3-26-2020

MC80: Quantifying Effect of Fleet Health on Sortie Execution in F-16 Fleet

Kyle R. Gladney

Follow this and additional works at: https://scholar.afit.edu/etd

Part of the Management and Operations Commons, and the Operations Research, Systems Engineering and Industrial Engineering Commons

Recommended Citation

This Thesis is brought to you for free and open access by the Student Graduate Works at AFIT Scholar. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of AFIT Scholar. For more information, please contact richard.mansfield@afit.edu.
MC80: QUANTIFYING THE EFFECT OF FLEET HEALTH ON SORTIE EXECUTION IN THE F-16 FLEET

THESIS

Kyle R. Gladney, Captain, USAF

AFIT-ENS-MS-20-M-151

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

DISTRIBUTION STATEMENT A.
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.
The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government. This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.
MC80: QUANTIFYING THE EFFECT OF FLEET HEALTH ON SORTIE EXECUTION IN THE F-16 FLEET

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 & Supply Chain Management

Kyle R. Gladney, BS
Captain, USAF

March 2020

DISTRIBUTION STATEMENT A.
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.
MC80: QUANTIFYING THE EFFECT OF FLEET HEALTH ON SORTIE EXECUTION IN THE F-16 FLEET

Kyle R. Gladney, BS
Captain, USAF

Committee Membership:

Lt Col John M. Dickens, Ph.D.
Chair

Lt Col Jason R. Anderson, Ph.D.
Member
AFIT-ENS-MS-20-M-151

Abstract

By order of the Secretary of Defense, all of the US Air Force's F-16 units were tasked to improve their fleet health to a Mission Capability (MC) rate of 80 percent, as part of a Department of Defense-wide push to make its Critical Aviation Platforms, and the units that employ them, more ready and lethal. This study uses historical fleet health and sortie execution data captured from LIMS-EV (Weapon System Viewer, 2020) to create a multiple regression model that quantifies the value of increased fleet health, defined as either MC rate or Aircraft Availability (AA) rate, in terms of increasing sortie output. It also uses forecasted near-future sortie demand to assess the utility of the 80 percent MC rate standard towards achieving desired sortie execution levels. This research concludes that both AA rate and MC rate correlate with increased aircraft utilization and that an increase in either fleet health metric correlates to increased annual utilization of roughly five sorties per aircraft. It also identifies AA rate as a more significant input to sortie execution than MC rate. Furthermore, it suggests that an AA rate standard of 71 percent is most appropriate for achieving the aircraft utilization levels needed to satisfy pilot training requirements.
Acknowledgments

I am deeply and sincerely thankful to my Logistics Program classmates for their endless commitment to help in any way they could, in matters inside and outside of the academic arena. The camaraderie we shared was certainly a major factor in all of our success, and it has been a pleasure to be a part of the team. I also extend my thanks to my thesis advisor, Lt Col John Dickens, and the rest of the department faculty, whose drive to examine problems through unique perspectives and find new truth through research was essential to my success in this program.

Kyle R. Gladney
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>iv</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>v</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>vi</td>
</tr>
<tr>
<td>List of Figures</td>
<td>viii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>ix</td>
</tr>
<tr>
<td><strong>I. Introduction</strong></td>
<td>1</td>
</tr>
<tr>
<td>- Background</td>
<td>1</td>
</tr>
<tr>
<td>- Problem Statement</td>
<td>3</td>
</tr>
<tr>
<td>- Purpose Statement and Research Questions</td>
<td>3</td>
</tr>
<tr>
<td>- Assumptions and Limitations</td>
<td>4</td>
</tr>
<tr>
<td>- Implications</td>
<td>4</td>
</tr>
<tr>
<td><strong>II. Literature Review</strong></td>
<td>5</td>
</tr>
<tr>
<td>- Chapter Overview</td>
<td>5</td>
</tr>
<tr>
<td>- Fleet Health and Maintenance Performance Metrics</td>
<td>5</td>
</tr>
<tr>
<td>- Sortie Demand and Pilot Training</td>
<td>8</td>
</tr>
<tr>
<td>- Linear Regression Applications to Military</td>
<td>14</td>
</tr>
<tr>
<td>- Summary</td>
<td>14</td>
</tr>
<tr>
<td><strong>III. Methodology</strong></td>
<td>16</td>
</tr>
<tr>
<td>- Chapter Overview</td>
<td>16</td>
</tr>
<tr>
<td>- Scope</td>
<td>16</td>
</tr>
<tr>
<td>- Data Gathering</td>
<td>17</td>
</tr>
<tr>
<td>- Model Creation</td>
<td>17</td>
</tr>
<tr>
<td>- Residual Analysis</td>
<td>19</td>
</tr>
<tr>
<td>- Model Employment</td>
<td>19</td>
</tr>
<tr>
<td>- Summary</td>
<td>20</td>
</tr>
<tr>
<td><strong>IV. Analysis and Results</strong></td>
<td>21</td>
</tr>
<tr>
<td>- Data Gathering and Description</td>
<td>21</td>
</tr>
<tr>
<td>- Non-Modeling Data Review</td>
<td>21</td>
</tr>
<tr>
<td>- Model Creation</td>
<td>23</td>
</tr>
<tr>
<td>- Residual Analysis</td>
<td>26</td>
</tr>
<tr>
<td>- Model Employment</td>
<td>28</td>
</tr>
<tr>
<td>- Summary</td>
<td>29</td>
</tr>
</tbody>
</table>
V. Conclusions and Recommendations .................................................................31
   Conclusions of Research ..................................................................................31
   Recommendations for Future Research .........................................................32
   Summary ..........................................................................................................32
Bibliography ........................................................................................................34
### List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1. Declining F-16 Fleet performance over Time</td>
<td>2</td>
</tr>
<tr>
<td>Figure 2. Calculation Differences in MC Rate vs. AA Rate</td>
<td>6</td>
</tr>
<tr>
<td>Figure 3. Operational Requirement Formula</td>
<td>7</td>
</tr>
<tr>
<td>Figure 4. RAP Lookback Flowchart</td>
<td>11</td>
</tr>
<tr>
<td>Figure 5. FLUG Sortie Demand by Mission</td>
<td>12</td>
</tr>
<tr>
<td>Figure 6. Data Collection Sourcing and Purpose</td>
<td>17</td>
</tr>
<tr>
<td>Figure 7. Declining Utilization of Sorties over Time</td>
<td>23</td>
</tr>
<tr>
<td>Figure 8. Residual Distribution Histogram</td>
<td>27</td>
</tr>
<tr>
<td>Figure 9. Residual by Prediction Plot</td>
<td>27</td>
</tr>
</tbody>
</table>
List of Tables

