	0.04114(	4)
-0.03107(	8)	0.047231(	3)
-0.02508(	2)	0.048669(	10)
-0.01871(	1)	0.052587(	6)

Missing Value	.
Count	5
% Count/Nobs	26.32

Appendix V: USAFE MAC Independent Variable Correlation Matrix

Correlation Analysis

10 'VAR' Variables:	A10	C130	C135	F4	F15
	F16	F111	OFF	AMN	TP2

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
A10	19	11992	1360	227854	9874	14426
C130	19	8600	825.2308	163397	7079	10724
C135	19	4445	405.3585	84459	3595	5099
F4	19	6477	3131	123059	2677	13074
F15	19	7257	828.2040	137885	5390	9055
F16	19	16564	3575	314710	11449	23118
F111	19	11051	759.7352	209964	9966	12716
OFF	19	10205	137.1963	193896	10004	10354
AMN	19	28213	185.2100	536042	27951	28470
TP2	19	0.6579	0.4730	12.5000	0	1.0000

Correlation Analysis

Pearson Correlation Coefficients / Prob > |R| under Ho: Rho=0
/ Number of Observations

	A10	C130	C135	F4	F15
A10	1.00000 0.0 19	-0.06584 0.7889 19	0.50482 0.0275 19	0.51813 0.0231 19	0.74610 0.0002 19
C130	-0.06584 0.7889 19	1.00000 0.0 19	0.04257 0.8626 19	-0.01769 0.9427 19	-0.11007 0.6537 19
C135	0.50482 0.0275 19	0.04257 0.8626 19	1.00000 0.0 19	0.52745 0.0203 19	0.47312 0.0408 19
F4	0.51813 0.0231 19	-0.01769 0.9427 19	0.52745 0.0203 19	1.00000 0.0 19	0.52715 0.0204 19
F15	0.74610 0.0002 19	-0.11007 0.6537 19	0.47312 0.0408 19	0.52715 0.0204 19	1.00000 0.0 19
F16	-0.00760 0.9754 19	-0.03300 0.8933 19	-0.34698 0.1455 19	-0.77757 0.0001 19	0.02661 0.9139 19
F111	0.57936 0.0093 19	0.02693 0.9129 19	0.18967 0.4367 19	-0.02839 0.9082 19	0.59208 0.0076 19
OFF	0.38753 0.1011 19	0.04812 0.8449 19	0.28040 0.2449 19	0.75881 0.0002 19	0.27106 0.2616 19
AMN	0.52065 0.0223 19	-0.05458 0.8244 19	0.14959 0.5410 19	0.62422 0.0043 19	0.42345 0.0708 19
TP2	0.18056 0.4595 19	0.12084 0.6222 19	0.21324 0.3807 19	0.62422 0.0043 19	0.18696 0.4434 19

Correlation Analysis

Pearson Correlation Coefficients / Prob > |R| under Ho: Rho=0 / Number of Observations

	F16	F111	OFF	AMN	TP2
A10	-0.00760 0.9754 19	0.57936 0.0093 19	0.38753 0.1011 19	0.52065 0.0223 19	0.18056 0.4595 19
C130	-0.03300 0.8933 19	0.02693 0.9129 19	0.04812 0.8449 19	-0.05458 0.8244 19	0.12084 0.6222 19
C135	-0.34698 0.1455 19	0.18967 0.4367 19	0.28040 0.2449 19	0.14959 0.5410 19	0.21324 0.3807 19
F4	-0.77757 0.0001 19	-0.02839 0.9082 19	0.75881 0.0002 19	0.62422 0.0043 19	0.62422 0.0043 19
F15	0.02661 0.9139 19	0.59208 0.0076 19	0.27106 0.2616 19	0.42345 0.0708 19	0.18696 0.4434 19
F16	1.00000 0.0 19	0.44026 0.0592 19	-0.68912 0.0011 19	-0.35033 0.1414 19	-0.65512 0.0023 19
F111	0.44026 0.0592 19	1.00000 0.0 19	0.05235 0.8315 19	-0.04921 0.8414 19	0.02069 0.9330 19
OFF	-0.68912 0.0011 19	0.05235 0.8315 19	1.00000 0.0 19	0.53393 0.0185 19	0.75588 0.0002 19
AMN	-0.35033 0.1414 19	-0.04921 0.8414 19	0.53393 0.0185 19	1.00000 0.0 19	0.07693 0.7543 19
TP2	-0.65512 0.0023 19	0.02069 0.9330 19	0.75588 0.0002 19	0.07693 0.7543 19	1.00000 0.0 19

