PREDICTING F-16 CAUSE CODE H MICAP HOURS USING JMP REGRESSION ANALYSIS

THESIS

Scott E. Carr, Captain, USAF

AFIT-ENS-MS-21-M-146

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENS-MS-21-M-146

PREDICTING F-16 CAUSE CODE H MICAP HOURS USING JMP REGRESSION ANALYSIS

THESIS

Presented to the Faculty

Department of Operational Sciences

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Logistics and Supply Chain Management

Scott E. Carr, BA

Captain, USAF

March 2021

AFIT-ENS-MS-21-M-146

PREDICTING F-16 CAUSE CODE H MICAP HOURS USING JMP REGRESSION ANALYSIS

Scott E. Carr, BA

Captain, USAF

Committee Membership:

Dr. Frank W. Ciarallo
Chair

Dr. Seong-Jong Joo
Reader

AFIT-ENS-MS-21-M-146

Abstract

Emergency demands for aircraft parts, MICAPs (Mission Impaired Capability Awaiting Parts) are one of the leading issues affecting mission capability supply rates for fighter aircraft. Cause code H MICAPS, those with a known demand level, but no supply available, are particularly troublesome and difficult to resolve. These MICAPS are due to failures by the supply chain to replenish stock levels within acceptable time limits. Numerous studies have identified a clear need for proactive measures to reduce MICAP hours. This would significantly improve aircraft availability. This study uses regression analysis implemented in JMP software to build models with a goal of predicting future MICAP hours for F-16 cause code H MICAPs. Four regression techniques, simple linear regression, multiple regression, stepwise multiple regression, and principal component analysis are applied. The data assembled for this analysis includes historical MICAP and demand data for F-16 NSNs (National Stock Number) identified as past top drivers of MICAPs for the F-16. The predictions from these models were systematically compared and filtered based on the error rates. This data assembly, model development, and model comparison effort establishes a methodology for developing accurate predictions of future MICAP hours. There seems to be significant potential for improvement in forecast accuracy with targeted models on subsets of NSNs. These predictive models can be used to direct proactive measures to reduce and prevent MICAPs, thereby increasing availability and readiness across a fleet of fighter aircraft.

Acknowledgments

This thesis is dedicated to the scholars who came before me:

My grandfather, aunt, and father.

Thank you to my mother, family, friends, classmates, and the AFIT faculty for their support in making this research possible.

A special thanks to Dr. Ciarallo, a lighthouse guiding me in my lifeboat as I tried to navigate through the rough waters of grad school during a pandemic.

Scott E. Carr, Capt, USAF

Table of Contents

List of Tables

List of Figures

PREDICTING F-16 CAUSE CODE H MICAP HOURS USING JMP REGRESSION ANALYSIS

I. Introduction

MICAPs (Mission Impaired Capability – Awaiting Parts) are emergency demands for a supply part or parts that affect the mission capability of an Air Force weapon system. These MICAPs can render an aircraft partially mission capable or non-mission capable - where an aircraft cannot be used until the needed part is received to render the aircraft fully mission capable. Thus MICAPS lead to a significant decrease in readiness and aircraft availability. Cause code H MICAPS are MICAPs where there is a historically established demand level for the given part needed, however, none are immediately available in the supply chain. In addition, the demand is not filled within the acceptable time limits set by the Uniform Materiel Movement and Issue Priority System (UMMIPS). This leads to the MICAPs accumulating high levels of MICAP hours, which is the length of time the MICAP is left unresolved. This results in extended downtime for the aircraft. These MICAPs have been a leading issue, especially for Fighter Aircraft, and the systemic resolution is an opportunity that has a great return on investment for the Air Force.

These MICAPs can be the result of a variety of issues including Diminishing Manufacturing Sources and Material Shortages (DMSMS) for the parts, long contract lead times in the production of the parts, and a system of delays in acquiring Awaiting Parts (AWPs) for a subcomponent of a larger end-item in demand. Numerous studies have recommended the need for proactive action to alleviate these issues and thus prevent and reduce future cause code H MICAPs and increase readiness across the fleet.

The purpose of this research is to determine if there is any correlation in historical cause code H MICAPs in the Air Force's fighter fleet that can be leveraged in a model. Such models could be used to enact changes in processes, practices, or resourcing to expedite the resolution of current MICAPs and reduce future ones. This would increase mission capability and aircraft availability.

The specific activities in this research are to create a set of predictive models to forecast future levels of MICAP hours for F-16 cause code H MICAPS. This was initiated by acquiring MICAP and demand data for leading NSNs (National Stock Numbers) with cause code H MICAPS from 2017 to 2020. This data was acquired from the 635 Supply Chain Operations Wing, aggregated monthly, and used in various linear regression techniques using JMP software. The four methods used were, simple linear regression of one historical variable, multiple linear regression of one variable type at multiple points in time, stepwise multiple regression of many variable types at various points in time, and the use of Principle Component Analysis (PCA) of all variables over the given time range to create a new set of variables with reduced correlation for use in multiple regression models. This is done in order to create sets of models at three intervals of prediction to forecast future MICAP hours for each NSN. The best models are identified by the lowest error rates. Each of the four regression techniques were used to develop numerous models in four phases and then the best of each technique at each forecast interval with the lowest errors were evaluated in a final phase to determine the best overall on a separate subset of the data.

Results from this study show relatively high error rates for all models developed compared to the least squared mean for the monthly MICAP Total Hours, with the Stepwise Multiple Regression models performing best across all three intervals of prediction. Application of the models, which were developed on the data set as a whole, showed potential for improvement when applied to various subsets of categorical variables with separate distributions of MICAP total hours and indicate the need for the future development of more accurate models, using specific subsets of NSNs.

The Literature Review in the following chapter reviews the three major sources of cause code H MICAPs, Diminishing Manufacturing Sources and Material Shortages (DMSMS), Awaiting Parts (AWPs) and Spares, and finally long contract lead times. It also covers relevant research on the Air Force's efforts to reduce MICAPs. In addition, the Literature Review summarizes similar studies where various forecasting models are used on multivariate data and compared by their error rates. The Data Collection chapter covers the collection of the data with the 635 Supply Chain Operations Wing (SCOW), how the data is structured, and the variables used along with information on their distribution and correlations. The Methodology chapter introduces the four model types used, simple linear regression, multiple regression, stepwise multiple regression, and principal component analysis, along with the experimental design for this study. The Analysis and Results chapter presents and compares the results of models created in the four phases using the validation set. The final comparison uses the test set along with categorical breakouts of the models. Finally, in the Conclusion chapter, final results are

discussed, including which method performed best. In addition, this chapter recommends some action in the supply chain and areas for further research and model development.

II. Literature Review

Overview

In response to the Secretary of Defense MC 80 Initiative to have fighter aircraft weapon systems at a mission capability rate of 80% by the end of fiscal year 2019, there was a massive effort to increase the readiness. The Perduco Group completed an analysis of the F-16 MICAPs from 2017 to 2019 for the Air Force Materiel Command (AFMC) to determine the effectiveness of their efforts. This analysis concluded that reducing future cause code H MICAPs needed support from the Air Force Sustainment Center (AFSC) working with sources of supply to implement sustainable and long-term corrective actions. It highlighted multiple issues within the supply chain that lead to an increase of cause code H MICAPs, including long lead time contracts, diminishing manufacturing sources and materiel shortages (DMSMS), and bits and pieces also known as awaiting parts (AWPs) (Johnstone & Chandler, 2019).

Diminishing Manufacturing Sources and Materiel Shortages and Obsolescence

Maddux (1999) describes the impact of Diminishing Manufacturing Sources and Materiel Shortages (DMSMS) on spare parts for the US Army Aviation & Missile Command. DMSMS is caused by the obsolescence and unavailability of spare parts and components for weapon systems. This has a negative impact on life cycle supportability of weapon systems. Assessment methods to mitigate DMSMS components need to be

continuously refined. In addition, alternative sources, parts, and substitutes need to be identified for unavailable or obsolete components. Lastly, opportunities for the introduction of new technologies into weapon systems must be.

Livingston (2000) describes DMSMS management strategies. It develops a breakdown of weapon systems' life cycles in six stages from introduction to phase out. DMSMS for components increases as systems age. Thus, supply chains must transition from reactive short-term actions to implementing proactive mitigation strategies driven by technology. Important aspects of this proactive strategy include improving the software on commercially open platforms, identifying critical technological components and finding alternatives for them, and lifetime buys of assets and bridge buys to compensate for DMSMS.

Kobren, Melnikow, and Robinson (2005) addresses DMSMS strategies arguing that the development of such a strategy is the responsibility of the program manager (PM). Furthermore, the strategy must be designed in close coordination with the original equipment manufacturer (OEM) of a component. Proactive and early planning for DMSMS is key in the production design with open systems. When DMSMS mitigation reactively, production extensions, redefined requirements, redesigning, and information sharing are all useful tools.

Sandborn, Prabhakar, and Ahmad (2011) studies microelectronics DMSMS in commercial markets. The recommendations focus on the necessity for forecast accuracy, and once again highlight the need to be more proactive than reactive. An important driver of success in managing DMSMS in this commercial context is communication

with the manufacturer on expected production, leading to increased forecast accuracy. In addition, past trends are valid predictors of obsolescence in the future. In particular, age is a driving factor leading to problems with DMSMS, accelerated by rapidly changing technology.

Miller (2012) conducts a multiple case study into DMSMS of electronic components for the US Air Force. The research includes issues from the F-22, B-2, E-3, and RQ-4B weapon systems. It establishes best practices necessary to handle the growing DMSMS problem. The case studies reinforce that both proactive and reactive mitigation strategies are necessary in practice. With that as a backdrop, a greater return on investment was observed from the use of proactive strategies. The key to proactive strategies is the evaluation and use of predictive tools that assess the health of the supply chain continuously.

Zolghadri, Addouche, Boisse, and Richard (2018) describes a case study from the automobile industry. The case involves the implementation of a Bayesian model for predicting parts obsolescence. This approach establishes three categories for mitigation strategies: (1) reactive strategies responding to current or upcoming obsolescence; (2) proactive strategies predicting the health of critical assets; and (3) strategic management which mitigates obsolescence risks during the design phase. The results emphasize that items with long life cycles need strategies focused on sustainment. Alternatively, short life cycle items benefit from mass production of assets. The implementation used software to model assets, representing cost and delay as networks of nodes.

Awaiting Parts and Spares

Awaiting Parts (AWPs) and Spares are an issue when the production, repair, or replacement of a larger end item is delayed by the procurement or repair of a subcomponent to that end item. These delays can have a compounding effect in the time it takes to return a weapon system back to its full mission capability, best represented by the build-up of MICAP hours for an NSN.

Huber (1985) investigates the sensitivity of Awaiting Parts (AWPs') at a base to variables describing the state of the depot repair cycle. Based on this investigation, Huber determined that little could be achieved from base-level initiatives. The analysis was based on using the Dyna-Metric supply system to more effectively manage reparable item inventories. The examples use data from ten Line Replaceable Unit (LRU) items in AWP status from the KC-135A fleet with long delays. Reparable in-transit days and depot shop flow were two variables that had the greatest impact on AWPs.

Levy and Jondrow (1990) discusses the effects of Readiness Based Sparing (RBS) processes on weapon system availability. In the past, supply requirements for weapon systems onboard a ship were determined by historical replacement rates. Levy and Jondrow studied two weapon systems, the AN/SPG-55B fire control radar, and the Mk 15 Close-in weapon system. Data from the Nonexpendable Shipboard Equipment Status Long (NAVSEA 4855 data) supported a comparison of the following before and after the RBS implementation: weapons system availability, AWP events, and time awaiting parts. This analysis determined that RBS increased weapon system availability by 3-4% and 3-5% respectively for these systems.

Eitan (1990) develops a model for Awaiting Parts Time for Naval aircraft spares. The goal of the model was to create more accurate Aviation Consolidated Allowance Lists (AVCAL) used for Readiness Based Sparing (RBS) on a ship. The model was based on the failure rates of Weapon Reparable Assemblies (WRA) and their Shop Replaceable Assemblies (SRA) subcomponents. One SRA failure was expected to accumulate Awaiting Parts Time very quickly. The final model from this effort was too computationally intensive to be used in a full-scale application.

Muno and Pezoules (1997) addresses the resupply of Line Replaceable Units (LRUs) for F-16 MICAPs. LRUs are items that are removed from a larger end item, replaced, and then sent to the depot for repair. The analysis used the theory of constraints to assess five stock numbers in the enterprise over three months of shipping records. The analysis showed that the most significant delays were caused by improper planning of asset shipments which exacerbated the bottlenecks in the transportation process.

Manship & Moore (2000) addresses the excessive levels of assets in Awaiting Parts status, with a particular focus on LRUs and SRUs. Analysis of AWPs from the EXPRESS system determined that many assets are in AWP status for extended periods at the base level. Manship and Moore proposed new business rules to redistribute AWP SRUs from bases with significant backlogs to other bases with lower AWP levels so that the assets can be repaired more quickly and returned for use.

Slay, Burleson, and Meyenburg (2000) studies the issue of spare part requirements for strategic airlift weapons systems. The findings focus on the benefits of improved computation of Mobility Readiness Spares Packages (MRSP) using the Aircraft

Sustainability Model (ASM). This model computes requirements in the tactical environment. Inaccuracies occur between the requirements of the In-place Readiness Spares Packages (IRSP) at the Main Operating Base (MOB) of the weapon system, and the MRSP for the aircraft in the deployed environment.

Carter and London (2000) explores excessive AWP Line Replaceable Units (LRU) and argues that insufficient information on the needed repair parts was entered into the EXPRESS system. This led to an enormous number of backorders in the system, with a negative effect on MICAPs. The recommended solution called for the LRU requirements to be used in a model to determine the repair part stockage levels. The model highlighted the potential to use probability theory to calculate the correct levels.

Long Contract Lead Times

Chen, Drezner, Ryan, and Simchi-Levi (2000) studies the bullwhip effect. The effect is measured by the amplification of demand variability as one moves further up a supply chain. This paper uses a model of a two-stage supply chain with one manufacturer and one retailer to demonstrate the effect. The model is then used to show that by centralizing demand information (making it available at every stage of the supply chain), the bullwhip effect can be significantly decreased.

Song, Yano, and Larssrisuriya (2000) studies the effects of demand and lead time uncertainty. The obsolescence of electronic components can lead to final order demand variability. This paper utilized a model of a make-to-order situation for a single customer with demand uncertainty, multiple components in production with random lead times,

and a fixed due date. The results showed that lead time uncertainty had a much more significant impact on performance than demand variability.

Dolgui and Ould-Louly (2002) surveys existing optimization models that include planned lead times for reducing expensive backlogs and holding costs. The survey included both single and multi-item models, with a single-level of production. Demand is constant and lead time is random in each of the models. The paper then proposes an original approach using an auxiliary Markov chain based on the Materiel Requirements Planning (MRP) Method. This leads to optimal control policies in some of the cases studied. However, more work was required to develop effective approaches for general use.

De Treville, Shapiro, and Hameri (2004) focuses on the roles of the transfer of demand information and supply lead time reduction in improving demand chain performance. The paper develops a typology of various supply chains according to lead times. The proposed framework allows for prioritized efforts in the reduction of supply lead times. Lead time reduction was identified as far more efficient than the sharing of demand information across the chain in improving performance.

Leng and Parler (2009) introduced a game theoretic model of a two-level supply chain with a supplier and a retailer, with a goal of lead-time reduction. The model partitions production lead time into three parts: set up time, production time, and shipping time. The results indicate that the implementation of profit-sharing contracts if designed properly, could be an effective way to reduce lead time and increase coordination within the supply chain.

Peltz et al. (2015) studies supply chain agility for the Defense Logistics Agency (DLA). This effort found that reduction of Production Lead Time (PLT) was not an area of focus in the DLA's acquisitions processes. With items that are unique to military applications that have a limited supply base, buyers do not consider PLT reduction in the bid selection process regularly. They generally accept lead times that are consistent with historical records.

Other Relevant Research

There are a number of reports on related topics, as well as related AFIT theses, that are relevant to forecasting MICAP cause code H hours. Greer and Moon (1981) studies base-level supply management indicators that were used at the time. This survey-based approach collected information about a comprehensive list of sixteen indicators across four major commands. The survey covered three major fields: stockage support indicators, not mission capable supply indicators, and warehouse shortage indicators. The overall conclusions emphasized that the majority of management indicators at the time were being vastly underutilized by the population of mostly young and inexperienced supply officers. These officers did not utilize or comprehend the indicators and were unprepared to take appropriate corrective actions when negative trends were shown by the indicators. Finally, the authors recommended the development of a comprehensive manual to document and aid decision makers.