Table 1. F-16 RAP Sortie Requirements ................................................................. 10
Table 2. F-16 RAP Mission Type Requirements .................................................... 10
Table 3. Descriptive Statistics ............................................................................. 21
Table 4. Descriptions of Initial Annual Models ..................................................... 24
Table 5. Descriptions of Final Annual Models ....................................................... 24
Table 6. Descriptions of Final Monthly Models .................................................... 25
Table 7. Expected Sortie Outputs at Various AA Rates .......................................... 29
I. Introduction

Background

In September of 2018, then-Secretary of Defense James Mattis directed the United States Air Force and United States Navy to increase the health of four fighter aircraft fleets – namely, the F-16, F-22, F-35, and F-18, which are cited as “Critical Aviation Platforms” – to a Mission Capability (MC) rate of 80 percent by the end of Fiscal Year (FY) 2019. This order came as a wake-up call to the flying and maintenance communities, who had all experienced declining fleet health in recent years. F-16 Wings’ problems mirrored that of most other wings operating legacy aircraft fleets: shrinking aircraft inventories, declining Aircraft Availability (AA) of the remaining fleet, and lower sortie utilization of the aircraft. For historical perspective, the last time that the active duty F-16 fleet averaged 80 percent MC over the course of a fiscal year was in FY10. It had a 24-month moving average of 72.5 percent MC at the time of the announcement. A visual depiction of the F-16 fleet’s declining performance over time is provided below. In the wake of Secretary Mattis’s order, operations and maintenance leaders across the Air Force at all levels began working initiatives, all of which collectively became known as MC80, in an effort to boost fleet health (Mehta, 2019).
Figure 1. Declining F-16 Fleet Performance over Time (LIMS-EV, 2019)

Upon the end of FY19, the deadline for Secretary Mattis’s directive, many indicators showed how the MC80 endeavor was successful within the active duty Air Force’s F-16 fleet. Largely as a result of the MC80 efforts, the FY19 annual MC rate average rose to 75.6 percent, the highest achieved since FY14. This success is even more evident when scoped down to just the active duty 291 combat coded F-16s, where FY19 yielded a 76.3 percent annual MC rate and included two months with an MC rate above 80 percent (June and July 2019), a feat that had not been achieved since April of 2013. Despite these limited successes, the conclusion of FY19 was far from “mission
accomplished” for MC80, as the expectation for FY20 and beyond is to continue working towards and beyond a sustained 80 percent MC rate (Mattis, 2018).

**Problem Statement**

Overall, a key problem with MC80 is that it appears to the unit-level maintainers tasked with achieving it to be an arbitrary goal: nobody would dispute that fighter availability can and should be improved, but why 80 percent MC rate, specifically? The MC80 order came out without a significant amount of explanation as to why that level of fleet health is more acceptable than the levels achieved in FY18, or of any level of fleet health in between. When a goal such as MC80 is set, those who are tasked with executing it would benefit from understanding the associate benefits. While it is reasonable to assume that a healthier fleet can execute more sorties than a less healthy one, this relationship is not adequately researched.

**Purpose Statement and Research Questions**

The objective of this research is to evaluate the effect of fleet health, an assumed binding constraint to USAF flying operations, on sortie execution rates. The questions that this research seeks to answer are as follows:

1) Can MC and/or AA rates be quantifiably correlated to increased aircraft utilization (SUTE)?

2) Is an 80 percent MC rate a valuable goal to achieve the necessary levels of flying? If not, what would a more functional goal be?
Assumptions and Limitations

This research is limited specifically to the Active duty F-16 fleet that carries an Aircraft Purpose Code (APC) of CC – Combat. The reasons for scoping perspective down from all of the named Critical Aviation Platforms to just this specific subset are explained in chapter III within this thesis. Furthermore, the analysis performed relies on some assumptions made about F-16 pilot manning levels, the experience mix of the pilot inventory, and the size of the F-16 fleet, in order to calculate a forecasted sortie demand. These assumptions are also explained in greater depth in chapter III.

Implications

This research aims to provide insight for the operations and maintenance communities concerning the relationship between fleet health and mission execution. This research will either validate the utility of the current fleet health goal or will provide an alternative that is based on achieving desired sortie production outcomes.
II. Literature Review

Chapter Overview

An extensive amount of base knowledge is required to understand the relationships between fleet health and fleet utilization. Fortunately, there is a wealth of academic literature examining Air Force operations and resource utilization. A review of the literature was conducted prior to and throughout the research process, covering the topics of fleet health, sortie demand and pilot training requirements, and linear regression applications to military problem sets.