Appendix W: PACAF and USAFE MAC Multiple Variable Network Output.

PACAF MAC FULL MULTIVARIABLE NETWORK OUTPUT

FY/Qtr	Actual Tons	Predicted Tons	SE	SY	e_t	$(e_t-e_{(t-1)})^2$	e_t^2
85/3	5938	6174	55535	11055	-236		55535
85/4	6084	5892	36992	63073	192	183178	36992
86/1	5829	5865	1279	15	-36	52025	1279
86/2	5569	6249	462267	69621	-680	414924	462267
86/3	6782	6421	130504	900872	361	1084006	130504
86/4	6285	6136	22175	204433	149	45089	22175
87/1	5926	6383	208886	8676	-457	367178	208886
87/2	6241	6171	4912	166581	70	277858	4912
87/3	6858	6394	215354	1050918	464	155220	215354
87/4	6244	6231	175	169038	13	203261	175
88/1	5739	6123	147465	8809	-384	157791	147465
88/2	5531	5508	525	91118	23	165591	525
88/3	4728	4835	11435	1220709	-107	16862	11435
88/4	3906	3961	3041	3712778	-55	2682	3041
					Total:	3125666	1300544

FY/Qtr	Actual Tons	Predicted Tons	AE
89/1	3841	3890	49
89/2	4124	3905	219
89/3	4056	3947	109
89/4	3841	3829	12
90/1	4305	3698	607

MAE:	199
MIN	12
MAX	607
SSE:	1300544
SSY:	7677696
RSQUARE:	0.83
DW:	2.40
Y BAR:	5832.86

PACAF MAC REDUCED MULTIVARIABLE NETWORK OUTPUT

FY/Qtr	Actual Tons	Predicted Tons	SE	SY	e_t	$(e_t-e_{(t-1)})^2$	e_t^2
85/3	5938	6139	40533	11055	-201		40533
85/4	6084	6088	16	63073	-4	38923	16
86/1	5829	5727	10422	15	102	11263	10422
86/2	5569	6176	368183	69621	-607	502493	368183
86/3	6782	6352	184577	900872	430	1074135	184577
86/4	6285	5993	85510	204433	292	18825	85510
87/1	5926	6214	82979	8676	-288	336960	82979
87/2	6241	6194	2211	166581	47	112279	2211
87/3	6858	6200	432393	1050918	658	372767	432393
87/4	6244	6009	55049	169038	235	178879	55049
88/1	5739	5842	10560	8809	-103	113829	10560
88/2	5531	5506	644	91118	25	16421	644
88/3	4728	4672	3107	1220709	56	921	3107
88/4	3906	3929	528	3712778	-23	6195	528
					Total:	2783890	1276712

FY/Qtr	Actual Tons	Predicted Tons	AE
89/1	3841	3783	58
89/2	4124	3749	375
89/3	4056	3845	211
89/4	3841	3888	47
90/1	4305	3739	566