Blazer (1984) studies the effectiveness of the Air Force's demand forecasting efforts. This Air Force Logistics Management Center report compares various forecasting methods that were in use at the time. The comparison is based on demand

averages, demand variation, and their impact on safety stocks. Blazer's study found that the demand pattern forecasts were satisfactory, however, the safety stock quantities from an economic order quantity model were inadequate. Additionally, the Department of Defense's policy constrained the allowable safety level and limited the Air Force from achieving its objective. The report concludes that the USAF could decrease MICAP occurrences by an average of four percent if computations of supply safety levels included accurate measures of demand variation and order shipment times.

Ingram (2020) studies the impact that Weapon System Sustainment (WSS) activities have on Aircraft Availability. A significant contribution of this AFIT thesis is establishing the lead times needed for leadership and decision-makers to positively impact Aircraft Availability. These lead times were established through a study of time lags using an Ordinary Least Squares Regression model. This research focused on four aspects of WSS: Technical Orders, Sustainment Engineering, Contractor Logistics Support, and Depot Purchased Equipment Maintenance. The model estimated the percentage impact each of these WSS aspects had on Aircraft Availability holistically. The study computed the lead time in months needed by these business processes to benefit Aircraft Availability. In addition, the analysis estimated the financial obligation to each process required to create a one percent improvement in Aircraft Availability.

Forecasting Methods

There are numerous examples of academic studies that apply a range of forecasting techniques to multivariate data sets for a wide range of practical applications. The studies that are the primary focus here are centered on creating multiple prediction

models. The studies then reduce the number of models, or rank their performance, through a comparison using an appropriate measure of forecast error. Techniques from the literature in the field of predictive analytics and forecasting form the base that the methodology in this study is built on.

Drucker et al. (1997) introduces a new regression technique based on Vapnik's concept of support vectors. Known as Support Vector Regression (SVR), Drucker et al. apply the technique to a data set from the Boston Housing UCI database to predict the median price of housing in the Boston area. The study compares SVR with a committee regression technique known as bagging. Data is split into three groups: 80 percent of the set is assigned to the training set, 16 percent is assigned to the validation set, and the final 4 percent is assigned to the test set. The training set is used to fit new models being developed. The validation set is used to determine which of a group of models performs best on out-of-sample data. The main benefit of this out-of-sample data is to identify models that are overfit on the training set. Finally, the test set is used to verify the effectiveness of a model. Drucker et al end their study by comparing models from the two techniques by measuring their Mean Squared Error (MSE) and determine that Support Vector Regression was more effective.

Smith, Williams, and Oswald (2002) develops models predicting traffic flow. Using data from the United Kingdom's MIDAS database of the London Orbital Motorway (M25), models represent continuous one-minute periods of traffic flow over three months using time-series sessional ARIMA, nonparametric regression, neural

networks, and heuristics. The final analysis compares the Mean Absolute Percent Error (MAPE) of the models and determined that the ARIMA models were the most accurate.

In a study of demand for intermittent spare parts for a Chinese petrochemical enterprise, Hua and Zhang (2006) develops a collection of Support Vector Machines, Logistic Regression, and bootstrap method models. With over fifteen thousand individual spare parts used in this industry, the research was limited to four thousand parts with useful historical data and further focused on only thirty parts that met the definition of intermittent demand. The final analysis compares the models based on the mean error ratio to zero values, as well as the MAPE. The analysis concluded that support vector machines were the most effective.

Inman, Anderson, and Harmon (2006) conducted a study to predict when new technology would be introduced in new weapon systems as military fighter aircraft in the United States' fleet. This study used a multivariate collection of data measuring the maximum mach number, mean flying hours between failures, payload, and range of missiles to predict the first year in flight of new weapon system designs from 1944 to 1982. They compared the use of Technology Forecasting using Data Envelopment Analysis (TFDEA) and regression-based modeling by measuring the deviation and residuals of the models and determined that TFDEA produced the most accurate results. Not only is Inman et al's study a great example of comparing various forecasting models in an application concerning US fighter aircraft, but the study is also a clear example of using a multivariate data set to build models.

Yip, Fan, and Chiang (2014) describes a study of maintenance cost for construction equipment. Using both univariate and multivariate methods applying time series ARIMA, Box-Jenkins, and General Regression Neural Networks (GRNN) the study developed methods to quantify costs using fuel consumption data and maintenance costs at different monthly intervals. Comparing the Mean Absolute Percent Error (MAPE) of the final models determined that the neural network models were the best at predicting the costs of equipment.

Kaytez et al. (2015) models energy consumption, comparing models based on Least Squared Support Vector Machines (LS-SVM), multiple linear regression, and neural networks. The study uses data from electrical power and energy systems, and the models include factors such as installed capacity, gross electricity generation, population, and total subscribership. Two distinct training and test sets are used in the analysis. Sum Squared Error (SSE), Root Mean Squared Error (RMSE), Max Error, and Mean Absolute Percent Error (MAPE) are computed for each model. In the final comparison, R^2 and Mean Squared Error are used to determine that the Least Squared Support Vector Machines were the most effective.

The studies summarized in the preceding paragraphs describe several aspects of related published research that motivated the methods used in the research described in the following sections. These include: (1) selection of study data from large datasets, (2) appropriate partitioning of data into time buckets, (3) the use of both univariate and multivariate approaches, (4) careful use of data to avoid overfitting of models, and (5)

selection of appropriate error metrics and benchmarks for model evaluation and comparison.

III. Data Collection and Variables

Data Collection

For this research, MICAP and demand data were collected for cause code H MICAPs for the F-16 weapon system, as was recommended in the findings of the Perduco report (Johnstone & Chandler, 2019). After discussions with leadership and subject matter experts from the 438 Supply Chain Operations Squadron (SCOS) on what data was needed, the analysis cell from the 635 Supply Chain Operations Wing (SCOW) was enlisted to collect the required data. Using LIMS-EV (Logistics, Installations, and Mission Support-Enterprise View) an overarching Air Force enterprise database, the SCOW provided data on every MICAP from all cause codes in the Air Force enterprise from January 2017 to December 2020. From this data, the scope of the research was narrowed to the NSNs that had cause code H MICAPs over the three years. 1,546 NSNs were identified from this time period, and the scope was further focused on the top 100 NSNs with the most cumulative number of MICAP incidents over the three years, and the top 100 NSNs with the highest cumulative total MICAP Hours. Since many of the NSNs fell into both categories, a total of 136 NSNs were identified to be the main focus of the research. As the NSNs accumulating the most incidents and hours, they are important drivers of the exponential growth in cause code H MICAPs.

This data identified many critical elements related to cause code H MICAPS. In particular, this data includes total MICAP hours, the aspect that is used as the dependent (Y) variable in the models developed in this thesis research. But more information was desired on the status of the supply chain during the time period in which the MICAPs occurred. Therefore, demand and availability data for these NSNs were needed. After further discussions with the SCOW analysis cell, acquiring availability data for each NSN over the three years was deemed a significant challenge. However, a workaround was provided by the SCOW within the demand data. They provided a list of all Issue Requests (ISU) and Due-outs (DUO) for the NSNs from 2017 to 2019, pulled from LIMS-EV for all 1,546 NSNs first identified for F-16 cause code Hs. The ISUs are supply transactions where the requested part is available and shipped. The DUOs are requests for parts when they are not available and the due-out or backorder is created until the part becomes available. These values can be combined to represent the total demand. Alternatively, they can be used separately to represent demand with availability and demand when the item is unavailable. This list was filtered to only include transactions for F-16 weapon systems.

Finally, the last set of data requested from the SCOW included the MICAP and demand data for the top 136 NSNs for 2020. The SCOW provided the data in late 2020, covering January 2020 to November 2020, however, the November 2020 data were excluded due to a high volume of open MICAPs. For this reason, the available November 2020 data would not represent a "full month" accurately. 32 of the 136 NSNs requested did not have MICAP incidents in 2020.

Variables

The following list provides a brief description of the continuous and categorical data collected for potential use in this study and for breakouts of the final models developed.

- NSN – National Stock Number – label applied to a supply item within the federal supply system (AF/A4LR, 2020b)
- Month – the time unit of measurement used to aggregate the continuous variables
- MICAP – Mission Impaired Capability Awaiting Parts – an emergency demand for a supply item for a weapon system affecting its mission capability. (AF/A4LR, 2020a)
- Cause Code H – MICAP with historically established demand level for an NSN with no assets available in the supply chain to fill the requirement in allowable time limits (AF/A4LR, 2017)
- MICAP Incidents (N) – The number of MICAPs that occur for an NSN in a given time period
- MICAP Quantity – The quantity of the NSN needed to fill the given MICAP
- Total MICAP Hours – The amount of time in hours that the MICAP remains unresolved
- Price – The unit cost of an NSN for the MICAP
- Total Demand Quantity – the sum of the ISUs and DUOs for an NSN in a given time period

- ISU – Item Customer Issue Requests – total quantity of the requisitions for an NSN where the asset was available and distributed to the customer when the requisition was initiated (AF/A4LR, 2017)
- DUO – Due-Out – total quantity of the requisitions for an NSN where the asset was not available and thus not distributed to the customer when the requisition was initiated (AF/A4LR, 2017)
- ERRC Code – alphanumeric code that indicates if the NSN is expendable, (the item is used in a weapon, removed upon failure, and then discarded), or a reparable asset, (the item is used in a weapon, removed upon failure, repaired, and then made available to be reinstalled in the weapon) (AF/A4LR, 2020a)
- PD – An unknown variable from some of the LIMS-EV scripts used in the data collection with categories of various types of NSNs

Distributions and Correlations of the Data

Cause Code H MICAPs are influenced by some of the most stubborn root causes, such as DMSMS, AWPs, and long contract lead times. For this reason, Cause Code H MICAPs tend to accumulate excessive amounts of MICAP hours as they sit unresolved. Recall that the data considered here have been filtered to the top 136 NSNs that are leading drivers of Cause Code H MICAPs for the F-16. For some instances in this filtered set, NSNs have in excess of 100,000 MICAP hours. These extreme outliers may influence or distort the results and error rates of the models developed in this study. However, they are included in the study because these extreme values are some of the

most important to accurately predict, given the severe impact they have on readiness, aircraft availability, and mission capability.

See Appendix A for Distribution of MICAP Hours in the Data Set

One of the limitations of this study was the use of all NSNs in the creation of models regardless of their variation of MICAP hours. This is a known source of inaccuracy in P-values for statistical tests. However, this limitation was not completely ignored. An Analysis of Variance (ANOVA) was conducted on MICAP hours with the categorical variables as independent variables. Below is a table of the Least Squared Means (LSM) from this initial ANOVA.

Table 1. Least Squared Mean of MICAP Total Hours by Categorical Variables

Least Squared Mean of MICAP Total Hours by Categorical Variables		
Overall Data Set		2251
ERRC Code	Expendable	1579
	Reparable	2922
PD	Accessories	1167
	Aircraft	1772
	Common Avionics	6109
	Electronic Warfare	360
	Instruments	943
	Landing Gear	3132
	Other	4329
	Power Systems	553
	Propulsion	662
	Unknown	3479

See Appendix B for the ANOVA of Categorical Variables

From this initial ANOVA analysis for the categorical variables, a connecting letters report was generated to identify significant differences between the categories. The initial ANOVA identified few significant differences. To remedy the identified non-

homogeneous variance of error across the categories, the MICAP hours variable was transformed by adding one and taking the logarithm of the value. This remedied the non-homogeneity of the variance. A second ANOVA on the transformed values showed a significant difference between the two ERRC Codes, and among many of the PD categories. (The Fisher's test in JMP was used for the comparisons with, $\alpha = 0.05$). This ANOVA analysis provided motivation for the breakout of the model performance by categories shown in the Categorical Breakout of Best Models section of the Analysis and Results.

Using the multivariate tool in JMP the correlations and a scatterplot of all variables were created. These correlations represent the linear relationship between the variables and are shown below:

Correlations

	N	MICAP QTY	MICAP TOTAL HRS	UNIT PRICE	Demand QTY, Total	Demand QTY, DUO	Demand QTY, ISU
N	1.0000	0.9316	0.6802	0.1851	0.2926	0.5624	-0.1251
MICAP QTY	0.9316	1.0000	0.6575	0.1435	0.2969	0.5765	-0.1326
MICAP TOTAL HRS	0.6802	0.6575	1.0000	-0.0045	0.0384	0.1803	-0.1176
UNIT PRICE	0.1851	0.1435	-0.0045	1.0000	0.3584	0.2290	0.2838
Demand QTY, Total	0.2926	0.2969	0.0384	0.3584	1.0000	0.6998	0.7342
Demand QTY, DUO	0.5624	0.5765	0.1803	0.2290	0.6998	1.0000	0.0287
Demand QTY, ISU	-0.1251	-0.1326	-0.1176	0.2838	0.7342	0.0287	1.0000

The correlations are estimated by Row-wise method.

Figure 1. Correlations of Continuous Variables

See Appendix C for Scatterplot of Continuous Variable Correlations

Since the dependent (y) variable being predicted in the regression models is MICAP Total Hours, the focus is on the other correlations between this variable and each of the independent (x) variables. MICAP Incidents (N) has the strongest correlation with MICAP Total Hours at 0.6802, followed by MICAP Quantity at 0.6575 then Due-Out

demand quantity at 0.1803, and finally total demand quantity with the lowest at 0.0384. Issue demand quantity and unit price both had negative correlations of relatively small magnitude. These variables with weak correlations to MICAP Total Hours, play almost no role in the models developed.

To take advantage of the correlation of MICAP Total Hours between time periods (i.e., time-series effects), its previous values should be useful in building regression models. Other variables with substantial correlations to MICAP Total Hours within a period were also promising (MICAP data, MICAP incidents (N), and MICAP Quantity). Although the due-out demand quantity and total demand quantity showed significantly lower correlations in the preceding analysis, they were still non-zero. Thus they were included in the model building process.

Data Structure in JMP

The raw data sets were lists of individual MICAP and supply transactions. These spreadsheets were copied from their original excel files into JMP spreadsheets. The *tabulate* and *join* functions in JMP were primarily used to organize and aggregate data as needed. First, the list of raw data was aggregated monthly using the JMP *tabulate* function. This allowed the creation of new tables with the sum value of all continuous variables for each NSN in every month in the data set. Then using the *join* tables function, another table was created using a cartesian join of a list of all NSNs and a list of every month throughout the time period, resulting in a row for every NSN at every month and no other data. This table was then joined with the tables with the numerical data so that months where a MICAP might have a value of zero in the continuous data would be

represented. Finally, this data was merged with the categorical variables using another cartesian *join* function so that the final sheet included all variables for every NSN and month in the data set.

Once the data table with all variables broken out by each NSN for each month was created, new columns were created to include previous months' data. This was done in order to have each month's MICAP Total Hours be in the same row with values for each variable from one to six months previous to the value. To do this, the columns were sorted by month, and then by NSN. Then six empty columns were created next to every variable. In those columns, formulas were entered with an offset in the table from one to six rows. Since the table was sorted by time, this added the required information to each row.

Because the first six entries for each NSN would be pulling information from a different NSN, the rows for January 2017 to June 2017 all had to be excluded on the spreadsheet, so their information would not be used in the model building. Also excluded were the row for November 2020. Since the MICAP and demand data for that month was received before the end of November 2020, it was incomplete and could misrepresent actual cumulative values for the month.

IV. Methodology

Research Questions

The purpose of this research was operationalized in the following research questions:

- Research Question 1: Can an accurate set of prediction models be developed to enable proactive efforts across the supply chain to reduce the occurrence of future cause code H MICAPs?
- Research Question 2: Are predictions further in the future more useful than models that generate more short-term predictions?
- Research Question 3: Should models be developed on a broad set of NSNs or should models be tailored to fit more narrow sets of data to improve accuracy?

Model Construction

Training, Validation, and Test Sets

Fundamental Principles

The data set was partitioned into three sets using the validation function in JMP: The first set is the *training set*, which is used to construct and fit regression models. The second set, *validation*, is used in the selection process to determine which of the many models created using the same technique is best. Selection within one technique is based primarily on the lowest Root Average Squared Error (RASE). Model R^2 values were also considered. Finally, the third portion of the data, the *test set*, is used in the final model comparison between different techniques to determine which of the various types of models is best. This final determination was on separate data from the training and validation sets. This three-tiered model evaluation and comparison procedure was utilized to avoid models that are overfit on the training data, leading to less useful models (Klimberg et al., 2018).

Operations in JMP

Similar to the study by Drucker et. al, ratios of 0.80 (80% of the data) for the training set, 0.16 (16%) for the validation set, and 0.04 (4%) for the test set were used to partition the data (Drucker et al., 1997). JMP's validation functionality allows for a categorical variable column to be added to the data set and used for this purpose. The preferred ratio is entered into a dialog in JMP. JMP randomly assigns rows in the data to one of the three categories (training, validation, test) based on the ratio. With this variable created, the data set was ready to be used to create numerous models using the techniques described in detail below.