Fleet Health and Maintenance Performance Metrics

It is of utmost importance to understand when conducting research about the outcomes of fleet health that the metrics by which fleet health is measured are clearly understood. MC rates are used throughout the Air Force as a key indicator of maintenance performance and are generally synonymous with fleet readiness (Oliver, 2001). MC rate is defined as the total amount of time that an aircraft or fleet accrues in fully mission capable or partially mission capable status as a percentage of the total time that the aircraft or fleet is possessed at a unit. MC rate is also a subset of a more encompassing metric, Aircraft Availability, which instead provides a percentage of total MC time for an aircraft or fleet as a percentage of the total time the aircraft is possessed by the Air Force at large, to include depot possession and unit possessed – not reported status. (Chapa, 2013). When calculating both, the numerator in each equation is the same, as AA hours or MC hours are the same thing; however, the difference is how to define
the denominator, as either possessed hours (MC rate) or Total Aircraft Inventory (TAI) hours (AA rate). AA rate is believed to give a more holistic perspective of fleet health, whereas MC rate focuses on the performance of the unit-level maintenance effort. Figure 2 illustrates how the different accounting of variables yields a difference between MC and AA rates.

The other heavily-discussed metric used throughout this research is aircraft utilization, which will be defined as the total sorties per aircraft per year. Annual sortie utilization (SUTE), is a measure of how well a unit is able to employ its assigned fleet. Dividing the total sortie output by the number of aircraft the unit has at its disposal corrects for fleet size and makes a more apt barometer of resource management between units. In the Maintenance Metrics Handbook published by the Air Force Logistics Management Agency, low fleet health (in terms of low MC rate and/or AA rate) and ineffective scheduling are pointed to as the driving factors behind low SUTE (Rainey et
al., 2001). This answer to the first research question will either support or disprove the causality of low fleet health to low SUTE.

Air Force Instruction 21-103 contains direction on how Combat Air Forces units and fleets are expected to calculate their AA standard. The first step is to determine “operational requirements” in accordance with the formula in Figure 3 below, which quantifies the requirement as a number of MC aircraft needed, on average, in a given day.

\[
\frac{S_o}{F_{do}} + \frac{(S_i)}{F_{di} \times T_u \times (1 - \alpha)} + G + S + A + R = OR
\]

**Figure 2. Formula to Calculate OR**  
AFI 21-103

This number of needed MC aircraft is divided by the TAI of the applicable unit(s) to determine the AA standard. This remains the current process for units and MAJCOMs to determine AA goals, and prior to MC80, this AA standard was also used as the baseline for determining MC rate standards.

Ritschel et al. (2018) found that there is a high correlation between AA hours and sorties flown. This is a logical conclusion, as it would make sense that an aircraft or a fleet that can be flown more will be flown more if the demand for sorties exists. The
authors determined that the general trend across all aircraft in the US Air Force is for every hour of increased availability an aircraft attains, five percent of that hour becomes a flying hour; this is a significant finding as it shows a relatively low return on the effort needed to increase AA. Ritschel et al. (2018) offered several potential explanations for this finding, primarily that most of AA hours occur at times that the aircraft cannot be utilized and is sitting idle awaiting the next flying opportunity. Another finding from this thesis was a 96.8 percent correlation between sorties executed and available hours in the whole Air Force F-16 fleet. Their study did not account for proportionality of AA hours or sorties executed to fleet size, however, such as correlating AA rate and SUTE, which is an objective that this research will attempt to achieve. The reason this distinction is important is to because when controlling for changes in fleet size, finding correlation between fleet health and utilization would indicate a cause-effect relationship, as opposed to coinciding decreases in both AA hours and total executed sorties that happen inevitably as fleet size decreases.

Sortie Demand and Pilot Training

Broadly speaking, the two requirements that drive sortie demand outside of real-world operations are upgrades and continuation training. The expectations for pilot training and progression are contained in two key documents; the most foundational of which is Air Force Manual (AFMAN) 11-2F-16V1. According to this manual, when a pilot trainee completes MDS-unique flying training and arrives at their operational unit, they must complete Mission Qualification Training (MQT) specific to their unit’s mission set. Once mission qualified, the pilot continues up a chain of upgrades: from
wingman to flight lead, from flight lead to instructor, and for some from instructor to evaluator. Each of these upgrades add value to a pilot’s utility in the squadron, as wingmen cannot fly without at least one other person of at least flight lead qualification in formation with them, and that instructor pilot(s) must fly in the formation to complete MQT or Flight Lead Upgrade sortie, and so on (FLUG) (Dahlman and Thaler, 2000).

Prior to beginning instructor pilot upgrade (IPUG), a flight lead-qualified pilot must first gain the designation of experienced pilot by accumulating a specific number of sorties in the F-16 as outlined in AFMAN 11-2F-16V1, which for most pilots is set at 250 but can be pro-rated for pilots who come from first-assignment instructor pilot duty or who cross over from another fighter airframe.

When not actively involved in an upgrade to earn flight lead or instructor/evaluator pilot status, pilots are continuously working towards attaining and maintaining monthly, quarterly, and annual flying requirements as directed by the second crucial flying requirements document; the Ready Aircrew Program (RAP) Tasking Memorandum (RTM). The number of sorties that each pilot must execute in a given month is dependent upon two criteria: first, whether they are designated as inexperienced or inexperienced, where the inexperienced pilots require more sorties than experienced; and second, if they are expected to maintain Combat Mission Ready (CMR) status or Basic MC status (BMC), where the CMR pilots require more sorties than BMC. Table 1 below illustrates monthly, quarterly, and annual sortie count requirements according to the RTM.
Table 1. F-16 RAP Sortie Requirements (ACC/A3T, 2018)

<table>
<thead>
<tr>
<th>ORG</th>
<th>Cycle</th>
<th>CMR Sorties Inexp/Exp</th>
<th>BMC Sorties Inexp/Exp</th>
</tr>
</thead>
<tbody>
<tr>
<td>RegAF</td>
<td>12 Month</td>
<td>108/96</td>
<td>72/60</td>
</tr>
<tr>
<td></td>
<td>3-Month Lookback</td>
<td>27/24</td>
<td>18/15</td>
</tr>
<tr>
<td></td>
<td>1-Month Lookback</td>
<td>9/8</td>
<td>6/5</td>
</tr>
</tbody>
</table>

Beyond simple sortie count, however, the RTM also specifies, on an annual basis, what specific mission profiles must be flown for sorties to qualify towards RAP. The table below is a depiction of the various mission types and amounts thereof that must be executed annually to stay proficient by RAP standards.