MAE:	251
MIN	47
MAX	566
SSE:	1276712
SSY:	7677696
RSQUARE:	0.83
DW:	2.18
Y BAR:	5832.86

USAFE MAC FULL MULTIVARIABLE NETWORK OUTPUT

FY/Qtr	Actual Tons	Predicted Tons	SE	SY	e_t	$(e_t-e_{(t-1)})^2$	e_t^2
85/3	10259	10234	624	592130	25		624
85/4	9377	9494	13671	12656	-117	20137	13671
86/1	9887	9680	42686	158006	207	104672	42686
86/2	9383	9339	1908	11342	44	26545	1908
86/3	10508	10521	160	1037342	-13	3173	160
86/4	10700	10399	90846	1465310	301	98631	90846
87/1	8932	9116	33992	310806	-184	235977	33992
87/2	9037	8873	27027	204756	164	121640	27027
87/3	11433	11213	48315	3777192	220	3070	48315
87/4	11167	11139	783	2814006	28	36794	783
88/1	9956	9991	1205	217622	-35	3931	1205
88/2	8292	8217	5598	1434006	75	11997	5598
88/3	6995	6974	460	6222530	21	2850	460
88/4	6927	6862	4163	6566406	65	1856	4163
					Total:	671274	271438

FY/Qtr	Actual Tons	Predicted Tons	AE
89/1	7569	6642	927
89/2	6005	6013	8
89/3	6260	6154	106
89/4	6432	6389	43
90/1	5721	5824	103

MAE:	237
MIN	8
MAX	927
SSE:	271438
SSY:	24824114
RSQUARE:	0.99
DW:	2.47
Y BAR:	9489.5

USAFE MAC REDUCED MULTIVARIABLE NETWORK OUTPUT

FY/Qtr	Actual Tons	Predicted Tons	SE	SY	e_t	$(e_t-e_{(t-1)})^2$	e_t^2
85/3	10259	10623	132792	592130	-364		132792
85/4	9377	9802	180960	12656	-425	3719	180960
86/1	9887	9605	79404	158006	282	500105	79404
86/2	9383	9045	114548	11342	338	3211	114548
86/3	10508	10912	163479	1037342	-404	551715	163479
86/4	10700	10268	186225	1465310	432	698666	186225
87/1	8932	9675	551548	310806	-743	1378746	551548
87/2	9037	9427	152280	204756	-390	124208	152280
87/3	11433	11322	12242	3777192	111	250876	12242
87/4	11167	10748	175462	2814006	419	95010	175462
88/1	9956	10278	103439	217622	-322	548340	103439
88/2	8292	8462	28866	1434006	-170	23019	28866
88/3	6995	6969	661	6222530	26	38264	661
88/4	6927	6983	3098	6566406	-56	6621	3098
					Total:	4222500	1885003

FY/Qtr	Actual Tons	Predicted Tons	AE
89/1	7569	6498	1071
89/2	6005	6462	457
89/3	6260	7096	836
89/4	6432	7275	843
90/1	5721	6247	526

MAE:	746
MIN	457
MAX	1071
SSE:	1885003
SSY:	24824114
RSQUARE:	0.92
DW:	2.24
Y BAR:	9489.50

Appendix X: PACAF and USAFE MAC Time Series Network Data

FY/QTR	PACAF MAC TONS Actual	PACAF MAC TONS Transformed	USAFE MAC TONS Actual	USAFE MAC TONS Transformed
83/1	5729	0.587733	8637	0.52192
/2	6768	0.8648	8242	0.45872
/3	6386	0.76293	9178	0.60848
/4	6203	0.714133	9327	0.63232
84/1	6020	0.665333	8704	0.53264
/2	6916	0.904267	9249	0.61984
/3	6050	0.673333	10058	0.74928
/4	5829	0.6144	9672	0.68752
85/1	5792	0.604533	9289	0.62624
/2	6188	0.710133	9439	0.65024
/3	5938	0.643467	10259	0.78144
/4	6084	0.6824	9377	0.64032
86/1	5829	0.6144	9887	0.72192
/2	5569	0.545067	9383	0.64128
/3	6782	0.868533	10508	0.82128
/4	6285	0.736	10700	0.852
87/1	5926	0.640267	8932	0.56912
/2	6241	0.724267	9037	0.58592
/3	6858	0.8888	11433	0.96928
/4	6244	0.725067	11167	0.92672
88/1	5739	0.5904	9956	0.73296
/2	5531	0.534933	8292	0.46672
/3	4728	0.3208	6995	0.2592
/4	3906	0.1016	6927	0.24832
89/1	3841	0.084267	7569	0.35104
/2	4124	0.159733	6005	0.1008
/3	4056	0.1416	6260	0.1416
/4	3841	0.084267	6432	0.16912
90/1	4305	0.208	5721	0.05536