Linear Regression

Fundamental Method

Linear Regression is a technique that forms models centered on the correlation between a dependent y variable and one or more independent x variables. The model can be expressed as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ \dots + \beta_n x_n + \varepsilon$$

where y is the dependent variable, n is the number of x variables considered, β_0 is the intercept, β_i is the rate of change for the independent variable x_i where $i \in \{1, \cdots, n\}$ and ε is the error term (Bowerman et al., 2005). Models that only use one independent x variable in the model are considered Simple Linear Regression. Models using two or more x variables are considered Multiple Regression.

Probability value or p-value for model terms is used to determine whether the value for β_n is statistically significant. This p-value is computed based on a *t* distribution

for a hypothesis test with a null hypothesis of H_0: $\beta_n = 0$ and an alternative hypothesis H_a: $\beta_n \neq 0$. It is typical to use $\alpha = 0.05$ as a threshold to determine whether to reject the null hypothesis H_0. If the p-value is greater than 0.05, then there is a failure to reject the null hypothesis and it cannot be disproven that $\beta_i = 0$. Therefore, the x_i associated with the parameter estimate β_i would not be significant in the regression equation, resulting in a product of zero. However, if the p-value is less than or equal to 0.05, then there is strong evidence to reject the null hypothesis. Based on this, the alternative hypothesis is concluded to be true. This process results in a useful estimate of the parameter β_i that can be used in a regression equation. (Bowerman et al., 2005).

The overall effectiveness is measured by the simple coefficient of determination (R^2) and the error rate. In this study, the Root Average Squared Error (RASE) is evaluated and used for model comparisons.

R^2 measures "the strength of the linear relationship between the response and the predictor" and can be calculated by dividing the explained variation, or variation explained by the model, by the total variation in the response of the model. It can be calculated using the following equation:

$$R^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2}$$

Where $\hat{y}_i$ is the predicted value, $\bar{y}$ is the sample mean, and y_i is the actual value. (Bowerman et al., 2005).

The Root Average Squared Error is "the square root of the mean squared prediction error" and is computed by finding the square root of the Sum Squared Error (SSE), the unexplained variation divided by the number of observations (n):

$$RASE = \sqrt{\frac{SSE}{n}} = \sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{n}}$$

Unlike the Root Mean Squared Error (RMSE) the RASE does not account for degrees of freedom (*The Model Comparison Report*, 2020).

Operations in JMP

Creating regression models in JMP is a simple process. First, a data table is created with one column of y variables and one or more columns of x variables. Then, the ‘fit model’ item is selected and then a ‘Model Specifications’ window appears with a list of columns from the data table. One column must be selected as the y variable and one or more columns can be selected as ‘model constructs’ which are the independent x variables. If there are two or more columns of data being used in the model, interactions among the x variables can also be introduced. This creates additional terms in the regression function. In the jump model dialogue, interaction terms are created by selecting all variables that will be included in the interaction terms, selecting macros, and the factorial to degree function, with degree set to a value of two. This adds model terms for all possible two-factor interactions. Once the columns have been selected a column can be selected for the training, validation, and test set. Separate models can also be created using a similar process when a categorical variable’s column is selected in the ‘by’ field. Finally, the “personality” item is set as ‘standard least squares’ and the emphasis is set at ‘minimal report’.

Once the model specifications are set, the user selects run, and a model report is generated. In this report all parameter estimates, p-values, the R^2 and RASE can be

found to evaluate the performance of the model. The model can be saved and a new column of prediction values for the model can be added to the spreadsheet for further evaluation.

Stepwise Multiple Regression

Fundamental Method

Creating collections of models from numerous continuous variables at multiple time intervals, and with potential interactions can quickly create large numbers of increasingly complex models. The complexity of navigating this enormous potential model space requires a more systematic approach to model creation and evaluation. For this reason, a stepwise regression tool was used. Stepwise regression allows a systematic exploration of a mix of many variables. It can be used to create very compact models only using as many variables as are needed. JMP implements stepwise regression and employs "an iterative model selection procedure, where at each step a single independent variable is added to or deleted from a regression model, and a new regression model is evaluated." (Bowerman et al., 2005).

With a set dependent (y) variable, the stepwise tool requires a list of potential independent (x) variables to consider in a multiple regression model. Once this is executed, a new window in JMP opens with the Stepwise Regression controls. The p-value from hypothesis tests for each potential model term can guide the search. The magnitude of the model coefficient for each variable is related to the correlation of the potential independent x variable to the dependent y variable. The user selects thresholds and these p-value thresholds are used as boundaries to create new models. Stepwise

regression can work both forward (adding terms) or backward (removing terms) from an existing model.

- In forward regression, the P-value for each potential new model term is compared to an entry bound (e.g. 0.05). If the P-value for a potential new model term is less than the entry bound, it is added to the current model.
- In backward regression, the P-value for each existing model term is compared to an exit bound (e.g. 0.25). If the P-value for an existing model term is greater than the exit bound, it is removed from the current model.

In its simplest form, forward regression starts with only a constant term in the model and adds x variables to the model using the entry bound. In backward regression, the model starts with all available x variables and removes them using the exit bound. A use of one or both of these approaches enables a focused exploration of the potential model space. (*Variable selection in multiple regression*, n.d.).

With a higher number of independent x variables being included in the model, there is a risk of multicollinearity. The Variance Inflation Factor (VIF) measures the multicollinearity of the j^{th} independent variable, x_j, with other x variables using the following equation:

$$VIF_j = \frac{1}{1 - R_j^2}$$

R_j^2 is the coefficient of determination for the regression model that relates x_j to all the other independent variables. x_j variables with a VIF greater than or equal to ten are

generally considered severely multicollinear and should be removed from the multiple regression model (Bowerman et al., 2005).

Operations in JMP

The model is first created using the *fit model* model specification. This opens up a new window like the one used to create a standard regression model described previously with set y value and validation selected. However, all potential x variables are selected for *construct model effects* and the model *personality* is changed to 'Stepwise'. Once this is executed, a new window in JMP opens with the Stepwise Regression controls. In this study, the stopping rule of 'P-Value threshold' is used set with the JMP default setting at a probability to enter and leave at 0.25 and 0.1 respectively. When the stepwise tool is executed in JMP, the best model resulting from forward or backwards steps can be saved and a new column of prediction formula values can be added to the data table.

The VIF of x variables can be found in the model report under 'parameter estimates'. If the multicollinearity of a variable is considered to be severe, the *fit model* dialogue can be reopened, the x_j can be removed, and then the model can be run again, saved, and the new prediction formula can be added to the spreadsheet.

Principle Component Analysis

Fundamental Method

Principle Component Analysis (PCA) is a technique with two objectives in a large multivariate dataset; first to reduce a large collection of variables into a smaller set, and second, to reduce correlations between variables. From two original x variables:

> PCA assesses the structure of correlations among the variables by defining a set of underlying dimensions called components… The method of PCA transforms the variables into uncorrelated variables (with no common variation) that still preserves two sources of unique variation so that the total variation in the two variables remains the same (Klimberg & McCullough, 2018).

Since the creation of the new components is based only on the variation of the *x* variables, the dependent *y* variable for the future regression model is not factored into the process.

If *k* number of *x* variables are selected to be transformed, *k* new components will be generated by the procedure. It is up to the user to determine a satisfactory number, *n*, out of the *k* components to use in the final regression models. This is not an exact science but is based on the following three methods evaluating eigenvalues and cumulative percentage of variation:

1. Finding an 'elbow' in the scree plot of eigenvalues by components and selecting only the components before the 'elbow'.
2. Only selecting components with eigenvalues greater than one.
3. Only selecting the first set of *n* components that account for 70 to 80 percent of the cumulative variation.

An eigenvalue is a weight of the measure of variation in the original set of variables for each new principal component (Klimberg & McCullough, 2018). Once a satisfactory number of principal components is determined, they can be used as a new set of *x* variables in regression models.

Operations in JMP

The set of principal components are automatically generated by JMP by selecting analyze, multivariate methods, then principal components. In the *model constructs* of the principal components, the list of *k* independent *x* variable columns is selected for the *y* columns of the PCA and then executed. The set of *k* components will be created in a new window with a list of eigenvalues, cumulative percentage of variation, and the scree plot.

See *A*ppendix D for Example List *o*f Eigenvalues *a*nd Scree Plot

Once the *n* components out of the k are chosen they can be saved as new columns on the original data spreadsheet. These new columns are used simply as new independent x variables to predict a dependent y variable in a multiple regression model.

Experimental Design

Every model in this study is created using regression in JMP. All models will be created to predict MICAP Total Hours; the dependent *y* variable at time *t*. Models will be used to predict MICAP hours in three intervals:

- three months in the future using data from three to six months previous to the y value
- four months in the future using data from four to six months previous
- five months in the future using data from five to six months previous.

Since the data is aggregated monthly, and each row in the spreadsheet representing one month for the NSN, the column of MICAP Total Hours is used as the *y* variable in the regression models. By having other columns of previous *x* values all on the same row, predictions for future values of the *y* variable can be calculated.

Table 2. Interval of Prediction for Regression Models

Interval of Prediction for Regression Models			
Interval between predicted value and latest predictor input	3 months in the future	4 months in the future	5 months in the future
Range of predictor input data	3 to 6 months previous	4 to 6 months previous	5 to 6 months previous

Four categories of models will be used to create models in four phases:

- Simple linear regression – using one column of a continuous variable from a month previous to the predicted value
- Multiple Regression – using columns of the same continuous variable from multiple months previous to the predicted value
- Stepwise Multiple Regression – using multiple continuous variables from multiple months previous to the predicted value
- Principle Component Analysis – multiple regression models using columns of principal components generated from the multiple continuous variables over the given interval of prediction

Table 3. Regression Model Types

Regression Model Types			
Phase I - SLR	Phase II - MR	Phase III - SW-MR	Phase IV - PCA
Simple Linear Regression of one *x* variable type from a single month previous to *y*	Multiple Regression of one *x* variable type from multiple months previous to *y*	Stepwise Multiple Regression of multiple types of *x* variables from multiple months previous to *y*	Multiple Regression of Principal Components created from multiple types of *x* variables from multiple months

The following continuous variables will be used throughout the regression models:

Table 4. Continuous Variables' Use in Regression Models

Continuous Variables' Use in Regression Models		
Variable Category	Variable Type	*x* or *y* variable
MICAP Data	MICAP Total Hours (*t*)	*y*
	MICAP Total Hours (*t-1,…t-6*)	*x*
	MICAP Incidents (*t-1,…t-6*)	*x*
	MICAP Quantity (*t-1,…t-6*)	*x*
	Unit Price	*x*
Demand Data	Total Demand Quantity (*t-1,…t-6*)	*x*
	Due-out Quantity (*t-1,…t-6*)	*x*

Numerous models will be created in each of these four stages over the three intervals of prediction. The RASE from the validation set will be evaluated for each model.

The best twelve models with the lowest justifiable validation RASE were selected for the analysis in the test set. Each of the four model categories will have one best model for each of the three intervals of prediction, totaling in twelve final models.

Phase I – Simple Linear Regression Models

This phase used Simple Linear Regression with single variables from one to six months previous to the *y* value. For example, the single *x* variable of MICAP Total Hours from $(t-3)$ was used to predict the y variable of present MICAP Total Hours in period *t*. There are six such models using previous Total MICAP Hours. A similar process was followed for MICAP Incidents (N), MICAP Quantity, Total Demand Quantity, and Due-Out Quantity. This resulted in a total of thirty simple linear regression models.

Table 5. Simple Linear Regression Models Created

Simple Linear Regression Models Created		
Model Number	*x Variable Type*	Months Previous to *y*
1	MICAP Total Hours	$(t-1)$
2		$(t-2)$
3		$(t-3)$
4		$(t-4)$
5		$(t-5)$
6		$(t-6)$
7	MICAP Incidents	$(t-1)$
8		$(t-2)$
9		$(t-3)$
10		$(t-4)$
11		$(t-5)$
12		$(t-6)$
13	MICAP Quantity	$(t-1)$
14		$(t-2)$
15		$(t-3)$
16		$(t-4)$
17		$(t-5)$
18		$(t-6)$
19	Total Demand Quantity	$(t-1)$
20		$(t-2)$
21		$(t-3)$
22		$(t-4)$
23		$(t-5)$
24		$(t-6)$
25	Due-Out Quantity	$(t-1)$
26		$(t-2)$
27		$(t-3)$
28		$(t-4)$
29		$(t-5)$
30		$(t-6)$

Phase II – Multiple Regression Models

Multiple Regression models in this study were created in three categories, (1) by the continuous variable used as *x* variables in the regression, (2) by the interval of

prediction, and (3) by models with and without interactions of the x values. The MICAP Total Hours, MICAP Incidents (N), MICAP Quantity, and Due-Out Quantity were all used in separate models, with data ranging from three to six months in the past, with and without interactions, for a total of twenty-four models (4×3×2).

Table 6. Multiple Regression Models Created

Multiple Regression Models Created			
Model Number	x Variable Type	Months Previous to y	Interactions between x variables
1	MICAP Total Hours	$(t-3), \ldots, (t-6)$	No
2			Yes
3		$(t-4), \ldots, (t-6)$	No
4			Yes
5		$(t-5), (t-6)$	No
6			Yes
7	MICAP Incidents	$(t-3), \ldots, (t-6)$	No
8			Yes
9		$(t-4), \ldots, (t-6)$	No
10			Yes
11		$(t-5), (t-6)$	No
12			Yes
13	MICAP Quantity	$(t-3), \ldots, (t-6)$	No
14			Yes
15		$(t-4), \ldots, (t-6)$	No
16			Yes
17		$(t-5), (t-6)$	No
18			Yes
19	Due-Out Quantity	$(t-3), \ldots, (t-6)$	No
20			Yes
21		$(t-4), \ldots, (t-6)$	No
22			Yes
23		$(t-5), (t-6)$	No
24			Yes

Phase III – Stepwise Multiple Regression

Sets of models were created with three varying factors: (1) time range from three to six months previous data included to only five to six months previous, to cover the three intervals of prediction; (2) only continuous MICAP variables included and those with due-out demand variables also included; and (3) with and without interactions of all x terms. Thus, a total of twelve models (3×2×2) were created through this process by running the groups of x variables through the forward and backward stepwise process. In the model validation phase, the best of the four models for each interval of prediction was selected based on the error rate on the validation set. The three best models from the validation phase, one for each of the three intervals of prediction, are compared in the final analysis in the test set.

Table 7. Stepwise Multiple Regression Models Created

Stepwise Multiple Regression Models Created			
Model Number	*x* Variable Categories Entered into Stepwise Tool	Months Previous to *y* Entered into Stepwise Tool	Interactions between *x* variables
1	Only MICAP Data	$(t-3), \ldots, (t-6)$	No
2			Yes
3		$(t-4), \ldots, (t-6)$	No
4			Yes
5		$(t-5), (t-6)$	No
6			Yes
7	MICAP Data and Due-Out Quantity	$(t-3), \ldots, (t-6)$	No
8			Yes
9		$(t-4), \ldots, (t-6)$	No
10			Yes
11		$(t-5), (t-6)$	No
12			Yes

Phase IV – Principal Component Analysis Models

Six sets of principal components were created using the PCA function in JMP. The final number *n* out of *k* components was selected using the rules for evaluating the Scree Plots, eigenvalues, and cumulative percentage of variation. These sets were created for each of the three prediction time periods using data from three to six months, four to six, and five to six months back. Each of the three time periods had two PCA sets created using only continuous MICAP data and a set with both MICAP and due-out demand data.

All of these models were evaluated using a range of their satisfactory principal components. Half of the sets had a max of four satisfactory components and the other half had a max of five components. Components selections considered eigenvalues, cumulative percentage of variation, and the scree plots. These components were used in creating a set of multiple regression models using the component's combinations of the independent x variables to predict MICAP Total Hours. Each of the six sets of principal components proved to be the most effective (lowest RASE) in the validation set when all satisfactory components were included in the regression models.

Below is a list of the six final models from the six sets of principal components, which use all of their satisfactory principal components:

Table 8. Final PCA Multiple Regression Models Created

Final PCA Multiple Regression Models Created		
Model Number	x Variable Categories Entered into PCA Tool	Months Previous to y Entered into PCA Tool
1	Only MICAP Data	$(t-3), \dots, (t-6)$
2		$(t-4), \dots, (t-6)$
3		$(t-5), (t-6)$
4	MICAP Data and Due-Out Quantity	$(t-3), \dots, (t-6)$
5		$(t-4), \dots, (t-6)$
6		$(t-5), (t-6)$

Final Analysis

Three models from each phase, one for each interval of prediction will be evaluated in the test set. The RASE in the test set for these twelve models will then be evaluated to determine the best overall prediction model.

Additionally, the final twelve models were reran broken out individually by the categorical variables ERRC Code and PD. This evaluation on more specific subdivisions of the data allowed for a comparison of the models' performance across the categories.

See Table 1. for List of Categorical Variables

To accomplish this, the models' prediction formulas were saved into a new column. These results were assessed using the JMP model comparison function and broken out by category. The test set RASE was used to determine which models have the lowest error rate and performed the best.

The final analysis includes an evaluation of a scaled ratio between the twelve models' RASE and the mean of MICAP Total Hours across the data set. This ratio is then compared to a similar ratio of a RASE taken from a naive model that predicts zero MICAP Hours for every NSN in every month. This comparison is also broken out for the

categorical results with ERRC Code and PD, and their corresponding RASE and means. This allows for a determination of whether the final results from the models are more useful than having no prediction at all.

V. Analysis and Results

Chapter Overview

Using the four regression methods described previously Simple Linear Regression, Multiple Regression, Stepwise Multiple Regression, and Principal Component Analysis, models were developed with the data set across the three intervals of prediction. The model comparison function was then used to determine the three best models, for the three intervals, for each method based on their errors in the validation set. The resulting twelve best models were then analyzed in the model comparison again but this time using the test set. This allowed for a determination of the overall best models in each interval. The twelve models were finally rerun broken out by categorical variables. The model comparison was again used to determine how the models performed within the various groups.

Simple Linear Regression

Simple linear regression was first used to show how prediction models performed worse as the time gap between the *y* value MICAP Total Hours and the *x* variables increased. Models using more recent data (eg. $t-3$)consistently performed best with the lowest RASE. These error rates consistently increased with each more distant month.

This was true for all models except for those based on Total Demand Quantity. Below are the RASE for the models for each variable over the three intervals of prediction:

Table 9. Simple Linear Regression Models' RASE from Validation Set

Simple Linear Regression Models' RASE in Hours from Validation Set			
	$(t-3)$	$(t-4)$	$(t-5)$
MICAP Total Hours	6189	7297	7798
MICAP Incidents	7580	7904	7930
MICAP Quantity	7370	7678	7798
Due-Out Quantity	7904	7925	7934
Total Demand Quantity	7993	7999	7993

See Appendix E for Complete Simple Linear Regression Model Comparisons

As assumed, the best individual predictor of MICAP Total Hours were previous month's MICAP Total Hours values. For $(t-3)$, $(t-4)$, and $(t-5)$, the RASE for each model were 6189, 7297, and 7798 hours respectively. The number of MICAP Incidents (N) and MICAP Quantity had slightly higher errors but mirrored the increasing trend as the interval from time *t* increased. For $(t-3)$, $(t-4)$, and $(t-5)$, MICAP Incidents had RASE of 7580, 7904, and 7930. For the same range, MICAP Quantity's models had a RASE of 7678, 7798, and 8110. Due-Out Quantity performed similarly to MICAP Quantity and Total Demand Quantity with RASE values of approximately 8000 for each month in the range with barely any variation across all six months with respect to its error rates. Based on this result, it was decided that this variable would not be included in any of the additional models developed. The variable may be flawed since it

is the summation of due-out and issue requests, and the issue request variables alone were shown to be uncorrelated with the *y* value, MICAP Total Hours.

Multiple Regression

Multiple regression was used to create a total of twenty-four models. Each model used only one type of *x* variable over multiple months. Four variable types were used in models with and without interactions between the time periods. Below are the RASE for each model created using Multiple Regression:

Table 10. Multiple Regression Models' RASE from Validation Set

Multiple Regression Models' RASE in Hours from Validation Set			
	$(t-3), \dots, (t-6)$	$(t-4), \dots, (t-6)$	$(t-5), (t-6)$
MICAP Total Hours	6176	7300	7509
MICAP Total Hours with interactions	6515	7487	7790
MICAP Incidents	7584	7908	7966
MICAP Incidents with interactions	7533	7911	7996
MICAP Quantity	7388	7709	7870
MICAP Quantity with interactions	7402	7741	7909
Due-Out Quantity	7916	7937	7967
Due-Out Quantity with interactions	8010	8040	7966

See Appendix F for Complete Multiple Regression Model Comparisons

Overall, MICAP Total Hours models performed the best, followed by MICAP Incidents and lastly by MICAP Quantity. Except for the interval $(t-3), \dots, (t-6)$ with MICAP Incidents, interactions in the model did not improve the error rate compared

to models without interactions. Also, Due-Out Quantity produced essentially equal results in the interval $(t-5),\ldots,(t-6)$ with and without interactions. Regardless, the three MICAP Total Hours models with RASE of 6176, 7300, and 7509 performed the best and are used in the final model comparison with the test set of data.

Stepwise Multiple Regression

For the stepwise multiple regression models, all continuous MICAP variables were considered in the stepwise tool, in addition to unit price for the first time. Additionally, models were run with and without due-out quantity, and with and without interactions. Separate models were developed with all these potential combinations across the three intervals. The models created using forward and backward stepwise regression were generally consistent and had the lowest RASE in the validation set. The RASE for the best models created can be found below:

Table 11. Stepwise Multiple Regression Models' RASE from Validation Set

Stepwise Multiple Regression Models' RASE in Hours from Validation Set			
	$(t-3),\ldots,(t-6)$	$(t-4),\ldots,(t-6)$	$(t-5),(t-6)$
Only MICAP variables	6210	7390	7562
Only MICAP Variables with interactions	6201	7358	8438
MICAP Variables and Due-Out Quantity	6239	7349	7554
MICAP Variables and Due-Out Quantity with interactions	6237	7315	8475

See appendix G for Complete Stepwise Multiple Regression Model Comparisons

The results from these models' error rates are not consistent. In the first intervals of prediction, interactions improve the models slightly, however in the $(t-5), \ldots, (t-6)$ interval, they make the models worse. However, when the due-out demand data is considered in the models, five out of six of the similar model pairs show that adding the due-out data makes the model worse. It was decided that the models to be used in the final test model comparison would be the three created only using MICAP variables without interactions. This was based on the philosophy that simpler is better, all other things being equal. The chosen models were the most simplistic, outweighing the minuscule improvements achieved by expanding to more complicated models.

Principle Component Analysis

Six sets of principal components were created over the three intervals of prediction using MICAP variables, both with and without Due-Out quantity. Sets of multiple regression models we first run in order to confirm that models with all the satisfactory number of principal components created the most accurate models.

See Appendix H for Comparisons of Models using Some to All Components

This resulted in the comparison of six multiple regression models from the six sets of principal components, with all justifiable components included as independent x variables. The comparison of RASE can be seen below:

Table 12. Principal Component Analysis Models' RASE from Validation Set

Principal Component Analysis Models' RASE in Hours from Validation Set			
	$(t-3), \ldots, (t-6)$	$(t-4), \ldots, (t-6)$	$(t-5), (t-6)$
Only MICAP variables	6747	7468	7721
MICAP variables and Due-Out Quantity	6857	7513	7749

See Appendix I for complete PCA Regression Model Comparisons

As expected, the models with data from the more recent time period performed better than those with less recent data. Across the board, the models with only continuous MICAP data performed better than those that also included due-out demand data.

Test Set Model Comparison

From the four categories of methodologies used, simple linear regression, multiple regression, stepwise multiple regression, and principal component analysis, three models were selected for the final analysis comparing the RASE of each model on the test set of data. One model in each of the four categories was selected for each range. These formed predictions of MICAP hours using x variables from three to six months before the predicted y values, four to six months before the y values, and five to six months before the y values. For both simple linear regression and multiple regression, the best models with the lowest error rates for all three intervals were the ones using previous MICAP Hours as the independent *x* variable to predict MICAP Hours. For the Stepwise multiple regression and Principal Component Analysis, the best models for all

time ranges used only continuous MICAP data (without due-out demand variables) and without interactions. These twelve models were then compared on the test set of data and their RASE are as follows:

Table 13. Overall Best Models' RASE from the Test Set

Overall Best Models' RASE in Hours from the Test Set			
	$(t-3), \ldots, (t-6)$	$(t-4), \ldots, (t-6)$	$(t-5), (t-6)$
Simple Linear Regression	7961	8460	8349
Multiple Regression	7746	8469	8276
Stepwise Multiple Regression	7657	8260	8255
Principal Component Analysis	8591	8554	8306

See Appendix J for Model Outputs of Twelve Overall Best Models

See Appendix K for Complete Overall Best Model Comparisons

Across the board for the three intervals of prediction, Stepwise multiple regression created the best models in the test set with the lowest error rate measured by RASE. The Simple Linear Regression models and Multiple Regression models performed similarly, and the Principal Component models had the highest error rates. All twelve models show higher error rates in the test set compared to the validation set, with the models for the interval $(t-3), \ldots, (t-6)$ see an increase in RASE from 23-29 percent compared to the validation set, $(t-4), \ldots, (t-6)$ with an increase from 12-16 percent, and $(t-5), (t-6)$ with a 7-10 percent increase RASE. This increase across

all intervals may be due to the smaller sample size of the test set sample size compared to the validation set.

To create a metric for comparison, a “completely uninformed” forecast was generated. This forecast consists of a value of “0” as a predicted value in all cases. The RASE for this case was calculated across all NSN for comparison to the other forecasts as a benchmark. This is shown in Table 14 in the “Zero Hours Prediction” column. In addition, a scaled metric is introduced to more accurately compare the effectiveness of all models by dividing the RASE by the least squared mean (LSM) of MICAP total hours in the sample set of data. This is shown in Table 14 in the row labeled “RASE/LSM”.

Table 14. Comparison of Overall Best Models with Scaled Ratio

Comparison of Overall Best Models vs. Predicting Zero MICAP Hours from Test Set					
Test Set Metrics	$(t-3), \dots,$ $(t-6)$	$(t-4), \dots,$ $(t-6)$	$(t-5),$ $(t-6)$	Zero Hours Prediction	
Overall	SW-MR	SW-MR	SW-MR		Model
	0.0919	-0.057	-0.055		R^2
	7656	8260	8255	8389	RASE
	2980	2980	2980	2980	LSM
	2.6	2.8	2.8	2.8	RASE/LSM

While the stepwise regression models from all three intervals have lower RASE than an overall prediction of zero MICAP total Hours, when compared using the scaled ratio, unfortunately only the model with data from the earliest interval $(t-3), \dots, (t-6)$ is a clear improvement at 2.6 when comparing the RASE of the models with the Zero MICAP Hours prediction and comparing the scaled ratio of RASE divided by LSM. The best model results for the intervals $(t-4), \dots, (t-6)$ and $(t-5), (t-6)$ both had a

scaled ratio of 2.8 equal to the Zero Hour Prediction, indicating that these models are likely no more accurate than the naive prediction of zero hours.

Categorical Breakout of Best Models

The twelve best regression models were rerun separately on each subset of the categorical variables ERRC Code and PD. This demonstrated how the models perform on smaller subsets of the data compared to the entire data set as a whole, from which the models were fit to predict overall. The RASE within each subset was compared using the test set data. Due to the small sample size of some of the PD categories, these results may require further investigation.

A dramatic split in RASE was found when the models were run on the two ERRC Code's test sets as seen below:

Table 15. Comparison of Best Models by ERRC Code with Scaled Ratio

Comparison of Best Models by ERRC Code vs. Predicting Zero MICAP Hours from Test Set						
Test Set Metrics		$(t-3), \ldots, (t-6)$	$(t-4), \ldots, (t-6)$	$(t-5), (t-6)$	Zero Hours Prediction	
ERRC Code	Expendable (45 NSNs)	PCA	SLR	PCA		Model
		-0.225	-0.135	-0.205		R^2
		4951	4766	4909	8389	RASE
		1579	1579	1579	2980	LSM
		3.1	3.0	3.1	2.8	RASE/LSM
	Reparable (91 NSNs)	SW-MR	SW-MR	SW-MR		Model
		0.1346	-0.045	-0.03		R^2
		8452	9286	9219	8389	RASE
		2922	2922	2922	2980	LSM
		2.9	3.2	3.2	2.8	RASE/LSM

For Expendable assets, two Principal Component Analysis models and a Simple Liner regression model were the most accurate over the three intervals of prediction and the RASE plummeted below 5000 hours for all three. Additionally, the ratio comparing errors to the zero hours predictor shows that these models are an improvement to having no prediction.

For reparable assets, the results worsened compared to the overall data set. The RASE for each of the three intervals increased by almost 1000 hours. The benchmark ratio to the zero hours predictor indicated that using these models might be worse than having no prediction at all. This also suggests that the variation in MICAP total hours for reparable items is greater than for expendable assets. Oddly, there are some models from the models using data further in the past with a RASE lower than the models with more recent data. This likely indicates that new models should likely be created to fit these specific data sets.

Compared to the best models', stepwise multiple regression, RASE in the overall test set, the expendable sample sees a large decrease in error by 35 percent for $(t - 3), \dots, (t - 6)$, 42 percent for $(t - 4), \dots, (t - 6)$, and 41 percent for $(t - 5), (t - 6)$ while reparable assets see increased error rates by 10, 12, and 12 percent for the same intervals, respectively.

Unfortunately, the scaled ratio for the expendable and reparable sets are greater than the scaled ratio in the overall data set. While this is not a direct comparison using the same LSM, it is an effort to show if there is any evidence that using the developed models could be more useful on specific sets of NSN. The increased ratio for the ERRC

codes may indicate that models should be developed on those smaller specific data sets, knowing that the mean of the two groups are significantly different.

There is a much greater variety of results when the twelve models are evaluated separately for each PD category as seen below in Table 16. No model type is consistently the best for each category or interval of prediction:

Table 16. Comparison of Best Models by PD with Scaled Ratio

Comparison of Best Models by PD vs. Predicting Zero MICAP Hours from Test Set						
Test Set Metrics		$(t-3), \dots, (t-6)$	$(t-4), \dots, (t-6)$	$(t-5), (t-6)$	Zero Hours Prediction	
PD	Accessories (15 NSNs)	MR	PCA	PCA		Model
		0.2379	0.2226	0.2881		R^2
		2625	2651	2537	8389	RASE
		1167	1167	1167	2980	LSM
		2.2	2.3	2.2	2.8	RASE/LSM
	Aircraft (31 NSNs)	SW-MR	PCA	SLR		Model
		-2.094	-2.082	-1.004		R^2
		8833	8815	7108	8389	RASE
		1772	1772	1772	2980	LSM
		5.0	5.0	4.0	2.8	RASE/LSM
	Common Avionics (12 NSNs)	PCA	PCA	PCA		Model
		0.116	0.0161	0.099		R^2
		7729	8154	7803	8389	RASE
		6109	6109	6109	2980	LSM
		1.3	1.3	1.3	2.8	RASE/LSM
	Electronic Warfare (3 NSNs)	PCA	SW-MR	PCA		Model
		0.3198	-5.53	-4.919		R^2
		206	638	608	8389	RASE
		360	360	360	2980	LSM
		0.6	1.8	1.7	2.8	RASE/LSM
	Instruments (18 NSNs)	SLR	SW-MR	PCA		Model
		-2.586	-1.307	-3.431		R^2
		1551	1244	1723	8389	RASE
		943	943	943	2980	LSM
		1.6	1.3	1.8	2.8	RASE/LSM

	Landing Gear (5 NSNs)	SW-MR	PCA	PCA		Model
		0.4196	-0.048	-0.019		R^2
		25568	34351	33884	8389	RASE
		3132	3132	3132	2980	LSM
		8.2	11.0	10.8	2.8	RASE/LSM
	Other (37 NSNs)	SLR	SLR	PCA		Model
		-0.36	-0.17	-0.01		R^2
		5653	5243	4870	8389	RASE
		4329	4329	4329	2980	LSM
		1.3	1.2	1.1	2.8	RASE/LSM
	Power Systems (4 NSNs)	MR	MR	PCA		Model
		-6.468	-9.061	-8.248		R^2
		900	1044	1001	8389	RASE
		553	553	553	2980	LSM
		1.6	1.9	1.8	2.8	RASE/LSM
	Propulsion (1 NSN)	MR/SW-MR/PCA	MR/SW-MR/PCA	MR/SW-MR/PCA		Model
		.	.	.		R^2
		1062	1030	823	8389	RASE
		662	662	662	662	LSM
		1.6	1.6	1.2	12.7	RASE/LSM
	Unknown (10 NSNs)	SW-MR	SW-MR	SLR		Model
		-0.053	-0.229	-1.417		R^2
		1527	1650	2313	8389	RASE
		3479	3479	3479	2980	LSM
		0.4	0.5	0.7	2.8	RASE/LSM

A variety of the four regression model types have the lowest RASE for each PD at each interval. There does not seem to be much change when the models are used to predict hours for aircraft and common avionics. RASE values are still in the high 7000s to low 8000s. The models performed much worse than the benchmark for the landing gear, with RASE well above 25000. This performance is poor and the models should not be used for forecasting for these items. The rest of the PD categories showed great improvement over the benchmark with much lower RASE and ratios to the zero hours

predictor. Electronic warfare and instruments performed best. This result may not be generalizable due to the low sample size in the test set and lower overall variation of actual MICAP total hours. Once again, we see some categories where the RASE from the models using data further in the past with a RASE lower than the models with more recent data. Additionally, the Propulsion category showed no clear best model type across all time ranges. This may also be due to the limited test set sample size.

Most of the categories see dramatic reductions in RASE for all three intervals. Electronic warfare has the most dramatic improvement with a 97 to 92 percent decrease compared to the overall stepwise multiple regression model results. Instruments and power systems both showed RASE decreases between 88 to 79 percent. The NSNs with unknown PD showed decreases by 80 to 72 percent. Accessories showed smaller but significant decreases between 69 to 66 percent and the other category had decreases of 41 to 26 percent. Common avionics and aircraft did not show consistent increases or decreases. Landing gear showed drastic increases in RASE from 234 to 316 percent. Propulsion showed decreases around 92 percent, however, only one NSN is in this category. Thus, the results are inconclusive for this subcategory.

A majority of the predictions split out by PD types had a lower scaled ratio compared to the 2.8 achieved by the zero MICAP hour prediction benchmark from the overall set. Accessories, common avionics, electronic warfare, instruments, other, power systems, and the unknown set all showed a lower ratio. So, given their different LSMs, the error rates were all comparatively lower than having no overall model (zero hours) or even the best of the overall models using Stepwise Multiple Regression.

Summary

While the overall results show high error rates and results not much better than a zero MICAP hours prediction, the categorical breakout shows that there is great potential for improved predictions when used on specific sets of NSNs like ERRC code or the PD.

VI. Conclusions and Recommendations

Conclusions of Research

The final error rates of the overall set of NSN monthly data from the test showed that the stepwise regression models across all three intervals of prediction had the lowest error rates of the four types of models. However, using the scaled ratio of RASE of the prediction of zero MICAP total hours divided by the least squared mean, a comparison of the best stepwise regression models shows that only the model from the interval predicting MICAP hours three months in the future is more useful than the naive prediction of zero for every NSN in every month.

The connecting letters report from the analysis of variance (ANOVA) on the categorical variables for transformed MICAP hours showed that there are groups of categorical variables with significantly different least squared mean MICAP hours (*see Appendix B for Connecting Letters Report*). This motivated the reevaluation of the twelve best models overall on these subsets of NSNs in the test set. The results show clear indications that models should be used and redeveloped on specific data sets based on the ERRC Code and potentially by PD. The expendable assets show the greatest area for improvement given their clear improvements with a well-defined category. Since the

PD are less well defined further investigation is needed before further application is considered. Models could be potentially built for specific PDs or groups of PDs that have similar variance, however, a larger data set would be needed with more NSNs. Landing gear should likely be excluded or separated from the data set given its extremely high error rates and likely wide variation in MICAP hours.

Results varied across the ERRC Code categories: The expendable assets' best models had error rates 35 to 42 percent lower than the RASE of the best models in the test set overall across the three intervals. The reparable assets' best models had error rates 10 to 12 percent higher than the RASE of the best models overall. Despite the contrast between the two categories, both had scaled ratios slightly higher than the zero MICAP hour prediction scaled ratio. These results, along with the two categories' statistically significant difference shown in the connecting letters report of the ANOVA on their transformed data, are a clear indication that new prediction models should be rebuilt and fit to separate data sets for these categorical variables.

Similar results were found in the categorical breakout by PD. A majority of the PD categories had error rates for their best models in the test set that were lower than the best models overall. The Other's RASE was 26 to 41 percent lower than the best models in the overall test set, the Accessories' were 66 to 69 percent lower, and the Electronic Warfare, Instruments, Power Systems, Unknown, and Propulsion were all 72 to 97 percent lower. The Common Avionics and Aircraft did not see a large change in RASE compared to the best overall models, and the Landing Gear's RASE was 234 to 316 percent higher than the overall best models across the three intervals of prediction.

Additionally, Accessories, Common Avionics, Electronic Warfare, Instruments, Other, and Power Systems all saw scaled ratios of their error rates much lower than the zero MICAP hour prediction. This data set will have to be expanded in future research to more reliably evaluate these findings for Electronic Warfare, Landing Gear, Power Systems, and Propulsion. These categories all have five or fewer NSNs in the test set, which is not likely a large enough sample size. However, the results indicate the clear potential for improved predictions if new models are developed and fit on subsets of NSNs based on some of these categories with similar variation in their monthly MICAP total hours.

Recommendations for Action

While the results of this study were fruitful, more work is needed to improve these prediction models before operational implementation can occur. The results of this study were built on the best performing models that only included MICAP data. Demand data that is more robust than the total monthly due-out quantity is likely needed for these NSNs to improve the forecast accuracy. Then new models should be developed on subsets of NSNs with historically similar variance of their monthly MICAP total hours.

Practical Application

The need for proactive action remains for the cause code H MICAPs for fighter weapon systems. So, the development of prediction models should not be ignored. If future efforts are successful in creating more accurate models, the SCOW can potentially implement a system to monitor the risk of surges in MICAP hours for NSNs based on the forecasts from prediction models. Thresholds of MICAP hour levels determined to be of

concern could be set to create alerts within supply systems to prompt personnel to begin inquiries across the supply chain, increase communication and initiate proactive action in an effort to prevent and reduce the occurrence of these cause code H MICAPs.

Procedure Established in this Study

While the models in this study are likely not ready for current operational implementation, this study was extremely successful in establishing a procedure to build future models using JMP software. First, raw data can be combined and aggregated into specific time intervals using the tabulate and join functions. Second, categorical variables should be run through an Analysis of Variance (ANOVA) to determine subsets of NSNs with similar variance in MICAP total hours to build models from. Third, data should be split into a training, validation, and test set. This is an essential step in the process ensuring that the evaluation of a model's success or usefulness is not based on the in-sample error rate from the original data it was fit on. Measuring the error rate on one or more sets of data independent from the original set used to create the model ensures that models are not overfit, i.e., too specifically tuned for the original data. Fourth, model types should be selected, then multiple models should be developed for each type, then error rates should be measured in the validation data set. Models with the lowest justifiable error rates should be selected for the final comparison against other model types. Fifth, once the best models for all model types have been identified, measure the models' error rates on the test data set. The models with the lowest error rates should be identified as the best models and model type for prediction. Finally, in the comparisons, a tool should measure the usefulness of models given their error rates.

This could be a scaled ratio to a naive prediction. Alternatively, this could be a scaled ratio to a measure currently in operational use.

Future Research

The results of this study point to many areas where future research can expand the work of general MICAP forecasting and the prediction of Cause Code H MICAP hours. First, the goal should be to develop models on subsets of NSNs with similar variance of MICAP hours, instead of developing models that encompass all the NSNs in the set. This can be done by separating models by the categorical variables, by groups of categorical variables identified in the ANOVA, or simply by a distribution of NSNs MICAP hours so that the NSNs with outliers of extremely high MICAP hours in the data set do not distort the results of the models.

Second, future studies should obtain more robust data on the NSNs. The data from this study were closely related to the end-user in the supply chain. Future data should better capture the three main sources that lead to Cause Code H MICAPs; DMSMS, AWPs and Spares, and Long Contract Lead Times. There is also potential for the introduction of other categorical variables related to these sources to create subsets of NSNs with similar variance for the development of future prediction models.

Third, the data set could be expanded beyond the 136 NSNs identified as top drivers. In the extreme, this includes expanding to all NSNs in a given time frame for which Cause Code H MICAPs occurred. It could also include NSNs for other weapon systems besides the F-16. Expanding the data set would also allow for a better sample size for the relatively small test set, and would also allow for a more reliable sample size

of the various PD categories, many of which had five or fewer NSNs in the test set. Further investigation into the origin and classification of PD is also needed if used in future research.

Additionally, models could be redeveloped to predict values other than MICAP hours. Hours were the focus of this study since the accumulation of MICAP hours was one of the defining aspects of the Cause Code H problem. However, models could be created to predict MICAP incidents, the total quantity of the NSNs needed to fill the MICAPs, or some other broad aspect of their impact on readiness, mission capability, or aircraft availability were the data set to be expanded.

Finally, more advanced predictive model types could be used to forecast MICAP hours. In JMP alone there are many useful capabilities besides the four types of linear regression used in this study. This includes the use of Neural Networks, also known as artificial intelligence, for a multivariate analysis framework, that can model non-linear phenomena through a network of inputs, layers of nodes, and outputs (Klimberg & McCullough, 2018). Similarly, Support Vector Regression, a machine learning technique using regression analysis, can be used for forecast development (Drucker et al., 1997). There is also potential for Time Series Regression, where a model relates the predicted *y* value as a function of time or the use of Box-Jenkins models (Bowerman et al., 2005).

Final Words

'All models are wrong, but some are useful.' There are numerous directions with the original data set from this study that future research can take to continue the efforts of MICAP prediction, and proactive action on the Cause Code H MICAPs. This study was

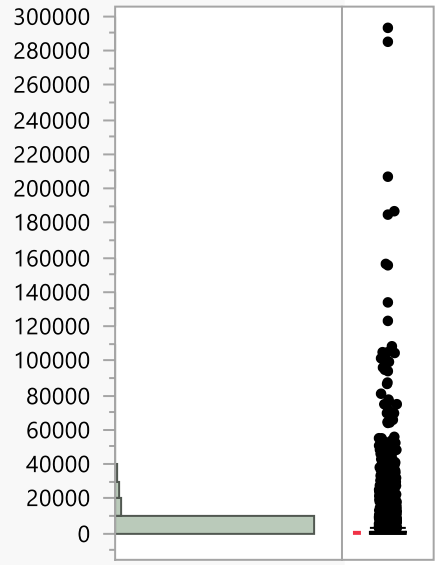
successful in taking some of the first steps to making the future implementation of proactive MICAP prevention and reduction possible. This study lays a potential groundwork for future researchers to take in this field as well as pointing in directions that will likely lead to more fruitful results, all in the effort to improve the readiness, availability, and mission capability of the Air Force's fighter aircraft fleet.

Appendices

Appendix A: Distribution of MICAP Hours in the Data Set

Distributions

MICAP TOTAL HRS



Quantiles

100.0%	maximum	293273
99.5%		74748.8
97.5%		26955.2
90.0%		6204.7
75.0%	quartile	1118.75
50.0%	median	175
25.0%	quartile	0
10.0%		0
2.5%		0
0.5%		0
0.0%	minimum	0

Summary Statistics

Mean	2980.4256
Std Dev	11908.645
Std Err Mean	166.42838
Upper 95% Mean	3306.6964
Lower 95% Mean	2654.1548
N	5120

Appendix B: ANOVA of Categorical Variables

ANOVA using untransformed data

Response MICAP TOTAL HRS

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	10	1.1212e+10	1.1212e+9	8.0143
Error	5109	7.1474e+11	139898874	**Prob > F**
C. Total	5119	7.2596e+11		<.0001 *

Effect Details

ERRC CODE

LSMeans Differences Student's t

α=0.050 t=1.96043

Level		Least Sq Mean
Repairable	A	2922.1185
Expendable	A	1579.4831

Levels not connected by same letter are significantly different.

PD

LSMeans Differences Student's t

α=0.050 t=1.96043

Level					Least Sq Mean
Common Avionics	A				6109.2759
Other	A	B			4329.2309
Unknown	A	B	C	D	3479.0175
Landing Gear		B	C		3132.3879
Aircraft		B	C	D	1771.7766
Accessories			C	D	1166.9790
Instruments				D	943.3352
Propulsion		B	C	D	662.1823
Power Systems				D	553.4073
Electronic Warfare				D	360.4156

Levels not connected by same letter are significantly different.

ANOVA using transformed data

Response Log(MICAP TOTAL HRS + 1)

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	10	4637.803	463.780	39.7774
Error	5109	59567.776	11.659	**Prob > F**
C. Total	5119	64205.579		<.0001 *

Effect Details

ERRC CODE

LSMeans Differences Student's t

α=0.050 t=1.96043

Level			Least Sq Mean
Repairable	A		4.6648839
Expendable		B	1.5872500

Levels not connected by same letter are significantly different.

PD

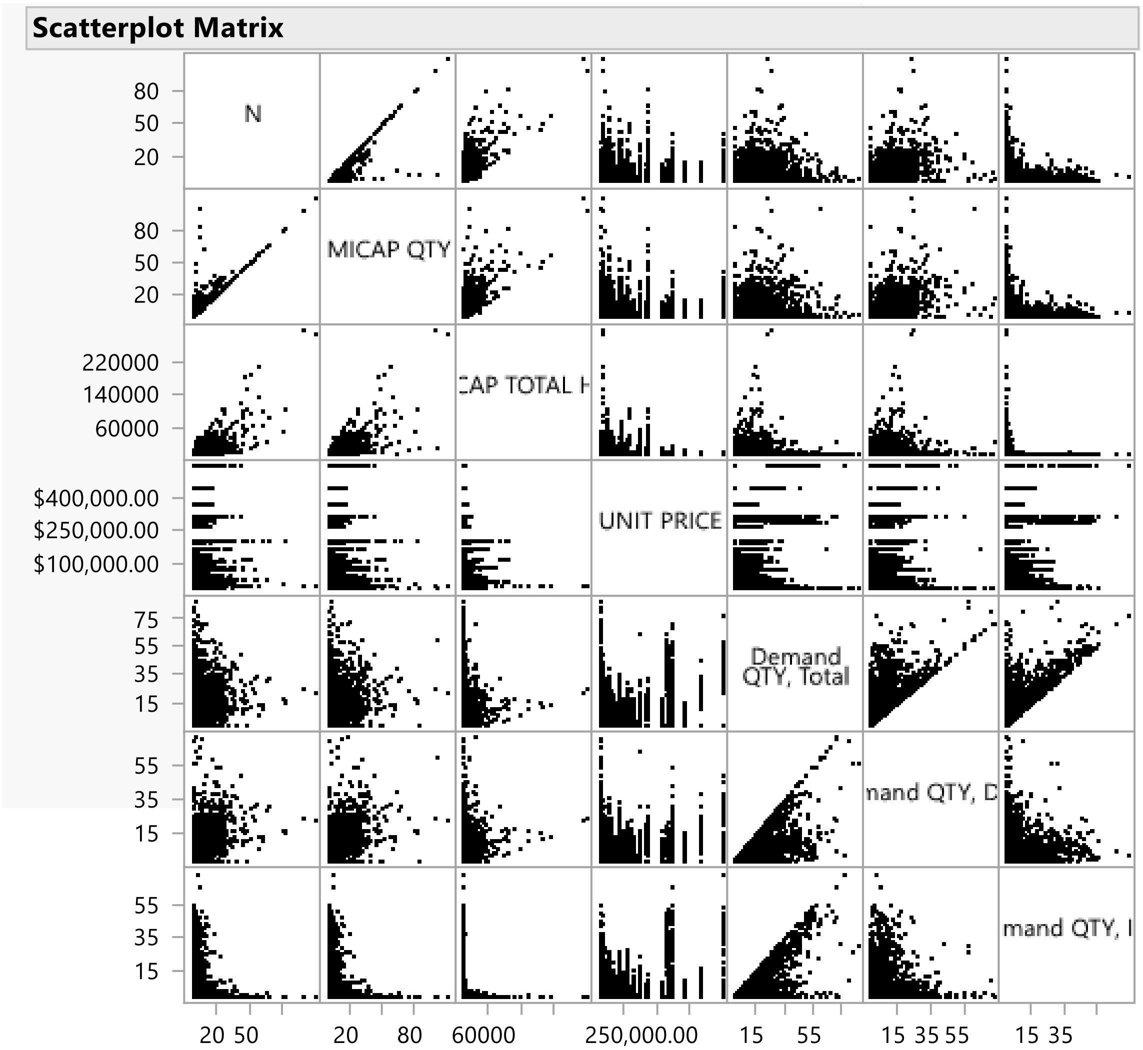
LSMeans Differences Student's t

α=0.050 t=1.96043

Level							Least Sq Mean
Other	A						4.820885
Common Avionics	A	B					3.961441
Electronic Warfare	A	B	C				3.913107
Unknown		B	C	D			3.671249
Power Systems		B	C	D			3.356638
Aircraft			C	D			3.345751
Instruments				D			3.244031
Accessories					E		2.643830
Landing Gear					E		2.503268
Propulsion						F	-0.199530

Levels not connected by same letter are significantly different.

Appendix C: Scatterplot of Continuous Variable Correlations

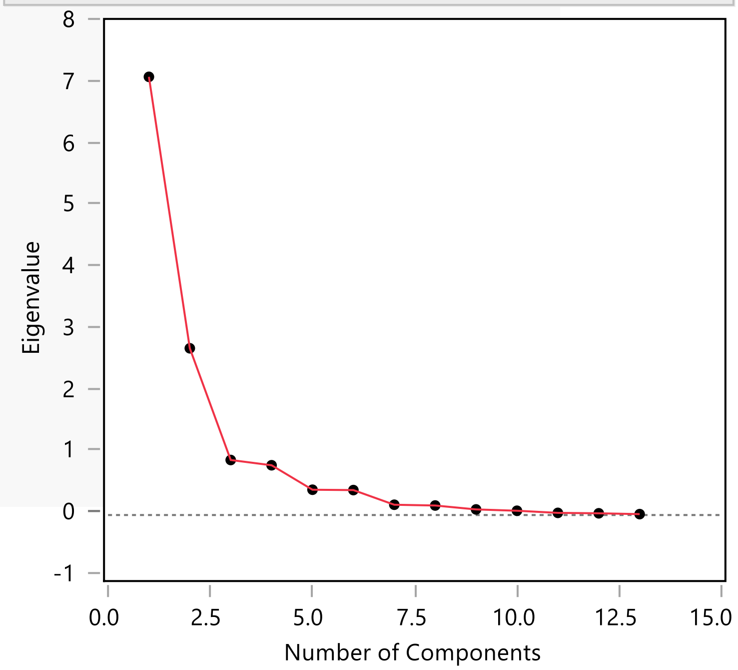


Appendix D: Example List of Eigenvalues and Scree Plot

Eigenvalues

Number	Eigenvalue	Percent	Cum Percent
1	7.1244	54.803	54.803
2	2.7155	20.889	75.692
3	0.9012	6.933	82.624
4	0.8158	6.276	88.900
5	0.4183	3.217	92.117
6	0.4125	3.173	95.291
7	0.1745	1.343	96.633
8	0.1621	1.247	97.880
9	0.0985	0.758	98.638
10	0.0769	0.592	99.230
11	0.0424	0.326	99.556
12	0.0355	0.273	99.829
13	0.0222	0.171	100.000

Scree Plot



Appendix E: Simple Linear Regression Model Comparisons

Models using MICAP Hours

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	1M SLR MICAP HRS PF	0.7980	5690.9	4133
Training	2M SLR MICAP HRS PF	0.5989	8019.8	4133
Training	3M SLR MICAP HRS PF	0.4258	9595.0	4133
Training	4M SLR MICAP HRS PF	0.2902	10668	4133
Training	5M SLR MICAP HRS PF	0.0829	12127	4133
Training	6M SLR MICAP HRS PF	0.0685	12221	4133
Validation	1M SLR MICAP HRS PF	0.6529	4704.8	777
Validation	2M SLR MICAP HRS PF	0.5944	5085.9	777
Validation	3M SLR MICAP HRS PF	0.3994	6188.9	777
Validation	4M SLR MICAP HRS PF	0.1651	7297.1	777
Validation	5M SLR MICAP HRS PF	0.0465	7798.4	777
Validation	6M SLR MICAP HRS PF	-0.031	8110.7	777
Test	1M SLR MICAP HRS PF	0.7765	3798.4	210
Test	2M SLR MICAP HRS PF	0.1940	7213.5	210
Test	3M SLR MICAP HRS PF	0.0184	7960.7	210
Test	4M SLR MICAP HRS PF	-0.109	8459.8	210
Test	5M SLR MICAP HRS PF	-0.007	8061.5	210
Test	6M SLR MICAP HRS PF	0.0047	8016.1	210

Models using MICAP Incidents

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	1M SLR N PF	0.3554	10166	4133
Training	2M SLR N PF	0.2507	10961	4133
Training	3M SLR N PF	0.1809	11460	4133
Training	4M SLR N PF	0.1300	11811	4133
Training	5M SLR N PF	0.0994	12017	4133
Training	6M SLR N PF	0.0809	12140	4133
Validation	1M SLR N PF	-0.053	8193.6	777
Validation	2M SLR N PF	0.1805	7229.7	777
Validation	3M SLR N PF	0.0991	7580.3	777
Validation	4M SLR N PF	0.0206	7903.7	777
Validation	5M SLR N PF	0.0141	7929.7	777
Validation	6M SLR N PF	-0.034	8122.5	777
Test	1M SLR N PF	0.1520	7399.2	210
Test	2M SLR N PF	-0.215	8855.7	210
Test	3M SLR N PF	-0.372	9411.4	210
Test	4M SLR N PF	-0.232	8918.9	210
Test	5M SLR N PF	-0.062	8281.6	210
Test	6M SLR N PF	-0.023	8125.4	210

Models using MICAP Quantity

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	1M SLR MICAP QTY PF	0.3227	10421	4133
Training	2M SLR MICAP QTY PF	0.2160	11212	4133
Training	3M SLR MICAP QTY PF	0.1552	11638	4133
Training	4M SLR MICAP QTY PF	0.1109	11940	4133
Training	5M SLR MICAP QTY PF	0.0829	12127	4133
Training	6M SLR MICAP QTY PF	0.0685	12221	4133
Validation	1M SLR MICAP QTY PF	0.0994	7579.1	777
Validation	2M SLR MICAP QTY PF	0.2699	6824.0	777
Validation	3M SLR MICAP QTY PF	0.1483	7370.5	777
Validation	4M SLR MICAP QTY PF	0.0756	7678.4	777
Validation	5M SLR MICAP QTY PF	0.0465	7798.4	777
Validation	6M SLR MICAP QTY PF	-0.031	8110.7	777
Test	1M SLR MICAP QTY PF	-0.327	9256.9	210
Test	2M SLR MICAP QTY PF	-0.091	8393.5	210
Test	3M SLR MICAP QTY PF	-0.198	8793.9	210
Test	4M SLR MICAP QTY PF	-0.140	8577.7	210
Test	5M SLR MICAP QTY PF	-0.007	8061.5	210
Test	6M SLR MICAP QTY PF	0.0047	8016.1	210

Models using Due-Out Quantity

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	1M SLR DUO QTY PF	0.0332	12451	4133
Training	2M SLR DUO QTY PF	0.0283	12482	4133
Training	3M SLR DUO QTY PF	0.0231	12516	4133
Training	4M SLR DUO QTY PF	0.0183	12546	4133
Training	5M SLR DUO QTY PF	0.0158	12563	4133
Training	6M SLR DUO QTY PF	0.0146	12570	4133
Validation	1M SLR DUO QTY PF	0.0355	7843.1	777
Validation	2M SLR DUO QTY PF	0.0291	7869.0	777
Validation	3M SLR DUO QTY PF	0.0204	7904.5	777
Validation	4M SLR DUO QTY PF	0.0153	7924.8	777
Validation	5M SLR DUO QTY PF	0.0131	7933.8	777
Validation	6M SLR DUO QTY PF	-0.005	8006.8	777
Test	1M SLR DUO QTY PF	-0.019	8111.2	210
Test	2M SLR DUO QTY PF	-0.061	8277.5	210
Test	3M SLR DUO QTY PF	0.0056	8012.4	210
Test	4M SLR DUO QTY PF	-0.024	8131.8	210
Test	5M SLR DUO QTY PF	-0.004	8052.1	210
Test	6M SLR DUO QTY PF	-0.011	8080.7	210

Models using Total Demand Quantity

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	1M SLR Demand QTY Total PF	0.0021	12650	4133
Training	2M SLR Demand QTY Total PF	0.0017	12652	4133
Training	3M SLR Demand QTY Total PF	0.0012	12655	4133
Training	4M SLR Demand QTY Total PF	0.0010	12657	4133
Training	5M SLR Demand QTY Total PF	0.0009	12657	4133
Training	6M SLR Demand QTY Total PF	0.0011	12656	4133
Validation	1M SLR Demand QTY Total PF	-0.003	7999.6	777
Validation	2M SLR Demand QTY Total PF	-0.004	8001.0	777
Validation	3M SLR Demand QTY Total PF	-0.002	7993.3	777
Validation	4M SLR Demand QTY Total PF	-0.003	7999.5	777
Validation	5M SLR Demand QTY Total PF	-0.002	7993.5	777
Validation	6M SLR Demand QTY Total PF	-0.005	8005.0	777
Test	1M SLR Demand QTY Total PF	-0.008	8067.8	210
Test	2M SLR Demand QTY Total PF	-0.018	8106.4	210
Test	3M SLR Demand QTY Total PF	-0.008	8067.3	210
Test	4M SLR Demand QTY Total PF	-0.011	8077.6	210
Test	5M SLR Demand QTY Total PF	-0.009	8072.1	210
Test	6M SLR Demand QTY Total PF	-0.009	8071.7	210

Appendix F: Multiple Regression Model Comparisons

Models using MICAP Hours

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	3-6M MR MICAP HRS PF	0.4411	9466.6	4133
Training	4-6M MR MICAP HRS PF	0.2997	10596	4133
Training	5-6M MR MICAP HRS PF	0.2013	11317	4133
Training	3-6M MR MICAP HRS FtD2 PF	0.4977	8974.4	4133
Training	4-6M MR MICAP HRS FtD2 PF	0.3438	10257	4133
Training	5-6M MR MICAP HRS FtD2 PF	0.2483	10978	4133
Validation	3-6M MR MICAP HRS PF	0.4020	6175.6	777
Validation	4-6M MR MICAP HRS PF	0.1644	7300.5	777
Validation	5-6M MR MICAP HRS PF	0.1159	7509.0	777
Validation	3-6M MR MICAP HRS FtD2 PF	0.3344	6515.5	777
Validation	4-6M MR MICAP HRS FtD2 PF	0.1211	7487.2	777
Validation	5-6M MR MICAP HRS FtD2 PF	0.0486	7789.8	777
Test	3-6M MR MICAP HRS PF	0.0707	7745.8	210
Test	4-6M MR MICAP HRS PF	-0.111	8468.4	210
Test	5-6M MR MICAP HRS PF	-0.061	8275.9	210
Test	3-6M MR MICAP HRS FtD2 PF	0.0338	7898.1	210
Test	4-6M MR MICAP HRS FtD2 PF	-0.297	9150.8	210
Test	5-6M MR MICAP HRS FtD2 PF	-0.270	9054.7	210

Models using MICAP Incidents

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	3-6M MR N PF	0.1840	11439	4133
Training	4-6M MR N PF	0.1341	11783	4133
Training	5-6M MR N PF	0.1026	11996	4133
Training	3-6M MR N FtD2 PF	0.2307	11107	4133
Training	4-6M MR N FtD2 PF	0.1511	11667	4133
Training	5-6M MR N FtD2 PF	0.1035	11990	4133
Validation	3-6M MR N PF	0.0981	7584.4	777
Validation	4-6M MR N PF	0.0194	7908.5	777
Validation	5-6M MR N PF	0.0050	7966.3	777
Validation	3-6M MR N FtD2 PF	0.1102	7533.5	777
Validation	4-6M MR N FtD2 PF	0.0188	7910.8	777
Validation	5-6M MR N FtD2 PF	-0.002	7995.7	777
Test	3-6M MR N PF	-0.342	9309.2	210
Test	4-6M MR N PF	-0.200	8802.4	210
Test	5-6M MR N PF	-0.057	8259.0	210
Test	3-6M MR N FtD2 PF	-0.317	9220.0	210
Test	4-6M MR N FtD2 PF	-0.175	8710.8	210
Test	5-6M MR N FtD2 PF	-0.070	8310.9	210

Models using MICAP Quantity

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	3-6M MR MICAP QTY PF	0.1606	11601	4133
Training	4-6M MR MICAP QTY PF	0.1166	11902	4133
Training	5-6M MR MICAP QTY PF	0.0876	12095	4133
Training	3-6M MR MICAP QTY FtD2 PF	0.1846	11434	4133
Training	4-6M MR MICAP QTY FtD2 PF	0.1278	11826	4133
Training	5-6M MR MICAP QTY FtD2 PF	0.0895	12083	4133
Validation	3-6M MR MICAP QTY PF	0.1442	7388.0	777
Validation	4-6M MR MICAP QTY PF	0.0683	7708.7	777
Validation	5-6M MR MICAP QTY PF	0.0289	7869.9	777
Validation	3-6M MR MICAP QTY FtD2 PF	0.1409	7402.3	777
Validation	4-6M MR MICAP QTY FtD2 PF	0.0606	7740.6	777
Validation	5-6M MR MICAP QTY FtD2 PF	0.0193	7908.8	777
Test	3-6M MR MICAP QTY PF	-0.186	8750.8	210
Test	4-6M MR MICAP QTY PF	-0.111	8469.2	210
Test	5-6M MR MICAP QTY PF	-0.007	8062.2	210
Test	3-6M MR MICAP QTY FtD2 PF	-0.195	8785.4	210
Test	4-6M MR MICAP QTY FtD2 PF	-0.103	8437.2	210
Test	5-6M MR MICAP QTY FtD2 PF	-0.023	8128.7	210

Models using Due-Out Quantity

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	3-6M MR DUO PF	0.0266	12493	4133
Training	4-6M MR DUO PF	0.0218	12524	4133
Training	5-6M MR DUO PF	0.0185	12545	4133
Training	3-6M MR DUO FtD2 PF	0.0321	12458	4133
Training	4-6M MR DUO FtD2 PF	0.0265	12494	4133
Training	5-6M MR DUO FtD2 PF	0.0202	12534	4133
Validation	3-6M MR DUO PF	0.0175	7916.1	777
Validation	4-6M MR DUO PF	0.0122	7937.5	777
Validation	5-6M MR DUO PF	0.0049	7966.8	777
Validation	3-6M MR DUO FtD2 PF	-0.006	8010.3	777
Validation	4-6M MR DUO FtD2 PF	-0.013	8039.8	777
Validation	5-6M MR DUO FtD2 PF	0.0051	7965.7	777
Test	3-6M MR DUO PF	-0.008	8068.7	210
Test	4-6M MR DUO PF	-0.022	8124.6	210
Test	5-6M MR DUO PF	-0.008	8066.8	210
Test	3-6M MR DUO FtD2 PF	0.0111	7990.1	210
Test	4-6M MR DUO FtD2 PF	-0.015	8096.7	210
Test	5-6M MR DUO FtD2 PF	-0.003	8046.1	210

Appendix G: Stepwise Multiple Regression Model Comparisons

Models using Variables from $(t-3), \ldots, (t-6)$

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	SW-MR MICAP 3-6M PF	0.4565	9335.4	4133
Training	SW-MR MICAP & DUO 3-6M PF	0.4630	9279.2	4133
Training	SW-MR MICAP 3-6M (w/o high VIF) PF	0.4499	9391.6	4133
Training	SW-MR MICAP & DUO 3-6M (w/o high VIF) PF	0.4575	9326.5	4133
Validation	SW-MR MICAP 3-6M PF	0.3953	6210.1	777
Validation	SW-MR MICAP & DUO 3-6M PF	0.3970	6201.5	777
Validation	SW-MR MICAP 3-6M (w/o high VIF) PF	0.3896	6239.4	777
Validation	SW-MR MICAP & DUO 3-6M (w/o high VIF) PF	0.3901	6236.9	777
Test	SW-MR MICAP 3-6M PF	0.0952	7643.0	210
Test	SW-MR MICAP & DUO 3-6M PF	0.1087	7585.6	210
Test	SW-MR MICAP 3-6M (w/o high VIF) PF	0.0919	7656.8	210
Test	SW-MR MICAP & DUO 3-6M (w/o high VIF) PF	0.1007	7619.6	210

Models using Variables from $(t-4), \ldots, (t-6)$

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	SW-MR MICAP 4-6M PF	0.3147	10483	4133
Training	SW-MR MICAP & DUO 4-6M PF	0.3184	10454	4133
Training	SW-MR MICAP 4-6M (w/o high VIF) PF	0.3123	10501	4133
Training	SW-MR MICAP & DUO 4-6M (w/o high VIF) PF	0.3164	10470	4133
Validation	SW-MR MICAP 4-6M PF	0.1438	7389.6	777
Validation	SW-MR MICAP & DUO 4-6M PF	0.1512	7357.7	777
Validation	SW-MR MICAP 4-6M (w/o high VIF) PF	0.1532	7349.1	777
Validation	SW-MR MICAP & DUO 4-6M (w/o high VIF) PF	0.1610	7315.2	777
Test	SW-MR MICAP 4-6M PF	-0.083	8361.6	210
Test	SW-MR MICAP & DUO 4-6M PF	-0.071	8316.0	210
Test	SW-MR MICAP 4-6M (w/o high VIF) PF	-0.057	8260.1	210
Test	SW-MR MICAP & DUO 4-6M (w/o high VIF) PF	-0.047	8223.4	210

Models using Variables from $(t-5),(t-6)$

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	SW-MR MICAP 5-6M PF	0.2100	11255	4133
Training	SW-MR MICAP 5-6M FtD2 PF	0.3948	9850.9	4133
Training	SW-MR MICAP & DUO 5-6M PF	0.2130	11233	4133
Training	SW-MR MICAP & DUO 5-6M FtD2 PF	0.4036	9779.4	4133
Validation	SW-MR MICAP 5-6M PF	0.1034	7562.2	777
Validation	SW-MR MICAP 5-6M FtD2 PF	-0.116	8438.2	777
Validation	SW-MR MICAP & DUO 5-6M PF	0.1053	7554.1	777
Validation	SW-MR MICAP & DUO 5-6M FtD2 PF	-0.126	8475.5	777
Test	SW-MR MICAP 5-6M PF	-0.055	8254.9	210
Test	SW-MR MICAP 5-6M FtD2 PF	-0.483	9785.1	210
Test	SW-MR MICAP & DUO 5-6M PF	-0.050	8233.1	210
Test	SW-MR MICAP & DUO 5-6M FtD2 PF	-0.485	9790.5	210

Appendix H: Comparisons of Models using Some to All Components

Models using a Range of Components for MICAP Variables from $(t-3), \ldots, (t-6)$

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	PCA MICAP 3-6M C1-3 PF	0.2982	10608	4133
Training	PCA MICAP 3-6M C1-4 PF	0.3230	10419	4133
Training	PCA MICAP 3-6M C1-5 PF	0.3323	10347	4133
Validation	PCA MICAP 3-6M C1-3 PF	0.2248	7031.6	777
Validation	PCA MICAP 3-6M C1-4 PF	0.2755	6797.7	777
Validation	PCA MICAP 3-6M C1-5 PF	0.2862	6747.1	777
Test	PCA MICAP 3-6M C1-3 PF	-0.210	8839.1	210
Test	PCA MICAP 3-6M C1-4 PF	-0.142	8586.4	210
Test	PCA MICAP 3-6M C1-5 PF	-0.143	8590.6	210

Models using a Range of Components for MICAP and Demand Variables from $(t-3), \ldots, (t-6)$

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	AAE	Freq
Training	PCA MICAP & DUO 3-6M C1-3 PF	0.2581	3847.7	4133
Training	PCA MICAP & DUO 3-6M C1-4 PF	0.3036	3788.1	4133
Training	PCA MICAP & DUO 3-6M C1-5 PF	0.3207	3697.1	4133
Validation	PCA MICAP & DUO 3-6M C1-3 PF	0.1521	3544.6	777
Validation	PCA MICAP & DUO 3-6M C1-4 PF	0.2494	3525.4	777
Validation	PCA MICAP & DUO 3-6M C1-5 PF	0.2628	3551.1	777
Test	PCA MICAP & DUO 3-6M C1-3 PF	-0.124	3526.1	210
Test	PCA MICAP & DUO 3-6M C1-4 PF	-0.207	3943.2	210
Test	PCA MICAP & DUO 3-6M C1-5 PF	-0.171	3775.8	210

Models using a Range of Components for MICAP Variables from $(t-4), \ldots, (t-6)$

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	PCA MICAP 4-6M C1-3 PF	0.2266	11136	4133
Training	PCA MICAP 4-6M C1-4 PF	0.2363	11066	4133
Validation	PCA MICAP 4-6M C1-3 PF	0.1122	7524.9	777
Validation	PCA MICAP 4-6M C1-4 PF	0.1255	7468.4	777
Test	PCA MICAP 4-6M C1-3 PF	-0.154	8631.4	210
Test	PCA MICAP 4-6M C1-4 PF	-0.133	8553.9	210

Models using a Range of Components for MICAP and Demand Variables from $(t-4), \ldots, (t-6)$

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	PCA MICAP & DUO 4-6M C1-3 PF	0.2033	11302	4133
Training	PCA MICAP & DUO 4-6M C1-4 PF	0.2233	11160	4133
Training	PCA MICAP & DUO 4-6M C1-5 PF	0.2347	11077	4133
Validation	PCA MICAP & DUO 4-6M C1-3 PF	0.0706	7699.1	777
Validation	PCA MICAP & DUO 4-6M C1-4 PF	0.1124	7524.2	777
Validation	PCA MICAP & DUO 4-6M C1-5 PF	0.1151	7512.7	777
Test	PCA MICAP & DUO 4-6M C1-3 PF	-0.110	8465.5	210
Test	PCA MICAP & DUO 4-6M C1-4 PF	-0.157	8641.9	210
Test	PCA MICAP & DUO 4-6M C1-5 PF	-0.131	8544.2	210

Models using a Range of Components for MICAP Variables from $(t-5), (t-6)$

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	PCA MICAP 5-6M C1-3 PF	0.1740	11509	4133
Training	PCA MICAP 5-6M C1-4 PF	0.1751	11501	4133
Validation	PCA MICAP 5-6M C1-3 PF	0.0571	7754.8	777
Validation	PCA MICAP 5-6M C1-4 PF	0.0654	7720.6	777
Test	PCA MICAP 5-6M C1-3 PF	-0.071	8315.1	210
Test	PCA MICAP 5-6M C1-4 PF	-0.069	8305.7	210

Models using a Range of Components for MICAP and Demand Variables from $(t-5), (t-6)$

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	PCA MICAP & DUO 5-6M C1-3 PF	0.1591	11612	4133
Training	PCA MICAP & DUO 5-6M C1-4 PF	0.1660	11564	4133
Validation	PCA MICAP & DUO 5-6M C1-3 PF	0.0290	7869.5	777
Validation	PCA MICAP & DUO 5-6M C1-4 PF	0.0584	7749.5	777
Test	PCA MICAP & DUO 5-6M C1-3 PF	-0.069	8306.2	210
Test	PCA MICAP & DUO 5-6M C1-4 PF	-0.067	8301.5	210

Appendix I: Principal Component Analysis Regression Model Comparisons

Models using Variables from $(t-3), \dots, (t-6)$

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	PCA MICAP & DUO 3-6M C1-5 PF	0.3207	10437	4133
Training	PCA MICAP 3-6M C1-5 PF	0.3323	10347	4133
Validation	PCA MICAP & DUO 3-6M C1-5 PF	0.2628	6856.8	777
Validation	PCA MICAP 3-6M C1-5 PF	0.2862	6747.1	777
Test	PCA MICAP & DUO 3-6M C1-5 PF	-0.171	8693.5	210
Test	PCA MICAP 3-6M C1-5 PF	-0.143	8590.6	210

Models using Variables from $(t-4), \dots, (t-6)$

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	PCA MICAP & DUO 4-6M C1-5 PF	0.2347	11077	4133
Training	PCA MICAP 4-6M C1-4 PF	0.2363	11066	4133
Validation	PCA MICAP & DUO 4-6M C1-5 PF	0.1151	7512.7	777
Validation	PCA MICAP 4-6M C1-4 PF	0.1255	7468.4	777
Test	PCA MICAP & DUO 4-6M C1-5 PF	-0.131	8544.2	210
Test	PCA MICAP 4-6M C1-4 PF	-0.133	8553.9	210

Models using Variables from $(t-5), (t-6)$

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	PCA MICAP & DUO 5-6M C1-4 PF	0.1660	11564	4133
Training	PCA MICAP 5-6M C1-4 PF	0.1751	11501	4133
Validation	PCA MICAP & DUO 5-6M C1-4 PF	0.0584	7749.5	777
Validation	PCA MICAP 5-6M C1-4 PF	0.0654	7720.6	777
Test	PCA MICAP & DUO 5-6M C1-4 PF	-0.067	8301.5	210
Test	PCA MICAP 5-6M C1-4 PF	-0.069	8305.7	210

Appendix J: Model Outputs of Twelve Overall Best Models

Best Simple Linear Regression Model for $(t-3)$

Response MICAP TOTAL HRS

Validation: Validation

Effect Summary

Source	LogWorth		PValue
MICAP TOTAL HRS - 3M	499.440		0.00000

Summary of Fit

RSquare	0.425841
RSquare Adj	0.425702
Root Mean Square Error	9597.272
Mean of Response	3055.364
Observations (or Sum Wgts)	4133

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	1110.9382	153.362	7.24	<.0001 *
MICAP TOTAL HRS - 3M	0.6800437	0.012286	55.35	<.0001 *

Best Simple Linear Regression Model for $(t-4)$

Response MICAP TOTAL HRS

Validation: Validation

Effect Summary

Source	LogWorth		PValue
MICAP TOTAL HRS - 4M	309.101		0.00000

Summary of Fit

RSquare	0.290188
RSquare Adj	0.290016
Root Mean Square Error	10670.96
Mean of Response	3055.364
Observations (or Sum Wgts)	4133

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	1440.2563	170.575	8.44	<.0001 *
MICAP TOTAL HRS - 4M	0.5613474	0.01366	41.10	<.0001 *

Best Simple Linear Regression Model for $(t-5)$

Response MICAP TOTAL HRS

Validation: Validation

Effect Summary

Source	LogWorth	PValue
MICAP TOTAL HRS - 5M	196.850	0.00000

Summary of Fit

RSquare	0.195645
RSquare Adj	0.195451
Root Mean Square Error	11359.4
Mean of Response	3055.364
Observations (or Sum Wgts)	4133

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	1732.7237	181.5543	9.54	<.0001 *
MICAP TOTAL HRS - 5M	0.4644294	0.014651	31.70	<.0001 *

Best Multiple Regression Model for $(t-3), \ldots, (t-6)$

Response MICAP TOTAL HRS

Validation: Validation

Effect Summary

Source	LogWorth	PValue
MICAP TOTAL HRS - 3M	203.687	0.00000
MICAP TOTAL HRS - 4M	6.346	0.00000
MICAP TOTAL HRS - 6M	0.642	0.22784
MICAP TOTAL HRS - 5M	0.562	0.27433

Summary of Fit

RSquare	0.441101
RSquare Adj	0.44056
Root Mean Square Error	9472.314
Mean of Response	3055.364
Observations (or Sum Wgts)	4133

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	1234.4952	152.2372	8.11	<.0001 *
MICAP TOTAL HRS - 6M	-0.034676	0.028749	-1.21	0.2278
MICAP TOTAL HRS - 5M	-0.045115	0.041265	-1.09	0.2743
MICAP TOTAL HRS - 4M	-0.207256	0.041005	-5.05	<.0001 *
MICAP TOTAL HRS - 3M	0.9250429	0.028627	32.31	<.0001 *

Best Multiple Regression Model for $(t-4), \ldots, (t-6)$

Response MICAP TOTAL HRS

Validation: Validation

Effect Summary

Source	LogWorth	PValue
MICAP TOTAL HRS - 4M	119.412	0.00000
MICAP TOTAL HRS - 5M	3.978	0.00011
MICAP TOTAL HRS - 6M	0.883	0.13091

Summary of Fit

RSquare	0.299733
RSquare Adj	0.299224
Root Mean Square Error	10601.54
Mean of Response	3055.364
Observations (or Sum Wgts)	4133

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	1522.3636	170.0939	8.95	<.0001 *
MICAP TOTAL HRS - 6M	-0.048608	0.032173	-1.51	0.1309
MICAP TOTAL HRS - 5M	-0.178395	0.045953	-3.88	0.0001 *
MICAP TOTAL HRS - 4M	0.757756	0.031448	24.10	<.0001 *

Best Multiple Regression Model for $(t-5), (t-6)$

Response MICAP TOTAL HRS

Validation: Validation

Effect Summary

Source	LogWorth	PValue
MICAP TOTAL HRS - 5M	73.980	0.00000
MICAP TOTAL HRS - 6M	7.133	0.00000

Summary of Fit

RSquare	0.201268
RSquare Adj	0.200882
Root Mean Square Error	11321
Mean of Response	3055.364
Observations (or Sum Wgts)	4133

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	1789.8169	181.25	9.87	<.0001 *
MICAP TOTAL HRS - 6M	-0.182469	0.03384	-5.39	<.0001 *
MICAP TOTAL HRS - 5M	0.6278178	0.033636	18.66	<.0001 *

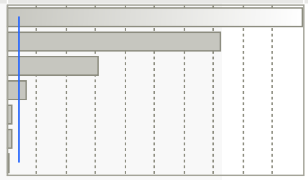
Best Stepwise Multiple Regression Model for $(t-3), \ldots, (t-6)$

Response MICAP TOTAL HRS

Validation: Validation

Effect Summary

Source	LogWorth	PValue
MICAP TOTAL HRS - 3M	310.024	0.00000
MICAP TOTAL HRS - 6M	36.187	0.00000
N - 6M	15.501	0.00000
N - 3M	3.411	0.00039
N - 5M	0.959	0.10986
MICAP QTY - 4M	0.849	0.14157
MICAP QTY - 3M	0.413	0.38666



Summary of Fit

RSquare	0.449922
RSquare Adj	0.448988
Root Mean Square Error	9400.687
Mean of Response	3055.364
Observations (or Sum Wgts)	4133

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	960.12588	185.7146	5.17	<.0001 *
N - 6M	345.58431	42.13947	8.20	<.0001 *
N - 5M	69.389923	43.3916	1.60	0.1099
N - 3M	-213.0311	59.99529	-3.55	0.0004 *
MICAP QTY - 4M	-49.98978	34.00063	-1.47	0.1416
MICAP QTY - 3M	-43.6603	50.42905	-0.87	0.3867
MICAP TOTAL HRS - 6M	-0.283062	0.022083	-12.82	<.0001 *
MICAP TOTAL HRS - 3M	0.8833	0.021453	41.17	<.0001 *

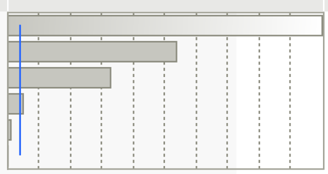
Best Stepwise Multiple Regression Model for $(t-4), \ldots, (t-6)$

Response MICAP TOTAL HRS

Validation: Validation

Effect Summary

Source	LogWorth	PValue
MICAP TOTAL HRS - 4M	163.673	0.00000
MICAP TOTAL HRS - 6M	26.999	0.00000
N - 6M	16.552	0.00000
N - 4M	2.459	0.00348
MICAP QTY - 4M	0.703	0.19807
N - 5M	0.048	0.89529



Summary of Fit

RSquare	0.312336
RSquare Adj	0.311336
Root Mean Square Error	10509.52
Mean of Response	3055.364
Observations (or Sum Wgts)	4133

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	1176.4669	204.6001	5.75	<.0001 *
N - 6M	413.70578	48.7198	8.49	<.0001 *
N - 5M	-6.663093	50.62101	-0.13	0.8953
N - 4M	-208.802	71.41575	-2.92	0.0035 *
MICAP QTY - 4M	-66.31803	51.51786	-1.29	0.1981
MICAP TOTAL HRS - 6M	-0.316929	0.028831	-10.99	<.0001 *
MICAP TOTAL HRS - 4M	0.8032896	0.028076	28.61	<.0001 *

Best Stepwise Multiple Regression Model for $(t-5),(t-6)$

Fit Group

Response MICAP TOTAL HRS

Validation: Validation

Effect Summary

Source	LogWorth	PValue
MICAP TOTAL HRS - 5M	69.692	0.00000
MICAP TOTAL HRS - 6M	15.588	0.00000
N - 6M	9.826	0.00000
N - 5M	0.962	0.10909
MICAP QTY - 5M	0.843	0.14342

Summary of Fit

RSquare	0.210025
RSquare Adj	0.209068
Root Mean Square Error	11262.86
Mean of Response	3055.364
Observations (or Sum Wgts)	4133

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	1389.1069	215.8243	6.44	<.0001 *
N - 6M	348.39007	54.24475	6.42	<.0001 *
N - 5M	-122.9693	76.7295	-1.60	0.1091
MICAP QTY - 5M	-80.49371	55.00285	-1.46	0.1434
MICAP TOTAL HRS - 6M	-0.3419	0.041567	-8.23	<.0001 *
MICAP TOTAL HRS - 5M	0.7384696	0.040832	18.09	<.0001 *

Best Principal Component Analysis Model for $(t-3),\ldots,(t-6)$

Response MICAP TOTAL HRS

Validation: Validation

Effect Summary

Source	LogWorth	PValue
Prin1 4	202.511	0.00000
Prin2 4	83.139	0.00000
Prin3 4	57.823	0.00000
Prin4 4	33.175	0.00000
Prin5 4	13.329	0.00000

Summary of Fit

RSquare	0.332265
RSquare Adj	0.331456
Root Mean Square Error	10354.86
Mean of Response	3055.364
Observations (or Sum Wgts)	4133

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	3048.1176	161.0946	18.92	<.0001 *
Prin1 4	1677.181	52.07249	32.21	<.0001 *
Prin2 4	-1745.011	87.88585	-19.86	<.0001 *
Prin3 4	2275.3189	138.811	16.39	<.0001 *
Prin4 4	1937.2073	158.1741	12.25	<.0001 *
Prin5 4	1582.466	209.1333	7.57	<.0001 *

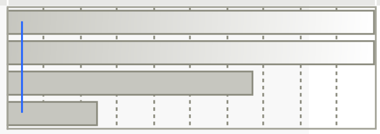
Best Principal Component Analysis Model for $(t-4), \ldots, (t-6)$

Response MICAP TOTAL HRS

Validation: Validation

Effect Summary

Source	LogWorth	PValue
Prin1 5	143.512	0.00000
Prin2 5	58.419	0.00000
Prin3 5	33.619	0.00000
Prin4 5	12.311	0.00000



Summary of Fit

RSquare	0.236326
RSquare Adj	0.235586
Root Mean Square Error	11072.45
Mean of Response	3055.364
Observations (or Sum Wgts)	4133

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	3011.187	172.242	17.48	<.0001 *
Prin1 5	1654.5765	62.1522	26.62	<.0001 *
Prin2 5	-1709.278	103.7182	-16.48	<.0001 *
Prin3 5	2193.5499	177.8587	12.33	<.0001 *
Prin4 5	1348.262	185.9174	7.25	<.0001 *

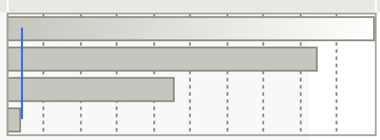
Best Principal Component Analysis Model for $(t-5), (t-6)$

Response MICAP TOTAL HRS

Validation: Validation

Effect Summary

Source	LogWorth	PValue
Prin1 6	106.925	0.00000
Prin2 6	42.370	0.00000
Prin3 6	22.788	0.00000
Prin4 6	1.730	0.01861



Summary of Fit

RSquare	0.175058
RSquare Adj	0.174259
Root Mean Square Error	11508.04
Mean of Response	3055.364
Observations (or Sum Wgts)	4133

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	3018.696	179.0161	16.86	<.0001 *
Prin1 6	1735.4151	76.42954	22.71	<.0001 *
Prin2 6	-1730.925	124.3206	-13.92	<.0001 *
Prin3 6	2164.8959	215.3129	10.05	<.0001 *
Prin4 6	-569.493	241.8973	-2.35	0.0186 *

Appendix K: Overall Best Model Comparisons

Best Models from $(t-3), \ldots, (t-6)$

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	SLR MICAP HRS 3M PF	0.4258	9595.0	4133
Training	MR MICAP HRS 3-6M PF	0.4411	9466.6	4133
Training	SW-MR MICAP 3-6M (w/o high VIF) PF	0.4499	9391.6	4133
Training	PCA MICAP 3-6M C1-5 PF	0.3323	10347	4133
Validation	SLR MICAP HRS 3M PF	0.3994	6188.9	777
Validation	MR MICAP HRS 3-6M PF	0.4020	6175.6	777
Validation	SW-MR MICAP 3-6M (w/o high VIF) PF	0.3896	6239.4	777
Validation	PCA MICAP 3-6M C1-5 PF	0.2862	6747.1	777
Test	SLR MICAP HRS 3M PF	0.0184	7960.7	210
Test	MR MICAP HRS 3-6M PF	0.0707	7745.8	210
Test	SW-MR MICAP 3-6M (w/o high VIF) PF	0.0919	7656.8	210
Test	PCA MICAP 3-6M C1-5 PF	-0.143	8590.6	210

Best Models from $(t-4), \ldots, (t-6)$

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	SLR MICAP HRS 4M PF	0.2902	10668	4133
Training	MR MICAP HRS 4-6M PF	0.2997	10596	4133
Training	SW-MR MICAP 4-6M (w/o high VIF) PF	0.3123	10501	4133
Training	PCA MICAP 4-6M C1-4 PF	0.2363	11066	4133
Validation	SLR MICAP HRS 4M PF	0.1651	7297.1	777
Validation	MR MICAP HRS 4-6M PF	0.1644	7300.5	777
Validation	SW-MR MICAP 4-6M (w/o high VIF) PF	0.1532	7349.1	777
Validation	PCA MICAP 4-6M C1-4 PF	0.1255	7468.4	777
Test	SLR MICAP HRS 4M PF	-0.109	8459.8	210
Test	MR MICAP HRS 4-6M PF	-0.111	8468.4	210
Test	SW-MR MICAP 4-6M (w/o high VIF) PF	-0.057	8260.1	210
Test	PCA MICAP 4-6M C1-4 PF	-0.133	8553.9	210

Best Models from $(t-5), (t-6)$

Model Comparison

Measures of Fit for MICAP TOTAL HRS

Validation	Predictor	RSquare	RASE	Freq
Training	SLR MICAP HRS 5M PF	0.1956	11357	4133
Training	MR MICAP HRS 5-6M PF	0.2013	11317	4133
Training	SW-MR MICAP 5-6M PF	0.2100	11255	4133
Training	PCA MICAP 5-6M C1-4 PF	0.1751	11501	4133
Validation	SLR MICAP HRS 5M PF	0.0985	7582.8	777
Validation	MR MICAP HRS 5-6M PF	0.1159	7509.0	777
Validation	SW-MR MICAP 5-6M PF	0.1034	7562.2	777
Validation	PCA MICAP 5-6M C1-4 PF	0.0654	7720.6	777
Test	SLR MICAP HRS 5M PF	-0.080	8348.7	210
Test	MR MICAP HRS 5-6M PF	-0.061	8275.9	210
Test	SW-MR MICAP 5-6M PF	-0.055	8254.9	210
Test	PCA MICAP 5-6M C1-4 PF	-0.069	8305.7	210

Bibliography

AF/A4LR. (2017). *Air Force Handbook 23-123 Materiel Management Handbook Volume One, Materiel Management Reference Information*. www.e-publishing.af.mi

AF/A4LR. (2020a). *Air Force Instruction 23-101 Materiel Management Policy Compliance*. www.e-publishing.af.mi

AF/A4LR. (2020b). *Air Force Manual 23-122 Materiel Management Procedures*. www.e-Publishing.af.mil

Blazer, D. J. (1984). *Demand Forecasting*. https://discover.dtic.mil/

Bowerman, B. L., O'Connell, R. T., & Koehler, A. B. (2005). *Forecasting, Time Series, and Regression: An Applied Approach* (Fourth Edi). Thomson Brooks/Cole.

Carter, M. W., & London, R. (2000). Why So Many AWP LRUs? *Air Force Jornal of Logistics*, *26*(4), 31–33.

Chen, F., Drezner, Z., Ryan, J. K., & Simchi-Levi, D. (2000). Quantifying the Bullwhip Effect in a Simple Supply Chain: The Impact of Forecasting, Lead Times, and Information. *Management Science*, *46*(3), 436–443.

De Treville, S., Shapiro, R. D., & Hameri, A. P. (2004). From supply chain to demand chain: The role of lead time reduction in improving demand chain performance. *Journal of Operations Management*, *21*(6), 613–627. https://doi.org/10.1016/j.jom.2003.10.001

Dolgui, A., & Ould-Louly, M. A. (2002). A model for supply planning under lead time uncertainty. *International Journal of Production Economics*, *78*(2), 145–152. https://doi.org/10.1016/S0925-5273(00)00180-8

Drucker, H., Surges, C. J. C., Kaufman, L., Smola, A., & Vapnik, V. (1997). Support vector regression machines. *Advances in Neural Information Processing Systems*, *1*, 155–161.

Eitan, Y. (1990). *“Awaiting-Parts Time” in a Multi-Indenture, Readiness-Based Sparing Model*. https://discover.dtic.mil/

Greer, J. A., & Moon, I. (1981). *Analysis and Use of Air Force Base Level Supply Management Indicators* [Air Force Institute of Technology]. https://discover.dtic.mil/

Hua, Z., & Zhang, B. (2006). A hybrid support vector machines and logistic regression approach for forecasting intermittent demand of spare parts. *Applied Mathematics and Computation*, *181*(2), 1035–1048. https://doi.org/10.1016/j.amc.2006.01.064

Huber, L. E. (1985). *A Dyna-Metric Analysis of Base Awaiting Parts (AWP) Sensitivity to Depot Repair Cycle Variables* [Air Force Institute of Technology]. https://discover.dtic.mil/

Ingram, M. D. (2020). *Explaining Weapon System Sustainment’s Impact to Aircraft Availability* [Air Force Institute of Technology]. https://discover.dtic.mil/

Inman, O. L., Anderson, T. R., & Harmon, R. R. (2006). Predicting U.S. jet fighter aircraft introductions from 1944 to 1982: A dogfight between regression and TFDEA. *Technological Forecasting and Social Change*, *73*(9), 1178–1187. https://doi.org/10.1016/j.techfore.2006.05.013

Johnstone, D., & Chandler, E. (2019). *Multi-Year F-16 MICAP Analysis Customer : Wyle*. https://discover.dtic.mil/

Kaytez, F., Taplamacioglu, M. C., Cam, E., & Hardalac, F. (2015). Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines. *International Journal of Electrical Power and Energy Systems*, *67*, 431–438. https://doi.org/10.1016/j.ijepes.2014.12.036

Klimberg, R., & McCullough, B. D. (2018). *Fundamentals of Predictive Analytics with JMP* (Second Edi). SAS Institute Inc.

Kobren, B., Melnikow, A., & Robinson, D. G. (2005). Mitigating Diminishing Manufacturing Sources and Material Shortages. *Defense Transportation Journal*, *June*, 62–66.

Leng, M., & Parlar, M. (2009). Lead-time reduction in a two-level supply chain: Non-cooperative equilibria vs. coordination with a profit-sharing contract. *International Journal of Production Economics*, *118*(2), 521–544. https://doi.org/10.1016/j.ijpe.2009.01.002

Levy, R. A., & Jondrow, J. M. (1990). *The Effect of Readiness-Based Sparing on WeaponSystme Availability*. https://discover.dtic.mil/

Livingston, H. (2000). GEB1: Diminishing Manufacturing Sources and Material Shortages (DMSMS) management practices. In *Proceedings of the 2000 DMSMS Conference*.

Maddux, G. A. (1999). *Investigate DMSMS Impact on Selected Spare Parts*. https://discover.dtic.mil/

Manship, W., & Moore, R. (2000). *Reducing Base-level Excess AWP Items*. https://discover.dtic.mil/

Miller, K. A. (2012). *Diminishing Manufacturing Sources and Material Shortages Mitigation Strategies: A Multiple-Case Study* [Air Force Institute of Technology]. https://discover.dtic.mil/

Muno, K. N., & Pezoulas, P. K. (1997). *Identification and Reduction of Bottlenecks Concerning MICAP Re-Supply of F-16 Weapons System Avionics Line Replaceable Units*. Air Force Institute of Technology.

Peltz, E., Cox, A. G., Chan, E. W., Hart, G. E., Sommerhauser, D., Hawkins, C., & Connor, K. (2015). *Improving DLA Supply Chain Agility*. papers3://publication/uuid/DAF3A5AB-2651-41BA-B81E-5A48E4EC88F3

Sandborn, P., Prabhakar, V., & Ahmad, O. (2011). Forecasting electronic part procurement lifetimes to enable the management of DMSMS obsolescence. *Microelectronics Reliability*, *51*(2), 392–399. https://doi.org/10.1016/j.microrel.2010.08.005

Slay, F. M., Burleson, R. E., & Meyenburg, J. D. (2000). Computing Wartime Spare Parts for Strategic Airlift. *Air Force Jornal of Logistics*, *24*(1), 31–34.

Smith, B. L., Williams, B. M., & Keith Oswald, R. (2002). Comparison of parametric and nonparametric models for traffic flow forecasting. *Transportation Research Part C: Emerging Technologies*, *10*(4), 303–321. https://doi.org/10.1016/S0968-090X(02)00009-8

Song, J. S., Yano, C. A., & Lerssrisuriya, P. (2000). Contract Assembly: Dealing with Combined Supply Lead Time and Demand Quantity Uncertainty. *Manufacturing and Service Operations Management*, *2*(3), 287–296.

https://doi.org/10.1287/msom.2.3.287.12346

The Model Comparison Report. (2020). SAS Institute Inc. SAS Institute Inc.SAS Institute Inc.

Variable selection in multiple regression. (n.d.). SAS Institute Inc. Retrieved January 29, 2021, from https://www.jmp.com/en_us/statistics-knowledge-portal/what-is-multiple-regression/variable-selection.html

Yip, H. L., Fan, H., & Chiang, Y. H. (2014). Predicting the maintenance cost of construction equipment: Comparison between general regression neural network and Box-Jenkins time series models. *Automation in Construction*, *38*, 30–38. https://doi.org/10.1016/j.autcon.2013.10.024

Zolghadri, M., Addouche, S. A., Boissie, K., & Richard, D. (2018). Obsolescence prediction: A Bayesian model. *Procedia CIRP*, *70*, 392–397. https://doi.org/10.1016/j.procir.2018.02.037

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
25-03-2021	Master's Thesis	Oct 2019 - March 2021

4. TITLE AND SUBTITLE
Predicting F-16 Cause Code H MICAP Hours Using JMP Regression Analysis

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Carr, Scott E, Capt

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/ENs)
2950 Hobson Way
Wright-Patterson AFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT-ENS-MS-21-M-146

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
635 Supply Chain Operations Wing
Capt Kalyn M Howard
304 W. Martin St Bldg 1665
Scott AFB, IL 62225
635SCOW.LG@us.af.mil

10. SPONSOR/MONITOR'S ACRONYM(S)
635 SCOW/LG

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution Statement A. Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Emergency demands for aircraft parts, MICAPs (Mission Impaired Capability Awaiting Parts) are one of the leading issues affecting mission capability supply rates for fighter aircraft. Cause code H MICAPS, those with a known demand level, but no supply available, are particularly troublesome and difficult to resolve. These MICAPS are due to failures by the supply chain to replenish stock levels within acceptable time limits. Numerous studies have identified a clear need for proactive measures to reduce MICAP hours. This would significantly improve aircraft availability. This study uses regression analysis implemented in JMP software to build models with a goal of predicting future MICAP hours for F-16 cause code H

15. SUBJECT TERMS
MICAP Hours, Cause Code H, Forecasting, JMP, Regression

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Dr. Frank W Ciarallo, AFIT/ENS
					19b. TELEPHONE NUMBER *(Include area code)*
U	U	U	UU	101	937-225-6565 x4702 frank.ciarallo@afit.edu

INSTRUCTIONS FOR COMPLETING SF 298

1. REPORT DATE. Full publication date, including day, month, if available. Must cite at least the year and be Year 2000 compliant, e.g. 30-06-1998; xx-06-1998; xx-xx-1998.

2. REPORT TYPE. State the type of report, such as final, technical, interim, memorandum, master's thesis, progress, quarterly, research, special, group study, etc.

3. DATE COVERED. Indicate the time during which the work was performed and the report was written, e.g., Jun 1997 - Jun 1998; 1-10 Jun 1996; May - Nov 1998; Nov 1998.

4. TITLE. Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

5a. CONTRACT NUMBER. Enter all contract numbers as they appear in the report, e.g. F33315-86-C-5169.

5b. GRANT NUMBER. Enter all grant numbers as they appear in the report. e.g. AFOSR-82-1234.

5c. PROGRAM ELEMENT NUMBER. Enter all program element numbers as they appear in the report, e.g. 61101A.

5e. TASK NUMBER. Enter all task numbers as they appear in the report, e.g. 05; RF0330201; T4112.

5f. WORK UNIT NUMBER. Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

6. AUTHOR(S). Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, J, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES). Self-explanatory.

8. PERFORMING ORGANIZATION REPORT NUMBER. Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES). Enter the name and address of the organization(s) financially responsible for and monitoring the work.

10. SPONSOR/MONITOR'S ACRONYM(S). Enter, if available, e.g. BRL, ARDEC, NADC.

11. SPONSOR/MONITOR'S REPORT NUMBER(S). Enter report number as assigned by the sponsoring/ monitoring agency, if available, e.g. BRL-TR-829; -215.

12. DISTRIBUTION/AVAILABILITY STATEMENT. Use agency-mandated availability statements to indicate the public availability or distribution limitations of the report. If additional limitations/ restrictions or special markings are indicated, follow agency authorization procedures, e.g. RD/FRD, PROPIN, ITAR, etc. Include copyright information.

13. SUPPLEMENTARY NOTES. Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

14. ABSTRACT. A brief (approximately 200 words) factual summary of the most significant information.

15. SUBJECT TERMS. Key words or phrases identifying major concepts in the report.

16. SECURITY CLASSIFICATION. Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

17. LIMITATION OF ABSTRACT. This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.