Table 2. F-16 RAP Mission Type Requirements (ACC/A3T, 2018)
The figure below is a visual representation of the continuation training concept, showing how CMR is assessed at the individual level, and how pilots can regain currency if they fail to meet RAP flying requirements. This figure is from a briefing that occurred at the US Air Force’s sortie production training summit in June 2018.

Figure 4. RAP Lookback Flowchart (Ross, 2018)

Conclusively, the various kinds of training and upgrading criteria that occur simultaneously add sensitivity and complexity to calculating and maintaining sortie demand levels. While any and all F-16 sorties count towards achieving the experience milestone, sorties only count towards achieving RAP if they can count towards one of the required mission type quotas. Upgrade sorties are even more complex and difficult to achieve, as they involve very specific scenarios that put a heavy demand on scheduling other pilots to fill support flying roles. Figure 5 shows the true extent of this “it takes a
“village…” concept, showing just how many sorties are required, in total, to achieve the relatively small number of scenarios for a single pilot in upgrade training.

![Figure 5. FLUG Sortie Demand by Mission (Walsh et al., 2019)](image)

The above figure shows how F-16 FLUG requires 62 sorties to accomplish 12 missions for the upgrade pilot. These numbers are comparable for MQT (52 sorties to accomplish 11 missions) and IPUG (62 sorties to accomplish 12 missions). Ideally, pilots are scheduled in these mission support sorties in a way that also satisfies one of their required RAP mission types, but a 2019 RAND research revealed that this does not always occur (Walsh et al., 2019).

A case study of the 388th Fighter Wing at Hill AFB, then an active duty F-16 wing, was conducted in 2000 by RAND researchers Dahlman and Thaler. In their report, the researchers found that balancing experience level and upgrade qualifications are extremely important to maintaining a consistent sortie demand. The report went on to describe the cycle of negative effects that occur when experience level skews lower than standard. First, maintenance units are overtasked to provide additional sorties demanded to keep an outsize proportion of inexperienced pilots flying at CMR rates. Additionally, flight leads and instructor pilots have to fly beyond their RAP requirements and focus on
training the surplus of inexperienced wingmen, consuming an outsize portion of available sorties. Collectively, this causes a shortage of available sorties remaining for other pilots’ continuation training, causing these pilots to miss RAP and regress. Individuals regressing and requiring additional sorties further continues the pattern of demand exceeding output, creating a “death-spiral” feedback loop. Dahlman and Thaler (2000) concluded that the proportion of CMR-experienced pilots in an F-16 flying squadron must stay at or above 52 percent experience level to avoid such a situation.

Nineteen years later, three more RAND Corporation researchers reassessed the relationship between sortie requirements, SUTE rate, manning/experience levels, and operational tempo by using a simulation of the 20th Fighter Wing’s operations at Shaw Air Force Base in South Carolina. Several pieces of information were gleaned from the findings of Walsh et al. (2019). The authors determined, along with other findings, that units are dependent upon the increased operations tempo of TDYs and deployments to achieve required sortie volume to maintain RAP standards. In addition, the authors reported that units must execute 5 to 10 percent more sorties than is dictated by squadron composition to account for non-RAP effective sorties. These ineffective sorties come from a variety of root causes, such as air aborts due to aircraft malfunction or weather, airshow flights, and sorties where training was unsatisfactory and requires reattempt. However, Walsh et al. (2019) primarily attributed the sortie waste to a mismatch of RAP mission type requirements that the pilots need for their own training versus the sortie profiles they get scheduled for in TDY and upgrade mission support scenarios. Most significantly, the study determined that an attempt to boost crew ratio in the squadron
through a surge in new FTU graduates would require increases in utilization to levels beyond the apparent ability of the maintenance enterprise in order to maintain training and upgrade timelines, and could cause a “death spiral” scenario as laid out by Dahlman and Thayer’s research from two decades earlier. Furthermore, several assumptions about flying unit makeup and operations tempo are carried forward from RAND’s simulations into this research, which will be discussed in chapter III within this thesis.

**Linear Regression Applications to Military**

Several theses produced by Air Force Institute of Technology have worked towards building regression models to explain and/or predict AA and MC rates in various platforms by using other metrics and measures of data (Oliver, 2001; Fry, 2010; Chapa, 2013). This thesis builds upon the research of these three by using similar modeling techniques but with a different aim: using AA as an independent input variable, along with other data, to predict sortie execution outputs. Also contained within these theses are a litany of other metrics and measures to consider and test as potential input variables to explain sortie execution variance in the model built in chapter IV.

**Summary**

In this chapter, the baseline understanding of Air Force flying requirements, as well as research regarding the concepts and methodologies applicable to this research effort, were discussed. This research attempts to fill a gap in the literature by testing the relationship between AA and sortie execution in a way that controls for changes to fleet size: this link, if established, will answer the first research question. Also, this chapter
provided the context for calculating and forecasting future sortie demands, which will be important in the employment of the regression model built to satisfy the second research question.
III. Methodology

Chapter Overview

This chapter will cover the scope of the research, data collection, and the steps taken to create and employ a regression model.

Scope

As mentioned in the introduction, this research will focus on the combat-coded F-16 fleet in the active duty US Air Force. This specific subset of the DoD’s aircraft inventory was chosen for several reasons. First, the field of options is narrowed down to F-16s and F-22s due to the simplicity of studying platforms that are exclusively US Air Force owned and operated within DoD. Next, the F-16 fleet is identified as preferable to F-22s for this study due to a greater amount of organic support (and thus simpler understanding of maintenance capability) and longevity of data in steady-state operations.