Variable	Transformation Equations
PACAF MAC	Transformed MAC = (MAC - 3900)(.4/1500) +.1
USAFE MAC	Transformed MAC = (MAC - 6000)(.4/2500) +.1

Appendix Y: PACAF and USAFE MAC Time Series Network Output

PACAF MAC TIME SERIES NETWORK OUTPUT

FY/Qtr	Actual Tons	Predicted Tons	SE	SY	e_t	$(e_t-e_{(t-1)})^2$	e_t^2
84/1	6020	5900	14456	9458	120		14456
84/2	6916	6213	494816	986546	703	340119	494816
84/3	6050	6247	38681	16193	-197	810191	38681
84/4	5829	5885	3089	8789	-56	19908	3089
85/1	5792	6054	68795	17096	-262	42729	68795
85/2	6188	6391	41187	70358	-203	3522	41187
85/3	5938	6277	114940	233	-339	18518	114940
85/4	6084	6095	127	26002	-11	107438	127
86/1	5829	6137	94628	8789	-308	87833	94628
86/2	5569	6183	377245	125139	-614	93995	377245
86/3	6782	6198	340548	738311	584	1434647	340548
86/4	6285	6372	7632	131225	-87	450140	7632
87/1	5926	5916	98	11	10	9456	98
87/2	6241	5893	120769	101283	348	114000	120769
87/3	6858	6324	284660	874693	534	34602	284660
87/4	6244	6180	4121	103202	64	220277	4121
88/1	5739	5828	7854	33764	-89	23355	7854
88/2	5531	6033	252151	153468	-502	171001	252151
88/3	4728	4796	4562	1427428	-68	188878	4562
88/4	3906	3810	9259	4067281	96	26821	9259
					Total:	4197430	2279617

FY/Qtr	Actual Tons	Predicted Tons	AE
89/1	3841	3868	27
89/2	4124	4014	110
89/3	4056	3987	69
89/4	3841	3893	52
90/1	4305	3887	418

MAE:	135
MIN:	27
MAX:	418
SSE:	2279617
SSY:	8899264
RSQUARE:	0.74
DW:	1.84
Y BAR:	5922.75

USAFE MAC TIME SERIES NETWORK OUTPUT

FY/Qtr	Actual Tons	Predicted Tons	SE	SY	e_t	$(e_t-e_{(t-1)})^2$	e_t^2
84/1	8704	10025	1745487	576385	-1321		1745487
84/2	9249	9526	76525	45882	-277	1091059	76525
84/3	10058	10409	123096	353787	-351	5508	123096
84/4	9672	10401	531004	43597	-729	142771	531004
85/1	9289	9522	54490	30346	-233	245291	54490
85/2	9439	9853	171127	586	-414	32488	171127
85/3	10259	10507	61485	633298	-248	27461	61485
85/4	9377	10441	1132841	7430	-1064	666489	1132841
86/1	9887	9620	71249	179606	267	1772293	71249
86/2	9383	10203	672482	6432	-820	1181515	672482
86/3	10508	10308	40093	1091607	200	1040974	40093
86/4	10700	10324	141010	1529674	376	30724	141010
87/1	8932	10044	1235974	282173	-1112	2211931	1235974
87/2	9037	9177	19591	181646	-140	944347	19591
87/3	11433	10712	519427	3880112	721	740772	519427
87/4	11167	11251	7054	2902934	-84	647542	7054
88/1	9956	9148	653399	242852	808	796233	653399
88/2	8292	8796	254262	1371709	-504	1722853	254262
88/3	6995	7711	512754	6092011	-716	44870	512754
88/4	6927	6306	385602	6432310	621	1787669	385602
					Total:	15132788	8408951

FY/Qtr	Actual Tons	Predicted Tons	AE
89/1	7569	6482	1087
89/2	6005	6115	110
89/3	6260	5946	314
89/4	6432	6012	420
90/1	5721	5986	265

MAE:	439
MIN:	110
MAX:	1087
SSE:	8408951
SSY:	25884379
RSQUARE:	0.68
DW:	1.80
Y BAR:	9463.20

Appendix Z: PACAF and USAFE MSC, MAC and Total Flying Hour Data

		PACAF			USAFE		
Year	Qtr	FLYING HOURS	MSC (TONS)	MAC (TONS)	FLYING HOURS	MSC (TONS)	MAC (TONS)
1978	1	35287	31163	5421	57480	46778	9828
	2	37575	30967	5801	52034	38450	8942
	3	34865	32924	6021	64399	47078	9793
	4	35922	32018	5754	69807	37450	9430
1979	1	35310	33469	5427	55423	48574	8821
	2	37593	30548	5456	56900	44156	8831
	3	35600	33046	5692	71221	44041	8841
	4	35615	34991	5940	71846	44296	9322
1980	1	35622	33145	6495	60461	42971	8496
	2	37403	30312	6647	58808	50901	8573
	3	36133	35818	6951	69735	56271	8867
	4	34429	32220	6406	68069	49844	9174
1981	1	36236	35198	5902	63546	53377	9244
	2	35292	30649	6016	65588	53945	8372
	3	36480	35193	6166	75263	60785	9170
	4	36718	35396	6818	75477	54941	8747
1982	1	36194	37343	6363	61878	55855	8610
	2	37007	41379	5860	70025	57543	8138
	3	38635	43392	5934	80973	63076	9270
	4	37534	42968	5368	78365	65851	9153
1983	1	37293	43039	5729	66217	91436	8637
	2	39678	49651	6768	70496	93263	8242
	3	39129	46352	6386	77627	83737	178
	4	37617	35398	6203	77944	83617	9327
1984	1	40018	38462	6020	69485	88315	8704
	2	40533	41800	6916	73093	86968	9249
	3	39523	48352	6050	78651	101701	10058
	4	38235	49203	5829	77702	85521	9672
1985	1	40802	47567	5792	69027	96200	9289
	2	40828	49835	6188	70119	71083	9439
	3	41344	59435	5938	83336	83702	10259
	4	40983	48235	6084	79858	77001	9377
1986	1	42905	49040	5829	75552	75830	9887
	2	41942	40829	5569	74412	79563	9383
	3	41476	42134	6782	80673	71583	10508
	4	41372	34675	6285	77628	55248	10700
1987	1	44270	42681	5926	75308	57088	8932
	2	42322	39408	6241	71368	63014	9037
	3	44381	35796	6858	87519	68675	11433
	4	43701	39293	6244	81329	69487	11167
1988	1	41308	42387	5739	75433	70569	9956
	2	43994	48394	5531	76434	70459	8292
	3	41724	55113	4728	74331	81479	6995
	4	38897	42250	3906	74752	76619	6927

1989	1	40676	43079	3841	68136	73847	7569
	2	41960	45086	4124	74152	63853	6005
	3	42353	44009	4056	85631	74806	6260
	4	36584	41224	3841	83012	88613	6432
1990	1	36672	42355	4305	68417	64655	5721

Bibliography

1. Abell, Joseph A. Evaluation of Bivariate (Linear) Regression for Use as Tonnage Forecasting Model. Unpublished report. Directorate of Material Requirements, DCS Logistics Operations, HQ AFLC/LORET, Wright-Patterson AFB OH, March 1982.