Once narrowed down to the F-16 fleet, the scope was further refined to active duty and combat coded APC aircraft. Focusing on this specific community is useful because all units within it have relative uniformity of sortie demand and of sortie production capacity. The uniformity of sortie demand comes from consistent manning, mission, training requirements, and scheduling of the flying squadrons across combat fighter wing. Lastly, the uniformity of sortie production capacity is defined in terms of the combat wings’ maintenance units maintaining consistent manning, funding, and scheduling.
Data Gathering

In order to build the regression model, twenty years of historical data (FY00-FY19) regarding fleet health, size, and flying execution was gathered through the Air Force’s LIMS-EV database. Data was collected from an annual (20 data points) and monthly perspective (240 data points). Then, to employ the model, current-day data regarding pilot inventory, experience mix, and RAP sortie requirements were gathered through the assignment management system database and through documents in the literature review, respectively. The below figure gives more granular detail on the data collected.

![Data used to build models and Data used to employ models]

Figure 6. Data Collection Sourcing and Purpose

Model Creation

In this phase of the research, a total of 6 models were built. The first step was to create three initial models by using simple bivariate analysis. The first of these three models cover the relationship already explored in the research accomplished by Ritschel et al. (2018), fitting raw annual sorties by raw annual AA hours. This serves as the baseline model that the other two relationships will be compared to. The next two
bivariate analyses were fitted using annual SUTE by annual AA rate and annual SUTE by annual MC rate, respectively. Once all three initial models are built, additional potentially-significant variables were added to see if they help explain more of the variation in the dependent variable while retaining a sufficiently low p-value indicating statistical significance. The outcome of this process is a linear equation creating a “line of best fit” that quantifies the relationship between the independent variable(s) and dependent variable. Equation 1 below demonstrates the standard form of a multiple regression model.

\[ Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \ldots + \beta_nX_n + \epsilon \]  

(1)

Where:

- \( Y \) = dependent variable (ex: SUTE rate)
- \( X_1, X_2, \ldots X_n \) = independent variables (ex: TAI, AA, crew ratio, etc.)
- \( \beta_1, \beta_2, \ldots \beta_n \) = coefficient of contribution of given independent variable
- \( \beta_0 \) = intercept
- \( \epsilon \) = error, unaccounted for by model

Once this process is completed, the two novel models (AA rate to SUTE and MC rate to SUTE) were compared against each other and the baseline model (AA hours to sorties) to evaluate these various relationships. This was also accomplished with the monthly data, for a total of 6 models. The only notable difference in creating these models is that monthly seasonality was accounted for with indicator variables, with October (the first month of each FY) as the baseline. When all six models were finalized, a single best fit model was selected for residual analysis and for use in determining the significance of either AA or MC rate in aircraft utilization.
**Residual Analysis**

In order to check the model’s validity, an analysis of the residuals must be accomplished. In this research, a residual analysis checked for the following characteristics: normality, independence, and constant variance. The normal distribution of residuals was assessed by checking the shape of a histogram of residuals for a “bell-curve” shape. Independence was assessed by utilizing a Durbin-Watson test and scoring a sufficiently high p-value. Finally, constant variance was assessed by plotting residuals over predicted value and looking for patterns that suggest the magnitude of residual is influenced by the magnitude of prediction, which would invalidate the model (Bowerman et al., 2005). If the model passes all three of these tests, it will be utilized in the final step of the analysis to answer the second research question.

**Model Employment**

In this final phase of analysis, the variables of the newly-built and validated multiple regression equation were populated with the calculated required SUTE and assumed values for the other independent variable to solve algebraically for the required AA or MC rate. The first step in the phase of analysis was to calculate the required SUTE by separating the pilot inventory into subsets based on experience mix and then multiplying each subset by its annual RAP requirement. These sortie requirement numbers were then aggregated, and an additional 10 percent in overage was added in accordance with the findings of Walsh et al. (2019). This final raw sortie demand was then divided by TAI to determine the SUTE requirement. This process of calculating SUTE requirement without a complete set of perfect data requires several assumptions.
The first assumption, as directed by the research sponsor, is that the pilot inventory is 100 percent of pilot authorizations at the combat wings in the Assignment Management System database. The second assumption, also directed by the research sponsor, is that TAI remains constant from the end of FY19. The final assumption is that the experience mix that Walsh et al. (2019) found at the 20th Fighter Wing and used in their simulation study can be employed within this thesis as the whole pilot inventory across the combat-coded community.

Once the SUTE requirement is calculated and all other independent variables (not including whichever fleet health metric) are populated, the equation can be solved for the fleet health metric target.

Summary

This chapter laid out the roadmap for successful model building, validation, and employment. It also laid out the necessary data collection and working assumptions for model development.
IV. Analysis and Results

Data Gathering and Description

Historical sortie execution and fleet health data were gathered, as described in the previous chapter. The below table describes the data collected for model creation.

Table 3: Descriptive Statistics (LIMS-EV, 2019)

<table>
<thead>
<tr>
<th></th>
<th>Min (Yr)</th>
<th>Annual Avg</th>
<th>Max (Yr)</th>
<th>Min (Mo)</th>
<th>Monthly Avg</th>
<th>Max (Mo)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAI</td>
<td>291 (FY19)</td>
<td>358.5</td>
<td>503 (FY00)</td>
<td>290 (Sep18)</td>
<td>393.2</td>
<td>504 (Apr00)</td>
</tr>
<tr>
<td>AA Hr</td>
<td>1.71M (FY18)</td>
<td>2.48M</td>
<td>3.25M (FY00)</td>
<td>132K (Feb18)</td>
<td>207K</td>
<td>288K (Mar05)</td>
</tr>
<tr>
<td>AA Rate</td>
<td>66.53% (FY17)</td>
<td>71.56%</td>
<td>74.83% (FY06)</td>
<td>52.6% (Sep16)</td>
<td>71.85%</td>
<td>79.2% (May05)</td>
</tr>
<tr>
<td>MC Rate</td>
<td>73.99% (FY17)</td>
<td>78.90%</td>
<td>84.01% (FY05)</td>
<td>69.41% (Sep16)</td>
<td>79.11%</td>
<td>86.33% (May05)</td>
</tr>
<tr>
<td>Sorties</td>
<td>41702 (FY19)</td>
<td>63737</td>
<td>94510 (FY00)</td>
<td>2919 (Sep12)</td>
<td>5311.4</td>
<td>9648 (Aug00)</td>
</tr>
<tr>
<td>Annual SUTE</td>
<td>123.4 (FY12)</td>
<td>158.9</td>
<td>192.7 (FY01)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Monthly SUTE</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>8.61 (Sep12)</td>
<td>13.39</td>
<td>20.05 (Aug01)</td>
</tr>
</tbody>
</table>