2. Armstrong, Scott J. Long-Range Forecasting From Crystal Ball to Computer (2nd Edition). New York: John Wiley and Sons, 1985.

3. Bails, Dale G. and Larry C. Peppers. Business Fluctuations - Forecasting Techniques and Applications. Englewood Cliffs NJ: Prentice-Hall Inc., 1982.

4. Black, Marlene, G033B System Operations Representative. Personal interviews. Logistics Operations Center, Aerospace Vehicle Distribution Office, HQ AFLC/LOC/XOYS/AVDO, Wright-Patterson AFB OH, 16 November 1989 through 28 November 1989.

5. Curtis, Helena. Biology (Third Edition). New York: Worth Publishers, Inc. 1979.

6. Department of Defense. Glossary of Terms Used in the Areas of Financial, Supply, and Installation Management. DOD Instruction 5000.8. Washington: Government Printing Office, 15 June 1961.

7. Dutta, Soumitra and Shase Shekhar. "Bond Rating: A Non-Conservative Application of Neural Networks," Proceedings of the 1988 IEEE International Conference on Neural Networks, Vol. II. 443-450. New York: IEEE Press, 1988.

8. Freund, Rudolf J. and Ramon C. Littell. SAS System for Regression 1986 Edition. Cary NC: SAS Institute Inc., 1986.

9. Gardner, Everette S. Jr. "How To Detect Trends and Seasonal Cycles," Lotus, 4: 44-50 (May 1988).

10. Gardner, Everette S. Jr. "Anticipating Business Cycles," Lotus, 5: 40-43 (October 1989).

11. Grayson, Major John. An Analysis of Budgeting for Second Destination Transportation. Unpublished report No. 0935-77. Air Command and Staff College, Maxwell AFB AL, May 1977.

12. Hecht-Nielson, Robert. "Theory of the Backpropagation Neural Network," An Invited Paper Presented at the 1988 INNS Annual Meeting. 1-23. San Diego CA: HNC Inc., 1988.

13. Hecht-Nielson, Robert. "Neurocomputing: Picking the Human Brain," IEEE Spectrum: 36-41 (March 1988).

14. Huffman, Brian J. and Thomas R. Hoffman. "Application of Neural Networks in Time Series Forecasting," Proceedings from the Twenty-First Annual Meeting of the Midwest Decision Sciences Institute. 162-164. 2-4 May 1990.

15. Kleinbaum, David G. and Lawrence L. Kupper. Applied Regression Analysis and Other Multivariable Methods. North Scituate MA: Duxbury Press, 1978.

16. Lamb, Captain Christopher J. and Captain Joseph B. Sarnacki. Formulation of a Model to Predict Second Destination Transportation Tonnage Estimates for Future Budget Requirements. MS thesis, LSSR 21-78B. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1978 (AD-A060490).

17. Lippman, Richard P. "An Introduction to Computing with Neural Nets," IEEE Acoustics Speech and Signal Processing, 4: 4-22 (April 1987).

18. Mc Clave, James T. and P. George Benson. Statistics for Business and Economics. San Francisco: Dellen Publishing Co., 1988.

19. Menker, George T., Financial Officer, William R. Lyons, Budget Analyst, Charles D. Ramsey, Budget Analyst, Martin S. Meeks, Budget Analyst. Personal interviews. Budget and Requirements Division, Directorate of Programs and Resources, DCS Distribution, HQ AFLC/DSXR, Wright-Patterson AFB OH, 9 November 1989 to 7 August 1990.

20. Narragon, E. A. and J. M. Neil. Second Destination Transportation: Phase 2. Logistics Management Institute Task 75-4. Washington: LMI, January 1976 (AD-A022602).