Also collected in this portion of the analysis were the pilot inventory of 440 pilots (Search Authorizations, 2020), projected aircraft inventory of 290 TAI (LIMS-EV, 2019), experience mix of 35 percent CMR-I, 47.5 percent CMR-E, 17.5 percent BMC-E (Walsh et al., 2019), and cost of business sortie requirement of 10 percent of calculated RAP requirement (Walsh et al., 2019).

Non-Modeling Data Review

A cursory exploration of the collected data revealed some trends that are worth understanding when moving forward with analysis. A consistent decline in TAI and a general downward trend in AA rate explain the consistent downward trend in total AA hours. Also, the consistent decrease in total sorties flown is a function of the decreasing TAI and of the general trend of decreasing SUTE. Another finding was that FY12 and
FY13 seemed to contain abnormally low sortie utilization, despite relatively healthy fleets. This is attributable to budget constraints in those years: according to conversations with contacts in ACC/A3T, the office that controls the flying hour program, FY12 was affected by an efficiency experiment where the flying units were intentionally allocated 10 percent fewer hours than the calculated RAP requirement, while FY13 was affected by the sequestration. One last significant finding was the existence of seasonality in the monthly data – certain months appeared, year over year, to yield significantly higher or lower SUTEs compared to other months. Both of these findings were significant in building regression models. What could not be gleaned from a data review is what drives the general trend of declining fleet health metrics (AA rate and MC rate), and what causal relationship – if any – exists between the general trends of declining fleet health and declining SUTE.

Another fascinating finding in this exploratory portion of the analysis was a declining utilization of AA hours across time. This ratio, which is defined as sorties per AA hours, is a different way of viewing utilization because it only accounts for the uptime of the fleet in the denominator instead of the TAI. What the graph in figure 7 below shows is that over time, the F-16 community has decreased its conversion of AA hours, which can be viewed abstractly as sortie potential, into kinetic sorties by a factor of 15-20 percent. After hitting a notable low point in FY13, it appears that the utilization of AA has leveled off at a level around .024 from FY14 through FY19. And while it is good that the downward trend is seemingly curtailing, it is still alarming that the new baseline is so low compared to the first ten years in the data sample.
The reason that this finding is so interesting is that it sits in stark contrast to what a person would expect if viewing the problem through a theoretical lens. In the face of an insatiable demand (in this case, the demand for sorties), one would assume that the resources to meet that demand (AA Hours) would be used at an increasing rate over time as they become increasingly scarce. The fact that this ratio has trended in the opposite direction is paradoxical and suggests that the Air Force has not been optimizing its existing AA. Future research should attempt to identify causal factors to this downward trend and suggest solutions for mitigating them.

**Model Creation**

As outlined in chapter III, the first step in model creation was to create three bivariate analyses, analyzing sorties to AA hours, annual SUTE to AA rate, and annual SUTE to MC rate. The results of these initial models are in the table below.
Table 4. Descriptions of Initial Annual Models

<table>
<thead>
<tr>
<th>Initial Models - Annual Data</th>
<th>Base Model</th>
<th>New Model 1</th>
<th>New Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable</td>
<td>Sorties Flown</td>
<td>Annual SUTE</td>
<td>Annual SUTE</td>
</tr>
<tr>
<td>Independent Variable</td>
<td>AA Hours</td>
<td>AA Rate</td>
<td>MC Rate</td>
</tr>
<tr>
<td>$\beta_1$ Coefficient</td>
<td>0.035</td>
<td>5.94</td>
<td>5.02</td>
</tr>
<tr>
<td>$p$-value</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td>0.0003</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.957</td>
<td>0.654</td>
<td>0.527</td>
</tr>
</tbody>
</table>

This table reveals some early insights into the process. First, the new models show a much lower adjusted R-squared value than the model that describes the relationship from the research published by Ritschel et al. (2018). Secondly, it appears that the AA rate seems to correlate better to SUTE than the MC rate does.

After several tests with other possible dependent variables, the one other significant variable that helped describe the additional variance in annual SUTE was the aforementioned budget constraint years, represented in a linear regression model with binary indicator variables. The final model, created with annual data, is described in the table below.

Table 5. Descriptions of Final Annual Models

<table>
<thead>
<tr>
<th>Final Models - Annual Data</th>
<th>Base Model</th>
<th>New Model 1</th>
<th>New Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable</td>
<td>Sorties Flown</td>
<td>Annual SUTE</td>
<td>Annual SUTE</td>
</tr>
<tr>
<td>Independent Variable 1</td>
<td>AA Hours</td>
<td>AA rate</td>
<td>MC rate</td>
</tr>
<tr>
<td>$\beta_1$ Coefficient</td>
<td>0.034</td>
<td>5.39</td>
<td>4.51</td>
</tr>
<tr>
<td>$p$-value</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td>0.0002</td>
</tr>
<tr>
<td>Independent Variable 2</td>
<td>- Budget Constraint as flying LIMFAC (binary) -</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_2$ Coefficient</td>
<td>-8026.45</td>
<td>-25.88</td>
<td>-27.78</td>
</tr>
<tr>
<td>$p$-value</td>
<td>0.0046</td>
<td>0.0024</td>
<td>0.0083</td>
</tr>
<tr>
<td>Adjusted R-Squared</td>
<td>0.972</td>
<td>0.791</td>
<td>0.653</td>
</tr>
</tbody>
</table>
This same process was then repeated with the monthly data, to similar findings, but with much lower R-squared values. This is most likely attributable to the outsize amount of noise encountered at such a granular level of data collection. One additional step that had to be taken with the monthly data was the addition of indicator variables for each month that would account for seasonality patterns. The final model created with monthly data is described in the table below.

### Table 6. Descriptions of Final Monthly Models

<table>
<thead>
<tr>
<th></th>
<th>Base Model</th>
<th>New Model 1</th>
<th>New Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable</td>
<td>Sorties Flown</td>
<td>Monthly SUTE</td>
<td>Monthly SUTE</td>
</tr>
<tr>
<td>Independent Variable</td>
<td>AA Hours</td>
<td>AA rate</td>
<td>MC rate</td>
</tr>
<tr>
<td>(x_1) Beta Coefficient</td>
<td>0.033</td>
<td>0.293</td>
<td>0.278</td>
</tr>
<tr>
<td>(x_1) p-value</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.87</td>
<td>0.48</td>
<td>0.474</td>
</tr>
</tbody>
</table>

**Note 1:**
In this model, 5 of the 11 months show significance as different from the baseline month of October. December, January, and September all have negative beta coefficients, signifying that the dependent variable underperforms (to varying degrees) in these months vs the baseline month, given all other inputs stay constant. Conversely, March and August have positive beta coefficients, signifying that SUTE overperforms in these months.

**Note 2:**
In these models, 7 of the 11 months show significance as different from the baseline month of October. November, December, January, February, and September all have negative beta coefficients, signifying that the dependent variable underperforms (to varying degrees) in these months vs the baseline month, given all other inputs stay constant. Conversely, March and August have positive beta coefficients, signifying that SUTE overperforms in these months.
When comparing the six final models in Tables 5 and 6, it becomes apparent that the relationship between availability and SUTE is better reflected over longer time domains; it appears that too much variability and noise exists in a granular monthly perspective to see the same levels of correlation. Also, the models in Figure 6 align with the ones in Figure 5 to suggest that AA rate is more applicable to explaining SUTE than MC rate is. For these reasons, the annual AA rate to the SUTE model was selected as the most appropriate model for moving forward with the research.

**Residual Analysis**

A residual analysis was performed to ensure the validity of the model prior to using it to create any further conclusions. The plots created below were used to check for normality and constant variance. While upon first glance the plots may have some possibility of being interpreted as non-normal or non-constantly varied (increasing residual magnitude as prediction magnitude increases); however, these tests are largely inconclusive, due in part to a low number of total residuals (20). The residual distribution histogram and a plot of residuals by predicted values are exhibited in the two figures below.
Since neither of these tests proved decisive in failure, the model’s validity could not be rejected. Such a judgment call was not required for detecting autocorrelation, meanwhile, as the Durbin-Watson test yielded a p-value above the .05 threshold (p=.0654) threshold and as such independence can safely be assumed. With all three of the criteria of the residual analysis satisfied, the model can finally be employed.
Model Employment

In equation form, the selected model for employment in solving for optimal AA rate is as follows:

\[ Y = -224.47 + 5.39X_1 - 25.88X_2 \] (2)

Where:

- \( Y \) = Annual SUTE (sorties/TAI)
- \( X_1 \) = AA Rate
- \( X_2 \) = Budget Constraint as Limiting Factor to Sortie Execution (binary)

Before solving for AA rate, the required SUTE must first be defined. This is accomplished by calculating the aggregate sortie demand based on pilot training requirements and dividing by the TAI. The 440 pilots are split into subgroups of 154 CMR-I, 209 CMR-E, and 77 BMC-E, which each have an annual demand of 108, 96, and 60 sorties, respectively. In summation, this yields a RAP sortie requirement of 41,316 sorties, which after the 10 percent overage allowance yields an annual sortie demand of 45,448 sorties. This number translates to an average of 156.2 sorties per aircraft per year.

With regards to the model equation, the target annual SUTE of 156.2 was substituted as the \( Y \) variable, and a value of 0 was input as \( X_2 \) to represent a lack of budget constraints affecting flying rates. Then, through simple algebra, the \( X_1 \) value was solved for which yielded a value of 70.6, meaning that an AA rate of 70.6 percent is the point prediction of fleet health requirement to achieve the 156.2 annual SUTE.

Now, it is necessary to acknowledge that the model’s point prediction does not guarantee this output with the given inputs; rather, it is much more appropriate to discuss
in terms of confidence intervals. When the model yields a 156.2 SUTE point prediction from inputs of a 70.6 percent AA rate and no budget constraints, this really means is that the model has 50 percent confidence that these inputs yield an output of at least that point prediction value. Anecdotally, for most Air Force decision-makers, 50 percent confidence is insufficient to sign off on 70.6 percent as the right goal for AA. Alternatively, A better way to attack this problem is to solve for an input that predicts the desired percent possibility of yielding at least the desired SUTE. This exercise was completed using the JMP statistical software package, with the results as depicted in the table below. In short, historical data suggests that Air Force leaders can be reasonably confident (80-95 percent) in achieving the desired SUTE rates with an AA rate of around 71 percent.

Table 7. Expected Sortie Outputs at Various AA Rates

<table>
<thead>
<tr>
<th>AA rate</th>
<th>SUTE Point Prediction</th>
<th>% confidence achieving 156.2 SUTE</th>
</tr>
</thead>
<tbody>
<tr>
<td>70.5</td>
<td>155.8</td>
<td>43.52%</td>
</tr>
<tr>
<td>70.6</td>
<td>156.3</td>
<td>52.32%</td>
</tr>
<tr>
<td>70.7</td>
<td>156.9</td>
<td>61.21%</td>
</tr>
<tr>
<td>70.8</td>
<td>157.4</td>
<td>69.72%</td>
</tr>
<tr>
<td>70.9</td>
<td>158.0</td>
<td>77.40%</td>
</tr>
<tr>
<td>71</td>
<td>158.5</td>
<td>83.93%</td>
</tr>
<tr>
<td>71.1</td>
<td>159.0</td>
<td>89.13%</td>
</tr>
<tr>
<td>71.2</td>
<td>159.6</td>
<td>93.03%</td>
</tr>
<tr>
<td>71.3</td>
<td>160.1</td>
<td>95.76%</td>
</tr>
</tbody>
</table>

Summary

This chapter walked through the data analysis and the results interpretation process employed to answer the research questions at the heart of this thesis. By
executing the plan laid out in chapter III, this research established empirical evidence supporting some compelling conclusions. First among these is that AA rate and MC Rate levels can explain a major amount of variance in fleet utilization, but that other factors beyond fleet health also can constrain SUTE. The second inference is that AA rate is a better lever than the MC rate for influencing sortie utilization outputs. Finally, this research indicates that an AA rate near 71 percent yields an appropriately high level of confidence for achieving the SUTE rates needed to meet sortie demand.
V. Conclusions and Recommendations

Conclusions of Research

The purpose of this research was to answer the research questions described in chapter I. These research questions are reiterated below, along with their respective answers.

*Research Question 1: Can MC and/or AA rates be quantifiably correlated to increased aircraft utilization (SUTE)?*

Correlation exists between these metrics. Changes in the annual AA rate explain 79 percent of the variance in annual SUTE, while changes in MC rate explain 65 percent of the variance. In the AA rate model, a 1 percent increase in fleet health correlates to an increase of 5.3 sorties/aircraft/year. While highly explanatory, both of these models have R-squared values below .80, suggesting that fleet health is not the only factor influencing sortie execution.

*Research Question 2: Is an 80 percent MC rate a valuable goal to achieve the necessary levels of flying? If not, what would a more functional goal be?*

AA rate appears to be the better of the two fleet health metrics to focus on in order to see effects in sortie execution rates. According to the model, a 71 percent AA rate yields the required annual SUTE at 84 percent confidence, which could be a more functional goal than MC80 to achieve the desired sortie outcomes. By happenstance, MC80 proved to be an effective goal to meet FY19 flying requirements due to the annual non-possession rate. However, future MC rate goals should be based on a starting point
of a calculated AA rate goal, that then factors in projections for depot possessed and unit possessed-not reported aircraft, which can vary significantly year-over-year.

**Recommendations for Future Research**

Research needs to be directed towards explaining the other binding constraints that have contributed to the decrease in SUTE over time, which potentially further explains the variance in SUTE not attributable to change in fleet health. Also, future researchers should investigate the counterintuitive phenomenon whereby the utilization of AA hours (sorties/AA hour relationship) is decreasing, as explicated in chapter IV. Finally, this methodology should be applied to other F-16 operational communities, as well as other mission design series fleets, to determine the potential similarities or differences in the relationship between fleet health and utilization based on mission type, fleet age, etc.

**Summary**

Improving fleet health has long been a leading focus of Air Force research and improvement projects. Such efforts are incredibly valuable for development and fine-tuning of tactics, techniques, procedures, perspectives, and best practices; however, these efforts come at the cost of time, effort, and dollars, and as such, must be measured. It is important that MC80 and other endeavors like it take historical performance data into account so that decision-makers can set realistic and practical goals. The MC80 goal was and continues to be, one of the more concerted and arguably successful endeavors that maintenance and sustainment units have undertaken. Moreover, great strides in the AA
rate and MC rate seem to suggest that a great deal of organizational learning has occurred and that those lessons will yield future long-term benefits. Going forward, MC80 should continue to be the goal for combat coded units in the Air Force’s active duty component, as it will both continue to push Air Force units past its current performance level and will likely yield the fleet health needed to achieve sortie execution goals for flying units. Finally, an understanding that fleet health is not the only constraint that can limit sortie execution must be considered, and a more holistic approach to the topic must be adopted to further maximize sortie output with the available resources.
Bibliography


34


Ross, Derby. "Training Design and Ops Challenges." Presented at Sortie Production Training Summit, Langley AFB VA. 19 June 2018


**MC80: Quantifying the Effect of Fleet Health on Sortie Execution in the F-16 Fleet**

**Gladney, Kyle R., Captain, USAF**

**Air Force Institute of Technology**
Graduate School of Engineering and Management (AFIT/ENS)
2950 Hobson Way, Building 641
WPAFB OH 45433-8865

**AFIT-ENS-MS-20-M-151**

**AFIT/ENS**

**HAF/A4P**

This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

**ABSTRACT**

By order of the Secretary of Defense, all of the US Air Force's F-16 units were tasked to improve their fleet health to a Mission Capability (MC) rate of 80 percent, as part of a Department of Defense-wide push to make its Critical Aviation Platforms, and the units that employ them, more ready and lethal. This study uses historical fleet health and sortie execution data captured from LIMS-EV to create a multiple regression model that quantifies the value of increased fleet health, defined as either MC rate or Aircraft Availability (AA) rate, in terms of increasing sortie output. It also uses forecasted near-future sortie demand to assess the utility of the 80 percent MC rate standard in terms of achieving desired sortie execution levels. This research concludes that both AA rate and MC rate correlate with increased aircraft utilization, and that an increase in either fleet health metric correlates to an increased annual utilization of roughly 5 sorties per aircraft. It also identifies AA rate as a more significant input to sortie execution than MC rate is, and suggests that an AA rate standard of 71 percent is most appropriate for achieving the aircraft utilization levels needed to satisfy requirements.