21. Neter, John and William Wasserman. Applied Linear Statistical Models: Regression, Analysis of Variance, and Experimental Designs. Homewood IL: Richard D. Irwin, Inc., 1974.

22. Pao, Yon-Han. Adaptive Pattern Recognition and Neural Networks. Reading MA: Addison-Wesley Publishing Co., Inc., 1989.

23. Pindyck, Robert S. and Daniel L. Rubinfield. Econometric Models and Econometric Forecasts (Second Edition). New York: McGraw-Hill Book Co., 1981.

24 Saylor, Robert, Traffic Management Specialist. Telephone interview. Army Materiel Command (AMC), Logistics Control Authority (LCA), Presidio, San Francisco CA, 29 May 1990.

25. Strom, Stephen L. An Analysis of the Air Force Logistics Command (AFLC) Forecasting Method for Predicting Second Destination Transportation (SDT). MS thesis, LSM 89S-59. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1989 (AD-A109881).

26. Tarr, Gregory L. Dynamic Analysis of Feedforward Neural Networks Using Simulated and Measured Data. MS thesis, ENG 88D-54. School of Engineering, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, December 1988 (AD-A202573).

27. Wall, Bill, Budget Analyst. Telephone interview. Navy Material Transportation Office, Service Wide Transportation (SWT), Naval Station Building, Norfolk VA, 1 March 1990 and 29 May 1990.

28. Wasserman, Phillip D. Neural Computing Theory and Practice. New York: Van Nostrand Reinhold, 1989.

29. Wonnacott, Ronald J. and Thomas H. Wonnacott. Econometrics. NY: John Wiley and Sons, 1979.

Vita

Captain Kevin R. Moore graduated from the University of Massachusetts, Amherst, Massachusetts in 1982 and entered the Air Force Officers Training School in May of 1985. After receiving his commission, he completed the basic transportation officer course at Sheppard AFB, Texas, and was subsequently assigned to Pope AFB, North Carolina. He served as an Air Terminal Operations Center duty officer until February, 1988, when he was reassigned to Kunsan Air Base, Korea as a Vehicle Maintenance Officer. Upon graduation from the Air Force Institute of Technology, he will be assigned to Headquarters Military Airlift Command, Scott AFB, Illinois.

[redacted]:

[redacted]

[redacted]

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 1990	Master's Thesis

4. TITLE AND SUBTITLE
FORECASTING AIR FORCE LOGISTICS COMMAND SECOND DESTINATION TRANSPORTATION: AN APPLICATION OF MULTIPLE REGRESSION ANALYSIS AND NEURAL NETWORKS

5. FUNDING NUMBERS

6. AUTHOR(S)
Kevin R. Moore, Captain, USAF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology, WPAFB OH 45433-6583

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFIT/GLM/LSM/90S-37

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The Air Force Logistics Command (AFLC/DSXR) currently uses a simple linear regression model to forecast overseas Second Destination Transportation (SDT) general cargo tonnage requirements for specific geographical areas. The independent variable for the model is the total flying hours for each geographical area while the dependent variable is the general cargo SDT tonnage requirement. This research explored the use of a multivariable approach for developing multiple regression and neural network models which was based on the breakout of the total flying hour variable into separate aircraft flying hours and the addition of military population variables. The purpose of this research was to develop multiple regression and neural network models for predicting Pacific (PACAF) and European (USAFE) Military Airlift Command (MAC) and Military Sealift Command (MSC) general cargo tonnage requirements that were more accurate forecasting models than the simple regression forecasting models presently used by AFLC/DSXR. The multiple regression and neural network models were compared to determine which type of model was statistically more accurate. Overall, the neural network models were the most accurate forecasting models.

14. SUBJECT TERMS
Artificial Intelligence, Neural Nets, Transportation, Logistics, Forecasting, Military Airlift

15. NUMBER OF PAGES
313

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL