NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

ANALYZING U.S. NAVY F/A-18 FUEL CONSUMPTION FOR PURPOSES OF ENERGY CONSERVATION

by

David Barnhill

March 2021

Thesis Advisor: Samuel E. Buttrey
Second Reader: Lyn R. Whitaker

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

<table>
<tr><td colspan="3">REPORT DOCUMENTATION PAGE</td><td>Form Approved OMB No. 0704-0188</td></tr>
<tr><td colspan="4">Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC, 20503.</td></tr>
<tr><td>1. AGENCY USE ONLY (Leave blank)</td><td>2. REPORT DATE
March 2021</td><td colspan="2">3. REPORT TYPE AND DATES COVERED
Master’s thesis</td></tr>
<tr><td colspan="3">4. TITLE AND SUBTITLE
ANALYZING U.S. NAVY F/A-18 FUEL CONSUMPTION FOR PURPOSES OF ENERGY CONSERVATION</td><td rowspan="2">5. FUNDING NUMBERS</td></tr>
<tr><td colspan="3">6. AUTHOR(S) David Barnhill</td></tr>
<tr><td colspan="3">7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000</td><td>8. PERFORMING ORGANIZATION REPORT NUMBER</td></tr>
<tr><td colspan="3">9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
N/A</td><td>10. SPONSORING / MONITORING AGENCY REPORT NUMBER</td></tr>
<tr><td colspan="4">11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.</td></tr>
<tr><td colspan="3">12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release. Distribution is unlimited.</td><td>12b. DISTRIBUTION CODE
A</td></tr>
<tr><td colspan="4">13. ABSTRACT (maximum 200 words)
Energy usage and conservation are perennial challenges facing the Naval Aviation Enterprise (NAE) and the U.S. Navy (USN) writ large. In order to promote USN energy conservation, the Naval Air Systems Command (NAVAIR) established the Air Energy Conservation (Air ENCON) program to further analytics-driven energy consumption assessment, and assist the USN to meet broader conservation goals.
This study used a flight sortie data set built by Deloitte Consulting, constructed from three separate data sources, to assess F/A-18 fuel consumption, aiding Air ENCON analysis goals. The data set, which was derived from aircraft memory unit (MU) recordings, Naval Aviation Flight Records (NAVFLIR), and the Sierra-Hotel Aviation Readiness Program (SHARP) records, consisted of more than 466,000 USN F/A-18 sorties spanning a four-year time frame.
This research evaluated the veracity of sortie data fuel output metrics and identified broad fuel consumption trends despite a significant proportion of missing or unused information. Furthermore, this thesis documents the effectiveness of the data to predict fuel consumption by use of original and generated predictors in combination with various imputation methods. Results suggest that while statistical inference is difficult due to the amount of missing data, broad trends related to sortie location are identifiable, and models using imputation coupled with original and generated predictors exhibit the best results for predictive effectiveness.</td></tr>
<tr><td colspan="3" rowspan="2">14. SUBJECT TERMS
fuel, fuel conservation, F/A-18, E-2D, E/A-18, naval aviation, machine learning, predict</td><td>15. NUMBER OF PAGES
131</td></tr>
<tr><td>16. PRICE CODE</td></tr>
<tr><td>17. SECURITY CLASSIFICATION OF REPORT
Unclassified</td><td>18. SECURITY CLASSIFICATION OF THIS PAGE
Unclassified</td><td>19. SECURITY CLASSIFICATION OF ABSTRACT
Unclassified</td><td>20. LIMITATION OF ABSTRACT
UU</td></tr>
</table>

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

ANALYZING U.S. NAVY F/A-18 FUEL CONSUMPTION FOR PURPOSES OF ENERGY CONSERVATION

David Barnhill
Commander, United States Navy
BS, U.S. Naval Academy, 2005
MS, Naval Postgraduate School, 2021

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
March 2021

Approved by: Samuel E. Buttrey
Advisor

Lyn R. Whitaker
Second Reader

W. Matthew Carlyle
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Energy usage and conservation are perennial challenges facing the Naval Aviation Enterprise (NAE) and the U.S. Navy (USN) writ large. In order to promote USN energy conservation, the Naval Air Systems Command (NAVAIR) established the Air Energy Conservation (Air ENCON) program to further analytics-driven energy consumption assessment, and assist the USN to meet broader conservation goals.

This study used a flight sortie data set built by Deloitte Consulting, constructed from three separate data sources, to assess F/A-18 fuel consumption, aiding Air ENCON analysis goals. The data set, which was derived from aircraft memory unit (MU) recordings, Naval Aviation Flight Records (NAVFLIR), and the Sierra-Hotel Aviation Readiness Program (SHARP) records, consisted of more than 466,000 USN F/A-18 sorties spanning a four-year time frame.

This research evaluated the veracity of sortie data fuel output metrics and identified broad fuel consumption trends despite a significant proportion of missing or unused information. Furthermore, this thesis documents the effectiveness of the data to predict fuel consumption by use of original and generated predictors in combination with various imputation methods. Results suggest that while statistical inference is difficult due to the amount of missing data, broad trends related to sortie location are identifiable, and models using imputation coupled with original and generated predictors exhibit the best results for predictive effectiveness.

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

Air ENCON	Air Energy Conservation
AOR	area of responsibility
AWG	Aviation Working Group
BUNO	bureau number
CCDR	Combatant Commander
CNAF	Commander, Naval Air Forces
CNO	Chief of Naval Operations
CONUS	continental U.S.
CSG	Carrier Strike Group
CVN	nuclear powered aircraft carrier
CVW	Carrier Air Wing
DECKPLATE	Decision Knowledge Programming for Logistics Analysis and Technical Evaluation
DoN	Department of the Navy
EAG	Energy Academic Group
FDNF	Forward Deployed Naval Forces
FHP	Flight Hour Program
FPC	flight purpose code
FRS	fleet replacement squadron
Hz	hertz

ICAO	International Civil Aviation Organization
iENCON	Incentivized Energy Conservation
IFR	in-flight refueling
KW	Kruskal-Wallis
LOCF	last observation carried forward
MAR	missing at random
MCAR	missing completely at random
MDS	Maintenance Data System
MI	multiple imputation
MNAR	missing not at random
MU	memory unit
NAE	Naval Aviation Enterprise
NALCOMIS	Naval Aviation Logistics Command Management Information System
NAOEP	Naval Aviation Operational Energy Program
NAVAIR	Naval Air Systems Command
NAVFLIR	Naval Aviation Flight Record
NAVSEA	Naval Sea Systems Command
NPS	Naval Postgraduate School
NRMSE	normalized root mean squared error
OFRP	Optimized Fleet Response Plan
OOMA	Optimized Organizational Maintenance Activity
PMA-265	F/A-18 and EA-18 program office

ROC/POE	Required Operational Capabilities/Predicted Operational Environment
SDR	Secure Data Repository
SFWSL	Strike Fighter Weapons School Atlantic
SFWSP	Strike Fighter Weapons School Pacific
SHARP	Sierra Hotel Aviation Readiness Program
SMART	Short-cycle Mission and Recovery Tanking
TACAIR	Tactical Aircraft
TFE	Task Force Energy
TMR	Total Mission Requirement
TMS	Type/Model/Series
TOD	Time of Day
TYCOM	Type Commander
USMC	U.S. Marine Corps
USN	U.S. Navy
VAQ	Electronic Attack
VFA	Strike Fighter
WMW	Wilcoxon-Mann-Whitney

THIS PAGE INTENTIONALLY LEFT BLANK

Executive Summary

Since 2010, the Naval Aviation Enterprise (NAE) has placed an increased focus on energy conservation and petroleum-based fuel consumption in its aircraft fleet. This followed implementation of new Department of the Navy (DoN) energy policy outlining methods to establish energy security and independence fleet-wide (Department of the Navy 2010a).

As a result of these energy conservation efforts, the NAE, through Naval Air Systems Command (NAVAIR), established the Air Energy Conservation (Air ENCON) program with the intent to utilize data-driven analysis to identify methods to reduce fuel consumption in the NAE fleet-wide (Olszewski et al. 2012). To date, these efforts have led to implementation of technological, material, procedural, and cultural solutions to help the NAE reduce fuel consumption across the community through various energy initiatives.

In 2019, NAVAIR's Naval Aviation Operational Energy Program (NAOEP) commissioned Deloitte Touche Tohmatsu Consulting to develop a Tableau Fuel Conservation Analytics Dashboard supporting assessment of fuel consumption metrics focused on F/A-18 platforms with the intent to extend the analysis capability and apply it to all NAE platforms. The Dashboard utilizes aircraft sortie information from 2016 to 2019 from three separate databases, merged to form a rich data set of approximately 466,000 sorties from F/A-18E, F/A-18F, and EA-18G platforms. The resultant output provides average and total fuel metrics for numerous aggregation categories.

In 2020, the NAOEP sponsored research through the Naval Postgraduate School (NPS) Energy Academic Group (EAG) to assess the Fuel Conservation Analytics Dashboard's underlying data set. In the resultant report, Barnhill et al. (2020) determined that the data set exhibited significant deficiencies rendering much of the Dashboard output unreliable. Specifically, the research showed that missing data is pervasive and unpredictable because the Dashboard incorporates only certain types of sorties in its metric calculations and in many cases sorties are not present where they should exist. This results in about 30% of known sorties being unincorporated and an unknown amount simply missing. Because of the nature of why data might not be incorporated and how much of the data is missing, detailed trend analysis is untenable and output metrics are unreliable due to bias.

This thesis continues the research begun by Barnhill et al. (2020). First, this study analyzes the effect of missing and unincorporated sorties, as well as missing data field values, on tractable assessment of Dashboard output metrics focused on deployed sorties. Additionally, this study identifies broad trends in fuel consumption despite missing and unincorporated deployed sorties by specifically focusing on deployment destination. Lastly, this research assesses prediction of deployed sortie fuel consumption metrics by use of predictors originally in the data set and predictors generated by the author. Additionally, we employ imputation methods to increase available sortie observations and use the imputed data sets to determine how well complete data may be used to predict fuel consumption.

The research results reiterate that the unpredictability of missing or unused data severely affects fuel consumption analysis. Further, results show that comparison between aggregation categories, even if they are similar in nature, is untenable because of differing proportions of missing observations. Despite missing information, broad trend analysis by deployment destination is possible. However, these trends are not detectable using the current categories in the sortie data and highlight the need for more detail in the data set. While trends are identifiable they should be considered only as an area for further research due to the missing information.

The use of variables original to the data set produce only marginal ability to predict fuel consumption. By contrast, models utilizing original predictors in combination with predictors constructed to capture more detailed information showed significant improvement. Imputation methods, specifically multiple imputation, employed to increase the number of available observations, amplify this improvement. However, the metrics only indicate how well predictors model known data and do not account for effects of missing sorties or data values.

While NAOEP's Analytics Dashboard provides a user-friendly interface for analysts to assess various fuel consumption metrics, deficiencies, specifically in missing data, prevent tractable analysis. However, the value of the data cannot be exaggerated and NAOEP should consider its utility in a variety of future energy conservation initiatives.

References

Barnhill D, Morgan B, Buttrey S, Whitaker L, Nussbaum D (2020) Evaluating NAVAIR's fuel dashboard. Unpublished technical report, Naval Postgraduate School, Monterey, CA

Department of the Navy (2010a) Department of the Navy's energy program for security and independence. Accessed December 6, 2020, https://www.secnav.navy.mil/eie/Documents/DoNEnergyProgramforSecurityandIndependence.pdf

Olszewski M, Quinn D, Noel B (2012) Navy air energy conservation (Air ENCON) program. NDIA Environment, Energy Security, and Sustainability (E2S2), New Orleans, LA. https://apps.dtic.mil/dtic/tr/fulltext/u2/a565706.pdf

THIS PAGE INTENTIONALLY LEFT BLANK

Acknowledgments

I would like to express my humblest gratitude to my advisor, Professor Buttrey. Thank you for the guidance throughout this project and providing rudder when needed.

To my second reader, Professor Whitaker, I appreciate very much your thoughts and insights as this project moved forward. Your direction was invaluable.

To Professor Nussbaum, thank you for including me on this project. It is a good feeling to be able to provide a modicum of support to naval aviation, a community I dearly love.

I will be forever grateful to CAPT Brian Morgan USN (Ret.). Your leadership, guidance, and input during the project was more important than I can intimate. This endeavor would have been incalculably more difficult without your help. Thank you.

To NAVAIR and NAOEP, with any luck we have provided you some insight and clarity to continue your great work supporting the fleet. Thanks for listening to our inputs with grace. You are all great professionals.

To my cohort, Maj Luke Flinn, LCDR Aaron Fish, LCDR Ian Brenner, LCDR Chris Norman, LT Dan Pulliam, LT Kevin Lutz, LT James Wiltshire, and LTJG Ben Liu. Thanks for the help, inspiration, tutoring, discussions, and clarity during our tenure here. I wish you all the best.

Finally, to my wife, Jennifer, and children, Meredith, Abigail, and John: you all are the strongest, most vibrant people I know. You provide the best support I could ask for in this and all my endeavors. I love you all.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 1:
Introduction

1.1 Background

Energy conservation has been a consistent strategic goal for the Department of the Navy (DoN) for over a decade. In 2009, former Secretary of the Navy Ray Mabus promulgated five strategic energy goals to achieve Energy Security and Independence for the DoN. Of these, the primary identified effort was to reduce the use of petroleum-based energy by 50% Navy-wide, replacing it with alternative sources (Department of the Navy 2010a). To this end, the Chief of Naval Operations (CNO) Admiral Gary Roughead, directed the establishment of Task Force Energy (TFE) to support the effort of reducing the use of petroleum-based liquid fuels. Task Force Energy consists of seven working groups with members from communities across the Navy, representing various energy interests (Department of the Navy 2010c).

Naval aviation consistently finds itself amongst the largest consumers of petroleum-based fuels in the DoN. In FY2010, NAE fuel consumption consisted of approximately 42% of all U.S. Navy (USN) fuel usage equating to 580 million gallons of fuel or, in aviation terms, approximately four billion pounds of aviation fuel (Olszewski et al. 2012).

From TFE and its Aviation Working Group (AWG), the Naval Air Systems Command (NAVAIR) established the Air Energy Conservation (Air ENCON) program in 2012 under the authority of the Naval Aviation Operational Energy Program (NAOEP). The Air ENCON program is a derivative of the Incentivized Energy Conservation (iENCON) program instituted by Naval Sea Systems Command (NAVSEA) to reward ship commands for identifiable energy conservation efforts (Olszewski et al. 2012). NAVSEA established iENCON in 1999. In 2008 alone, iENCON helped the U.S. Navy achieve $136 million in fuel cost savings (Salem et al. 2009).

As with iENCON, the goals of the Air ENCON are to identify and implement techniques and procedures, as well as changes to Naval Aviation Enterprise (NAE) community culture, to reduce fuel use across the NAE while maintaining safety and readiness standards (Olszewski et al. 2012). Through the Air ENCON program, NAOEP identifies four broad areas of focus:

1. Develop innovative fuel-savings processes for deployment NAE-wide.
2. Establish fuel metric and reporting elements to assess fuel consumption at the unit level.
3. Communicate aviation energy policy and process changes to NAE leadership and personnel.
4. Incentivize conservation efforts to instill a community-wide "Culture of Conservation." (Olszewski et al. 2012)

To engage in these areas of focus, NAOEP uses data-driven analysis to identify new fuel conservation methods. Since the inception of the Air ENCON program, NAOEP has spearheaded many fuel conservation initiatives that have since been incorporated in NAE standard operating procedures. These initiatives vary from general award recognition of energy-conscious units to NAE-wide procedural, technological, and material implementation.

NAOEP and Air ENCON efforts have proved most effective in altering ground and maintenance procedures and encouraging non-operational flight events to be conducted in high-fidelity simulators. However, as large amounts of longitudinal operational and specifically deployed sortie data is collected, NAOEP seeks to evaluate in-flight fuel consumption trends to identify operational flight profiles resulting in greater fuel expenditure. The overall goal is to introduce procedure change recommendations in order to increase fuel conservation efforts without degrading operator proficiency or safety. While fuel conservation is viewed largely from a cost perspective, NAOEP intends these efforts to help inculcate a culture of conservation within the NAE (Olszewski et al. 2012).

1.2 Fuel Conservation Analytics Dashboard

In 2019 NAOEP, in support of the Air ENCON program, contracted Deloitte Touche Tohmatsu to develop a Tableau dashboard tool to be used for analysis of F/A-18 platform variant fuel consumption (Tableau Software, Inc 2020). In developing the dashboard, Deloitte professionals merged sortie data sets originating from three separate databases:

1. The aircraft memory unit (MU) database.
2. The Sierra Hotel Aviation Readiness Program (SHARP) database.
3. The Naval Aviation Flight Record (NAVFLIR) database.

MU data is derived from in-flight aircraft sensor recordings and is drawn from the NAVAIR 6.8.4 Secure Data Repository (SDR). SHARP and NAVFLIR data are taken from records logged post-flight by the pilot. Both SHARP and NAVFLIR data sets were pulled from Decision Knowledge Programming for Logistics Analysis and Technical Evaluation (DECKPLATE). The three data sets were compiled into a single data set by the Deloitte Analytics consultants supporting the NAOEP (NAVAIR 4.4) team for ultimate display in a Tableau F/A-18 Fuel Conservation Analytics Dashboard. The resulting data set consisted of 466,401 unique sorties split among three different Type/Model/Series (TMS) aircraft, specifically, F/A-18E and F/A-18F Super Hornets and EA-18G Growlers, spanning a four year period from 2016 to 2019.

In addition to the database of unique sorties, Deloitte incorporated a second database of MU output of 33 variables measuring various parameters and aircraft system output. This data is recorded at a 60 hertz (Hz) refresh rate while an aircraft is airborne. In pre-processing, MU sensor information was aggregated into ten-second intervals, resulting in approximately 160,000,000 observations.

In 2020 NAVAIR engaged the Naval Postgraduate School (NPS) through the Energy Academic Group (EAG) to execute a comprehensive study of the Dashboard and data fidelity. Specifically, the study conducted by Barnhill et al. (2020) examined the Dashboard's underlying data and resultant output metrics.

1.3 Research Focus

In completing the analysis of the underlying Dashboard data, the NPS team found that the data suffered from a non-trivial amount of missing information making use of statistical output generated by the Dashboard tool tenuous at best. However, in the analysis, Barnhill et al. (2020) suggested that analyses is still viable, though only at high aggregation levels (e.g., TMS by month). To support future analysis, this thesis extended prior research by examining the unique sortie database, focusing on three areas:

1. Continued analysis of the effect of missing sorties, data field values, and MU data on analysis of Dashboard output metrics.

2. Identify broad trends in fuel consumption despite the amount of missing and unincorporated data. The study assessed whether there is a correlation between higher fuel consumption and deployed area of responsibility (AOR) (i.e., the Middle East or western Pacific) and what types of sortie operations may have affected higher fuel consumption.
3. The study assessed prediction of MU fuel consumption response variables using random forest models based on original and generated predictors and using various imputation methods.

This study reaffirmed that the Fuel Conservation Analytics Dashboard output metrics derived from the merged data set are highly uncertain due to the amount of missing MU data. However, despite the missing data, the study gives evidence that broad trends are identifiable when information outside the data set is included. Specifically, there is a correlation between higher fuel consumption and sorties occurring in the Middle East AOR with a direct relation to the types of missions conducted. Lastly, the study determined that using imputation techniques and a combination of generated and original predictors yielded more effective prediction results than relying exclusively on resident information. However, analysts must be cautious in interpreting these results as the reasons for sortie data being missing cannot be ascertained directly, potentially resulting in bias.

CHAPTER 2:
Prior Research

2.1 Literature Review

This section summarizes research in three areas: policy and research regarding fuel consumption in the USN writ large and the NAE specifically; the effects of missing data in a data set; and techniques used to impute missing information to increase data fidelity and enhance predictive ability.

2.1.1 NAE Energy Policy

In 2010 the Department of the Navy DoN published its Energy Program for Security and Independence (Department of the Navy 2010a). The primary objective of the program as outlined by the Secretary of the Navy at the time the Honorable Ray Mabus, is two-fold: 1) Energy Security and 2) Energy Independence. The policy defines Energy Security as “utilizing sustainable sources that meet tactical, expeditionary, and shore operational requirements and force sustainment functions, and having the ability to protect and deliver sufficient energy to meet operational needs.” Energy Independence is defined as relying on energy sources not “subject to intentional or accidental supply disruptions” (Department of the Navy 2010a).

According to the documentation, the NAE used 42% of petroleum consumed in FY 2008 within the DoN (Department of the Navy 2010a). For the NAE to meet the established energy goals, DoN leadership placed focus on technology development, specifically developing more fuel-efficient engines and employing the use of bio-fuel as a method for energy conservation. Additionally, the Energy Program sought to reduce non-operational fuel consumption by substituting “in-aircraft” training with simulators.

Also in 2010, the DoN published A Navy Energy Vision for the 21st Century with the intent to “identify ends, ways, and means for increasing energy security for the Navy” (Department of the Navy 2010b). Of the ten tenets driving the Energy Vision, two apply directly to the the way the NAE operates. These tenets are to consider energy demands in broad, strategic planning and to incorporate energy-efficient operating procedures and technology (Department of the Navy 2010b).

From these policies, the USN has seen consistent technological investment to support energy conservation in the NAE whether by implementing alternative fuel types or increasing the efficiency of aircraft engines. At the time of writing, the Energy Vision had been emphasizing the need to continue shifting non-operational flying to training simulators across the aviation fleet. The NAE has seen prior success in this area where increased use of simulators saw a reduction in fuel costs of $45 million in the USN's aviation training community in 2008 alone (Department of the Navy 2010b). The Department of the Navy (2010b) sought to extend the use of simulators to operational squadrons, as well.

While technological improvements and training shifts help the energy conservation effort, a change in "culture" is necessary to effect long-term meaningful impacts in the NAE. The Department of the Navy (2010b) acknowledged the need for a culture shift but did not indicate how this might come about except in vague terms of increasing awareness and disseminating best operational practices throughout the fleet.

The USN Air ENCON Program was developed in 2012 in response to the Energy Vision and is administered by NAVAIR's NAOEP. The Air ENCON program is modeled on NAVSEA's iENCON program put in place to enhance fuel conservation in the surface community (Olszewski et al. 2012). Similarly, the Air ENCON program emphasizes fuel conservation in the NAE by defining the broad categories discussed in the Energy Vision, namely, to increase awareness and implement best practices.

The Air ENCON program uses a metric-focused approach to measure success in fuel conservation. The program tracks fuel consumption by TMS and squadron, accounting for both training and operational fuel usage. Ancillary focus is placed on implementing best practices. To date, NAOEP, through the Air ENCON program, has identified multiple technological implementations and ground operation practices such as utilizing "cold" truck refueling and employment of Short-cycle Mission and Recovery Tanking (SMART) leading to decreased fuel consumption (Olszewski et al. 2012). While these procedural implementations have yielded identifiable benefits, there has been little research done on NAE- or TMS-wide in-flight consumption trends or analysis.

2.1.2 Fuel Conservation

Fuel conservation in the NAE has been the subject of several research initiatives, though there is little detailed analysis associated with operational fuel consumption.

One study conducted by Salem et al. (2009) investigated the effectiveness of the various energy programs. The goal of the study was to identify best practices and recommend extensions across the USN. Of particular interest was the evaluation of the surface Navy's iENCON program and its effect on fuel conservation. The researchers conducted the study prior to Air ENCON being established and it specifically recommended the program be extended to the aviation community (Salem et al. 2009).

Salem et al. (2009) also identified cultural aspects of separate warfare areas by interviewing representatives from the different communities. As applicable to the NAE, the authors noted that pilots and aircrew possess significant latitude in how fuel is consumed because usage is considered a safety concern. Pilots are seen to be the primary stakeholders in balancing safety with fuel loading and usage. Indeed, the aviation community is seen to base its conservation decisions on perceived risk (Salem et al. 2009). Importantly, the results from the study highlighted aspects of flying that focused less on analysis of in-flight procedure and more on improved use of infrastructure (e.g., runways usage) or protocol than on procedural analysis. However, one specific recommendation that concurs with the recommendations provided by Barnhill et al. (2020) was to utilize operational staffs to identify technological and process improvement in the NAE (Salem et al. 2009). The extension of this recommendation is that operator input is necessary to drive procedural or cultural change. Additionally, staff input informs decision makers who can enforce conservation efforts. This highlights the current position of NAOEP where the office makes recommendations through Air ENCON but cannot mandate change without enforcement from operational decision makers.

2.1.3 NAE Database Fuel Analysis

Since establishing the Air ENCON program, NAVAIR and NAOEP have spearheaded numerous efforts to identify and implement procedures to reduce fuel consumption. Only recently, has NAVAIR leveraged available sortie databases to analyze fuel usage. Nonetheless, several research teams have conducted a small number studies using sortie databases.

In their 2016 white paper, Eger et al. (2016) sought to identify alternative metrics for readiness by assessing consumable fuel while enhancing pilot proficiency. To accomplish this task, the authors accessed several, separate naval aviation databases used by NAE entities to report readiness metrics. Among these databases were the SHARP and Naval Aviation Logistics Command Mangagement Information System (NALCOMIS) and its Optimized Organizational Maintenance Activity (OOMA). NALCOMIS is the software program pilots use to record all NAVFLIR data. Inputs to these systems are post-flight logs completed by the aircrew and which ultimately serve to record flight hours both for operational (SHARP) and maintenance (NALCOMIS) tracking.

Pertinent to this thesis, Eger et al. (2016) encountered numerous issues with the veracity and availability of the data in determining alternative fuel metrics. First, in developing new metrics, the authors identified the difficulty in amalgamating data from separate systems managed by different agencies. While necessary, this amalgamation led to an increase in variability, negatively affecting the accuracy of the new metrics (Eger et al. 2016). Additionally, Eger et al. (2016) addressed the accuracy of the data sets themselves. Because both SHARP and NALCOMIS records are completed after a flight, input information is based solely on aircrew memory. Each record is therefore subject to human error potentially introducing variability or bias into the analysis (Eger et al. 2016).

Barnhill et al. (2020) analyzed F/A-18 fuel data in a NAOEP-sponsored project assessing the fidelity of a data set constructed by Deloitte Touche Tohmatsu from SHARP, NAVFLIR, and MU data stored in the NAVAIR 6.8.4 SDR. The three data sets were subsequently merged to form a single data set. In addition, NAOEP requested a trend analysis on a variety of fuel metrics. The merged data set would be used in the aforementioned Tableau Fuel Conservation Analytics Dashboard. From this analysis, the authors determined that any resultant output could not be used for meaningful conclusions due to the amount of unused or missing data. This further affected the desired trend analysis, negatively. Given the amount of missing data and its unpredictability, trend analysis could only be conducted in broad categories of little value or in very specific instances where the analyst is aware of the proportion of sorties available (Barnhill et al. 2020).

2.1.4 Missing Data

Missing information is a pervasive concern in data analysis. Rarely is data received fully formed and "missingness" is manifested in many ways. The underlying data supporting the Tableau Fuel Conservation Analytics Dashboard is no different. Analysis showed that of approximately 466,000 unique sorties, only 313,000 or approximately 67% are incorporated into the Dashboard output (Barnhill et al. 2020). This section discusses the effects of missing data on statistical output and assesses acceptable proportions of missing data. Additionally, this section defines different patterns of missing data, and the relationship between pattern and proportions of missing data.

Missing information in a database may manifest itself in several ways. In some cases it is simply missing values in database categories while in others it may exist as missing observations where observations should be present. Kang (2013) identified four areas of concern when information is missing in a data set. First, statistical power is weakened in conducting hypothesis testing. Specifically, the probability of rejecting the null hypothesis when it is false is reduced. Second, using data with missing values can cause statistical bias in estimating output parameters. This can be true whether portions of or entire observations are missing. Third, missing values lessen the likelihood that a sample is representative of the population. Lastly, is that missing data can complicate the analysis of a study (Kang 2013).

The proportion of missing data is an important aspect to consider when addressing "missingness." However, there is evidence that the pattern, not the proportion, of missing data is predominant in evaluating a data set with missing values (Tabachnick and Fidell 2013). Unfortunately, there is no defined threshold for acceptable amounts of missing data. Tabachnick and Fidell (2013) suggest that if 5% of the data is missing in a random pattern, then there is little cause for concern. A 5% threshold was also used by Schafer and Graham (2002). Others suggest that 10% is an acceptable threshold (Bennett 2001). Regardless, all agree that missing data can insert bias into any output. This is of primary concern in this study as output metrics derived from the Fuel Conservation Analytics Dashboard display no obvious indication of missing information.

In determining how to work with missing data, an analyst must determine the pattern of "missingness" if possible. Rubin (1976) defined these patterns in three forms: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR). MCAR data has little effect on parameter bias because the missing information is independent of either the observed data or the missing data so the observed data may be treated as a random sample of the complete data (Dong and Peng 2013). By contrast MAR data indicates that missing information is predictable from other variables in a data set. That is, if variable values are missing in a data set, “missingness” can be predicted by other variables in the observable data set (Tabachnick and Fidell 2013). The final category as defined by Rubin (1976) is MNAR. MNAR is defined such that the probability of data being left out is different for separate categories and the reasons for this are unknown (Van Buuren 2018). While missing pattern categories have clear definitions, missing data rarely fall exclusively into a single category. Further, it can be very difficult to differentiate between the three as reasons for “missingness” are myriad and can be dependent on factors not captured in the data.

It may be intuitive to assume that an analyst can determine the proportion of missing data and assess if too much is missing to make statistical inference. While the proportion of missing data is important, determining the pattern of missing data is more so (Tabachnick and Fidell 2013). In determining the pattern, the analyst can determine whether the missing information is ignorable or not. Broadly, MAR data is ignorable or non-informative while MNAR data is not (Schafer and Graham 2002). Barnhill et al. (2020) show that there are patterns of missing data for certain squadrons and over specific time frames, strongly indicating that the sortie data is MNAR. This study assumes nothing different.

2.1.5 Imputation

To avoid or limit bias due to missing data fields, imputing missing data with other values can be a useful strategy. While there are a number of imputation strategies, the ones considered in this thesis involved single-value imputation and multiple imputation (MI). Using these strategies to fill out missing data values, the study then predicted known MU fuel consumption values using random forest models and assessed the prediction capability for data sets with and without imputed data.

While imputation is the primary focus in this section, it is worth discussing why deleting observations may not be effective. When handling missing data it can be tempting to remove those observations with missing values. Ignoring missing data may increase or induce more bias if the data is not considered MCAR or, in few cases, MAR (Schafer and Graham 2002). MCAR data can be considered representative of the population so deletion can be acceptable. In MAR data, deletion is not recommended except in very specific circumstances. One example is using complete cases (ones in which all desired fields are populated) in a data set to estimate parameters in a response which is the only variable exhibiting "missingness." While the parameter estimates are acceptable, other inferential statistics exhibit concerning bias (Schafer and Graham 2002). Schafer and Graham (2002) state succinctly that results from case deletion are likely biased because the complete cases are not representative of the population.

As an alternative to case deletion, single-value imputation allows the analyst to estimate the missing values in a category with one value. This value may be derived in several ways such as the last observation carried forward (LOCF) where missing values are imputed using the last value observed from a participant or category. Alternatively, a measure of central tendency may be used to impute data. Oftentimes, the mean is chosen but the median may be used as well, depending on the data observed. All of these cases require the assumption that imputed values are similar to those observed. Single-value imputation is easy and simple but may induce bias and result in underestimated variance (Jakobsen et al. 2017). In the case of this study the mean value was employed to impute missing numeric fields for one of the four prediction models.

MI involves calculating or determining multiple estimates for any particular missing value and combining the estimates to determine the imputed value. MI typically has three phases: 1) construct complete data sets with imputed values several times; 2) estimate desired parameters using any desired statistical method; and 3) pool the estimates into a single output. By taking several imputations the analyst can reduce the uncertainty associated with any single imputation. MI is often used for MAR data; however, differentiating between MAR and MNAR data is difficult since the missing data is unknown and therefore cannot be assessed as to whether it can predict the unknown data well or not (Jakobsen et al. 2017). If MI is used and MAR cannot be proven, analysts should conduct sensitivity or qualification analysis on the result if deriving inferential statistics (Leurent et al. 2018).

While there are numerous MI methods, of particular interest in this paper are those found in the “missForest” package in the R programming language (Stekhoven and Buhlmann 2011; R Core Team 2017). The “missForest” package was developed by Stekhoven and Buhlmann (2011) to implement MI methods to handle mixed-type (continuous and categorical) data sets with few assumptions. The method uses random forest methodology developed by Breiman (2001) to impute missing values only using the observed portions of the data set. Essentially, the algorithm treats each variable with missing values as a response and uses remaining variables as predictors. A random forest is applied to predict the missing values of the data set variables in iterative fashion. After each iteration, normalized root mean squared error (NRMSE) is computed for numeric variables and the percentage of incorrect classification for categorical ones (Stekhoven 2011). When both of these values increase from the previous iteration, the algorithm stops and the imputed data set is taken to be the result of the next-to-last iteration (Stekhoven 2011).

This study employs MI using “missForest” to develop a model to further assess how well predictors resident in the original data set and generated by this author model fuel consumption response variables. While MI is effective, in this data set, because of its MNAR categorization, the analyst must consider that the model only assesses how well the predictors model known responses and should not be considered as an assessment of unknown observations.

CHAPTER 3: Data and Methodology

3.1 Introduction

This chapter describes the data supporting the Fuel Conservation Analytics Dashboard. The chapter begins with a discussion of data sources and paths, data inclusion, variables constructed through merged data sets, and derivation of fuel metrics. A discussion of the Fuel Conservation Analytics Dashboard aggregation categories and output metrics follows. Lastly, we discuss data cleaning, generated categories, and methodologies for further analysis of the data set.

The underlying data consisted of two separate data sets, a unique sortie data set and a continuous fuel data set, though we will only consider the unique sortie data set in this analysis. Much of the content is taken from NPS analysis conducted by Barnhill et al. (2020) in support of NAOEP.

3.1.1 Data Sources

The unique sortie data set supporting the Dashboard is compiled from three separate data sources: MU, SHARP, and NAVFLIR databases.

MU Data: The MU data set contains aircraft information as recorded by organic sensors relating information from aircraft instruments and systems. After each flight, the aircrew removes the on-board recording device so the data may be incorporated into a central database. Different codes are used to provide the systems' statuses information for the duration of the flight. Event-driven information is recorded in binary form (Deloitte Analytics Team 2020).

SHARP Data: The SHARP record is a post-flight log, manually recorded by the aircrew. SHARP records not only go into aircrew logbooks but also directly report squadron readiness (Deloitte Analytics Team 2020). Several fields of the SHARP log are similar to those in the MU data record, such as fuel burned or used, but accuracy is reliant on aircrew memory and is subject to human error. A SHARP record is required to be recorded for each flight (CNAF 2016).

NAVFLIR Data: The NAVFLIR is a record of aircraft utilization for an individual sortie. Like SHARP these logs are recorded post-flight by the aircrew. Specifically, this record is used for maintenance purposes to track necessary maintenance actions as an element of the Maintenance Data System (MDS). A NAVFLIR utilization report must be completed after each flight (CNAF 2017).

Ideally, all three data sources would possess records related to every sortie conducted, and where two sources report the same field, these values would be the identical. Because SHARP and NAVFLIR records are filled out by the aircrew, these records are subject to human error. For this reason, the MU data, where present, is considered “truth” when examining fuel consumption.

Figure 3.1 illustrates the flow of both MU and NAVFLIR records to various data repositories. MU data is recorded during the flight and the MU card is provided to squadron maintenance department. The maintenance department transfers the MU files to the F/A-18 and EA-18 program office (PMA-265) repository where it can then be provided to Boeing for data storage. The record is also incorporated into a Hadoop cluster serving as a SDR aviation data warehouse where it may be accessed via DECKPLATE (NAVAIR 2020).

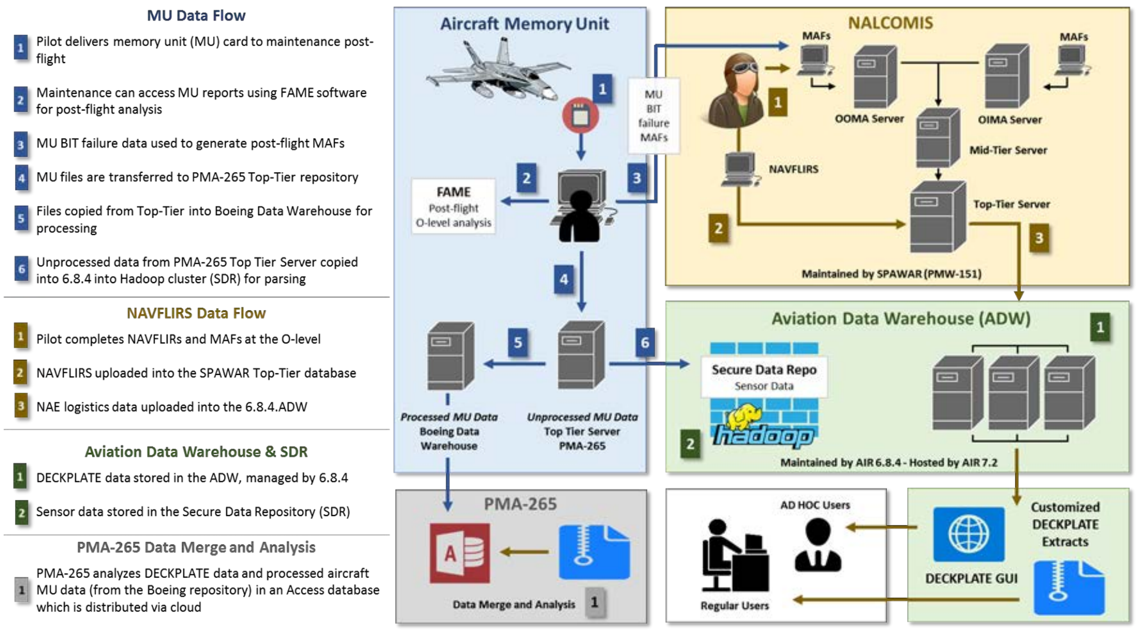

Figure 3.1. Fuel Analytics Dashboard Source Flow. Source: NAVAIR (2020).

NAVLFIR logs are filled out electronically by pilots upon completion of a sortie. The squadron uses the NAVFLIR to record aircraft flight time for maintenance and inspection purposes. The electronic record is ultimately uploaded to the same SDR as the MU data. The SHARP data path is not shown but follows a similar path as the NAVFLIR as it is also accessible using DECKPLATE (NAVAIR 2020).

3.1.2 Data Merging and Analysis

The three data sources are merged together by matching fields shared by each data set to form a single sortie observation. Primary methods of matching and merging are by algorithmic comparison of launch and land dates, launch and land locations, bureau number (BUNO), flight duration, incorporation of daylight savings time, and timezone offsets. In combining the data sets, a "fidelity" label is assigned as to the confidence at which the records match (i.e., "Low", "Avg", "Good", "High") (Deloitte Analytics Team 2020). Figure 3.2 illustrates the data processing flow Deloitte employed for ultimate visualization in the Dashboard.

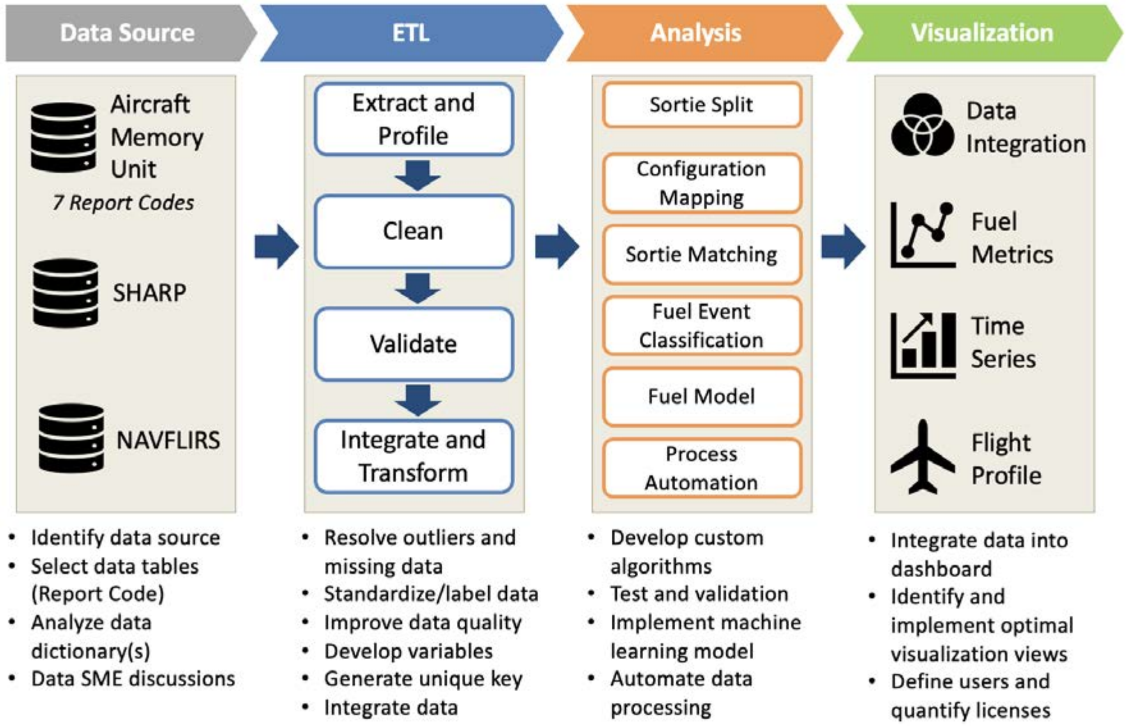

Figure 3.2. Fuel Analytics Dashboard Data Flow. Source: NAVAIR (2019).

Figure 3.2 shows how MU, SHARP, and NAVFLIR data were first extracted, cleaned, and standardized to prepare for integration. Next the resultant data set was analyzed to ensure appropriate sortie matching and event classification. Upon merging and matching, the data was then incorporated into the Dashboard through use of custom algorithms, testing and validation, machine learning tools, and, ultimately, automated processing (NAVAIR 2019). After this process functionality was applied for the ultimate visualization.

Merging the data sources incorporates detail into the data that might not be captured in a single data source with limited categories. If instead a single source were used, this would implicitly suggest that the used source is "correct" without consideration of other data sources. While the MU data set is considered "correct" when constructing the Fuel Conservation Analytics Dashboard output metrics, incorporating SHARP and NAVFLIR sources allows us to determine how well these other databases capture sortie information (Barnhill et al. 2020).

In executing the processing steps, care must be taken to consider vulnerabilities associated with incorrect logging or discrepancies in MU recordings. Section 3.1.1 describes the ideal situation that all databases match, even given the human error issues inevitable in the SHARP and NAVFLIR data sets. Barnhill et al. (2020) show otherwise; sorties are absent, observation field values are missing, and categories are mis-assigned. While the analysis conducted by Barnhill et al. (2020) does not focus on how to make the data collection process "better," the process is of vital importance, as missing or erroneous information in the beginning negatively affects output metrics for the end-user. Specifically, NAOEP must ensure a high capture rate of sorties across the NAE regardless of data source.

Additionally, how these data sets are merged is essential to the proper working of the Fuel Conservation Analytics Dashboard and any follow-on metric analysis. Assessment of how the data sources are combined and the specific mechanism of matching must be evaluated to ensure the sortie matching is as correct as possible. Incorporation of naval aviation expertise is also necessary to ensure broad generalizations do not inadvertently blur detail in the data merge, such as assuming all deployed squadrons fly similar missions or that Combatant Commander (CCDR) requirements are universal regardless of AOR.

3.1.3 Unique Sortie Data Fields

The unique sortie data set consists of 466,401 individual sorties matched among the MU, NAVFLIR, and SHARP data sets for three separate TMS platforms: F/A-18E, F/A-18F, and EA-18G aircraft. The data set is comprised of 79 fields taken from each data source or generated by comparing the databases (Deloitte Analytics Team 2019).

The unique sortie data set is composed of fields corresponding to data collected by aircraft sensors and recorded on the MU, as well as pilot post-flight logging information in SHARP and NAVFLIR databases. Most fields go unused aside for matching sorties when the data set is compiled. These fields, or variables, fall into three broad super-categories: labels that provide a naming convention for a particular sortie, "global" variables that apply to sorties regardless of data source, and source-specific variables that are unique to the original database.

Of the 79 variables, eight are considered "master" variables providing information applying to a sortie regardless of database source. These "master" variables are typically either a global variable or labels. There are an additional three fields indicating the fidelity of pairwise matching between each database. They provide little in the way of meaningful information for statistical output. Further, to verify that the unique BUNO assigned to an aircraft is correct, there are three indicator variables showing whether the BUNO is present in the particular database for that sortie. Again, this provides information to the analyst about the raw data but matters little for the Dashboard statistical output. The remaining 65 variables are database-specific aggregators. These variables provide information related to the sortie as it was captured in the data record.

Deloitte added several informational and categorical fields used for aggregation in the Fuel Conservation Analytics Dashboard. Of these, 60 are determined to provide meaningful information in statistical analysis when evaluating the Dashboard. Those 60 fields are taken from a data dictionary provided by Deloitte Analytics Team (2020) and are defined in Appendix A.1 in detail.

Ultimately, the Fuel Conservation Analytics Dashboard uses approximately 20 fields for aggregation. All fuel metric output is derived from recorded MU data only.

3.1.4 Dashboard Data Inclusion

While there are 466,401 sorties represented in the data set, only those that are also present in the MU database are used for input into the Dashboard. This results in 313,279 sorties being incorporated into Dashboard calculations while 153,122 sorties provide no input whatsoever. To determine which type of sorties will be included, we must examine the "master_matchstatus" variable. This variable is the primary indicator as to whether a sortie is included or not.

The unique sortie database was compiled in such a way that each sortie falls into one of seven categories depending on whether the sortie appeared in one or multiple databases. These categories are "MU Only," "NAVFLIRS Only," "SHARP Only," "MU-SHARP," "MU-NAVFLIRS," "NAVFLIRS-SHARP," and "Triple Match" (sortie appears in all three databases). The naming convention indicates in which databases the sortie appears. Table 3.1 shows the sortie count by data source category and platform.

Table 3.1. Sortie Count by Database Source Category and Platform.
Source: Barnhill et al. (2020)

Sortie Type	**Total by Category**	**F/A-18E**	**F/A-18F**	**EA-18G**
MU Only	10,149	3,167	2,159	4,823
SHARP Only	20,469	9,266	9,511	1,692
NAVFLIRS Only	26,737	5,250	3,572	17,915
MU-SHARP	20,207	11,368	7,613	1,226
MU-NAVFLIRS	55,241	7,900	3,572	41,030
NAVFLIRS-SHARP	105,916	55,892	39,249	10,775
Triple Match	227,682	120,835	88,335	18,512
Total Sorties	466,401	213,678	156,750	95,973

MU fuel metric are used exclusively for the Dashboard because these values are received directly from the aircraft sensors and are not subject to human error. While SHARP and NAVFLIR data provide information supporting sortie assessment, this information results from post-flight logging and is therefore reliant on aircrew memory.

While Table 3.1 indicates the primary categories used to determine whether observations are included in the Dashboard output, sorties may fall into a final category where the sorties were flown but a record of them does not appear in any database. While there is no mechanism to directly observe sorties in this final category, their presence is inferred by comparison to sortie counts in similar squadrons (i.e., by TMS). However, comparing similar squadrons across separate Carrier Air Wings (CVWs) requires the assumption that the squadrons are doing very similar operations given other observable characteristics.

Using only sorties containing MU information (which we will call “MU sorties”) for fuel metric calculation in the Fuel Conservation Dashboard results in a loss of approximately 1/3 of the total observed sortie count with each TMS experiencing a similar proportional loss. Table 3.2 provides a comparison between total sortie counts and MU sorties by TMS, together with the proportion of sorties that have MU data.

Table 3.2. MU Sortie Count by Platform.

Platform	Total Sorties	MU Sorties	Proportion
F/A-18E	213,678	143,270	0.670
F/A-18F	156,750	104,418	0.666
EA-18G	95,973	65,591	0.683
Total	466,401	313,279	0.671

While approximately 30% of MU data is unavailable, this loss is not spread evenly across squadrons or CVWs. Further, in several cases, squadrons appear to be missing sorties from all databases suggesting a higher percentage of missing MU sorties than can be inferred directly from the data available. These are sorties that fall into the eighth category of "no database."

3.1.5 Data Aggregation Categories

The Fuel Conservation Analytics Dashboard user is given the choice of 15 separate categorical variables as defined below. The user may choose a combination of these variables for aggregation.

Match Status: The sortie match status represents the type of match for each unique sortie. There are seven types of matches: "MU Only," "SHARP Only," "NAVFLIR Only," "NAVFLIRS-SHARP," "MU-SHARP," "MU-NAVFLIRS," and "Triple Match." The Dashboard only uses sorties with available MU data so only four match types are used: "MU Only," "MU-SHARP," "MU-NAVFLIRS," and "Triple Match."

TMS: The TMS of the particular aircraft. The Dashboard allows analysis from three different TMS: F/A-18E, F/A-18F, and EA-18G. The Dashboard uses TMS from the MU data set.

Fleet: Identifies the Type Commander (TYCOM) for an associated sortie. The Dashboard allows for five separate assignments of TYCOM: "PAC Navy," "LANT Navy," “CNARF Navy,” “NASC FS,” and for unassigned sorties, “NULL” is allowed. This category assignment is determined by the “master_command” field.

Airwing: Identifies the operational CVW, or aircraft carrier aviation complement, to which a squadron is assigned. Assignment is based on BUNO and Year/Month/Day of launch date. The Navy operates nine CVWs, each containing a of combination of aircraft types. A typical CVW consists of five jet aircraft squadrons: three F/A-18E squadrons, one F/A-18F squadron, and one EA-18G squadron. However, during the time frame analyzed several CVWs still incorporated F/A-18C aircraft instead of the newer F/A-18E with flight of the last F/A-18C occurring in 2019. The CVW assignment is taken from the “master_airwing” field.

Squadron: Associates a specific sortie to a USN squadron. Each operational squadron is assigned one TMS, while training and support squadrons often fly multiple TMS. Like CVWs, several squadrons still flew the F/A-18C aircraft during the time frame examined. Because the Dashboard only examines F/A-18E/F and EA-18G aircraft, data from squadrons that transitioned from F/A-18C to F/A-18E are not captured prior to the transition. Squadron assignment is taken from the MU data set.

BUNO: Identifies the BUNO associated with a specific aircraft. The BUNO assignment originates from the MU.

Sortie Length: Indicates sortie duration. This value is taken from the MU data set.

OFRP Phase: Identifies the particular phase of the Optimized Fleet Response Plan (OFRP) cycle in which a sortie occurs. Phases are defined and assigned as described in section B.1.5. OFRP assignment is derived from the “master_ofrp” field in the unique sortie data set.

Configuration: Indicates the type of external stores an aircraft is carrying. F/A-18E/F and EA-18G register 11 pylons where stores may be placed. There are 15,972 separate configurations given the different store or weapon combinations across TMS. This category is recorded in the “configuration” field in the MU data set.

Mission: Sortie mission is derived from SHARP data. SHARP is the only data source that assigns a specific mission to a sortie. There are 33 mission options in the data across TMS. If a SHARP record does not exist, mission is assigned a "NULL" value. "NULL" mission assignment is prevalent across all TMS but is most problematic for EA-18G where only 14% of sorties possess a mission assignment. Further, of the available EA-18G missions, approximately 50% show fewer than 30 sorties per mission type. F/A-18E and F/A-18F sorties fare better with regard to mission assignment. F/A-18E sorties show a mission assignment rate of 82% and of those approximately 90% exhibit more than 30 observations. F/A-18F sorties show a rate of approximately 80% with 83% showing more than 30 sorties.

Total Mission Requirement (TMR) Code: The TMR code is a three-character alphanumeric indicator that defines the intended sortie purpose. The Dashboard uses 173 different TMR codes for aggregation with those sorties with unassigned TMR given a "NULL" value. The TMR is closely related to the sortie mission assignment because several missions are grouped within a single TMR category. The TMR may be derived from either the NAVFLIR or SHARP record; however, the Dashboard exclusively uses the NAVFLIR record when the NAVFLIR is present. While TMRs in the NAVFLIR are valid, TMR mismatch rate between NAVFLIR and SHARP is approximately 26% suggesting some uncertainty as to the correct TMR.

Fuel Event: Fuel events on a sortie come in three types: "fuel dump," "fuel give," or "fuel take." Additionally, the Dashboard identifies whether a fuel dump event occurs in the first half of a sortie. This is derived from the MU data set.

Flight Event: A flight event comes in three types: a single engine event, arresting hook event, or canopy/ladder event. A single engine event is defined as when an aircraft has only one engine operating. An arresting hook event indicates whether an arresting hook is lowered. The canopy/ladder event indicates whether the aircraft canopy is open and/or the ladder is down.

Year: Indicates the year in which a sortie took place. If a squadron transitioned from a TMS not included in the data to one that is (e.g., F/A-18C to F/A-18E), no information is present prior to the transition.

Month: Indicates the month in which the sortie took place.

3.1.6 Fuel Metric Determination

Dashboard fuel output metrics are derived from the eight fuel-specific fields from the MU data. However, as observed in Appendix A.1, eight fuel fields in the SHARP data mirror those similarly named categories in the MU data. Because MU data is derived directly from the aircraft and SHARP data is subject to human error, MU data is considered as the benchmark for fuel metric calculations.

Fuel Used: Fuel used is defined as the total fuel used (lbs) over an entire sortie. The difference between starting and ending fuel is combined with any fuel that is taken, given, or dumped in flight. Average fuel used is also calculated across sorties by any category chosen by the user (e.g., by "Airwing," "Squadron," or "TMR").

Fuel Burn: Fuel burned is defined as the total fuel burned (lbs) during a flight. This is calculated by measuring only the amount of fuel consumed by the engine during flight. Average fuel burned is also calculated across sorties by any category chosen by the user.

Fuel Burn Rate: Fuel burn rate is calculated by dividing the total fuel burned by the sortie duration.

Fuel Dump: Total fuel dumped is calculated for each chosen category by measuring the total fuel released from the aircraft where a dump event occurs. In addition, average fuel dumped is calculated across sorties where a dump event occurs by chosen category or categories. Differentiating between fuel dumped and fuel given in normal flight operations required a supplemental sensor flag indicating a dumping event because the pattern of dumping fuel is similar to giving fuel.

Fuel Give: Total fuel given is calculated by measuring the total fuel an aircraft gives to other aircraft during the giving aircraft's sortie for any given aggregation category or collection of aggregation categories. Average fuel given is also calculated by any category chosen by the user. As stated, differentiating between dumping and giving fuel proved challenging because of similar quantity change patterns. In addition to the change pattern for a give event, a "give" indicator is captured by the MU. Metrics are only calculated for those sorties in which a give event took place.

Fuel Take: Total fuel taken is calculated by measuring the fuel an aircraft receives in flight. Average fuel taken is computed similarly to other output metrics. This is indicated in two ways, a measurable increase in fuel level during the flight and a "take" indicator captured by the MU. Metrics are only calculated for those sorties where a take event took place.

Total Sorties: The total sortie variables count the number of sorties specific to the aggregation category chosen by the user.

Flight Hours: Total and average flight hours are calculated and displayed for user-chosen aggregation categories.

3.2 Data Preparation and Methodology

This section describes techniques for data cleaning, variable construction, and methods of analysis used in this thesis for the unique sortie data set.

3.2.1 Data Cleaning

Meaningful fuel consumption analysis requires understanding the nuances of naval aviation. While the original unique sortie data set is rich, it does not capture these nuances well. In order to capture sortie-specific characteristics more effectively, several new variables were constructed. These serve to incorporate information from sorties that do not possess MU data as well as to correct mistaken assumptions made in the construction of certain data fields in the original unique sortie data set.

Duration: Sortie duration is determined by subtracting the launch date-time from the land date-time and converting it to hour units.

Mon: This field is a surrogate for the month of a particular year that a sortie took place. It is derived from the launch date and comes in the form of MM (e.g., “01” is the month of January).

Year: Similar to month determination, the “Year” field captures the year that the sortie occurred. The year is only captured for basic analysis and not for prediction modeling.

TMS: A new TMS field is constructed to include TMS assignment from non-MU sorties. When more than one record is present there is little disagreement between the three records. Between NAVFLIR and SHARP data there are only a total of 198 instances of mismatched TMS and of those instances, 55% are from training squadrons and only one instance appears to occur for operational squadrons while deployed.

Airwing: The squadron CVW assignment in the data set is based on the “master_airwing” field. To correct mis-assigned sorties or unassigned sorties a new field of “Airwing” is constructed. Of sorties assigned to a CVW on deployment, there is a 99.7% match between the old and new field. However, 4,982 sorties are unassigned to CVWs using the “master_airwing” field. To correct for this, the assigned squadron is examined and if the launch date falls within a known deployment time frame, the sortie is assigned to the correct CVW. Table 3.3 shows the comparison between the two fields. The “master_airwing” results are shown in rows while the newer “Airwing” category values are represented in the columns.

Table 3.3. Sortie Comparison Between Original and Constructed Airwing Fields.

		Constructed "Airwing" Category								
		CVW-1	CVW-11	CVW-17	CVW-2	CVW-3	CVW-5	CVW-7	CVW-8	CVW-9
"master_airwing" Category	None	2,961	18	129	21	0	1,345	458	8	42
	CVW-1	5,221	0	0	0	0	3	0	0	0
	CVW-11	0	5,213	0	0	0	2	0	0	1
	CVW-17	0	1	5,989	0	0	1	0	0	0
	CVW-2	0	86	0	7,014	0	1	0	0	1
	CVW-3	84	0	0	0	7,208	0	0	0	0
	CVW-5	25	0	0	0	0	28,083	0	0	0
	CVW-7	0	0	0	1	0	0	17,012	0	1
	CVW-8	7	0	0	0	0	1	0	6,562	0
	CVW-9	0	31	0	0	0	0	0	0	17,668

Squadron: Squadron assignment is generally consistent across MU, NAVFLIR, and SHARP when present. Like the CVW field, the Squadron field ensures a squadron assignment for records that possess no MU data.

This field also serves to rectify mismatches between squadron fields reported in separate flight databases. For example, of the roughly 73,500 deployed sorties identified by OFRP, F/A-18E squadrons exhibit the largest number of mismatched squadron assignments at approximately 130. F/A-18F squadrons show approximately 30 mismatches and EA-18G squadrons only six. Most of these mismatches occur between MU and SHARP or NAVFLIR and SHARP records.

MU and NAVFLIR squadron assignments largely match as these are based on maintenance-related records belonging to the squadron that owns the aircraft. For SHARP records, the pilot will record the flight with his or her squadron for flight-hour accountability and personal logbook maintenance. Mis-assignment can occur if a training command like Naval Air Warfare Development Center (NAWDC), fleet replacement squadrons (FRS) (VFA-106, VFA-122, or VAQ-129), or the associated weapons schools (Strike Fighter Weapons School Pacific (SFWSP) and Strike Fighter Weapons School Atlantic (SFWSL)) utilize operational squadron aircraft. While the maintenance records reflect the squadron that owns the aircraft, the SHARP record reflects the command that flew the aircraft. This is done for flight hour management purposes. If aircrew from one command uses another squadron's aircraft, the flight hours will be applied to the using command.

BUNO: The BUNO is provided in each of the three records. This variable is filled based on precedence of records available. MU is considered correct if present, SHARP is considered correct if MU is not available, and finally NAVFLIR BUNO is used if the other two are unavailable.

Deployed: This variable is used in lieu of the "master_ofrp" to determine whether squadrons are deployed or not. Appendix B.1 discusses USN policy regarding OFRP assignment for any deployable unit. The "master_ofrp" variable exhibits obvious errors related to phase assignment where in many cases the OFRP phase is left out or squadrons are assigned multiple phases for a given month. Because of the problems with the original variable, it is worth more discussion below.

Every operational activity is always in some phase of the OFRP as discussed in section B.1.5. OFRP is an important attribute as it informs decision makers at the squadron-level and above about fuel consumption during each readiness phase. However, it is clear that the data processing methodology does not assign OFRP phase correctly (Barnhill et al. 2020). This becomes apparent, for example, in the case of CVW-5, which is the only CVW that is part of Forward Deployed Naval Forces (FDNF). Because it is part of FDNF, CVW-5 squadrons may only be assigned to the "Deploy" and "Sustain" phases. This fact was previously confirmed by the author in direct discussions with Commander, Naval Air Forces (CNAF) representatives. CVW-5 consists of five squadrons that fly F/A-18 or EA-18G aircraft (Barnhill et al. 2020).

To examine OFRP issues further, we offer VFA-102 as an example. VFA-102 is one of the squadrons assigned to CVW-5 and flies the F/A-18F aircraft. Figure 3.3 shows sortie OFRP assignment by month from 2016 to 2019. First, we observe that even though the sorties should only be assigned the "Deploy" or "Sustain" phase, VFA-102 is not consistently assigned the correct phase until the middle of 2017. In addition, we note that several months show no sorties at all. For example, Figure 3.3 shows that VFA-102 records no sorties with OFRP assignment for the months of June and July 2017, as well as October of 2019 (i.e., “2017-06,” “2017-07,” “2019-10”). While it is reasonable to expect low numbers of sorties in a month, as in April of 2019 (“2019-04”), the number of sorties will never be zero. This is because any naval aircraft not flown for 30 days requires an in-depth squadron-conducted inspection and follow-on maintenance flight. To avoid these maintenance requirements, squadrons need only fly the aircraft one time in a 30-day period which is common practice. We also observe in Figure 3.3 that multiple OFRP phases are assigned in a single month. For VFA-102, this occurred in November of 2016, January of 2017, and April of 2018 (i.e., “2016-11,” “2017-01,” “2018-04”). OFRP phase is assigned on a monthly basis and, if the squadron is due to proceed to the next phase, the subsequent phase does not start until the beginning of the following month.

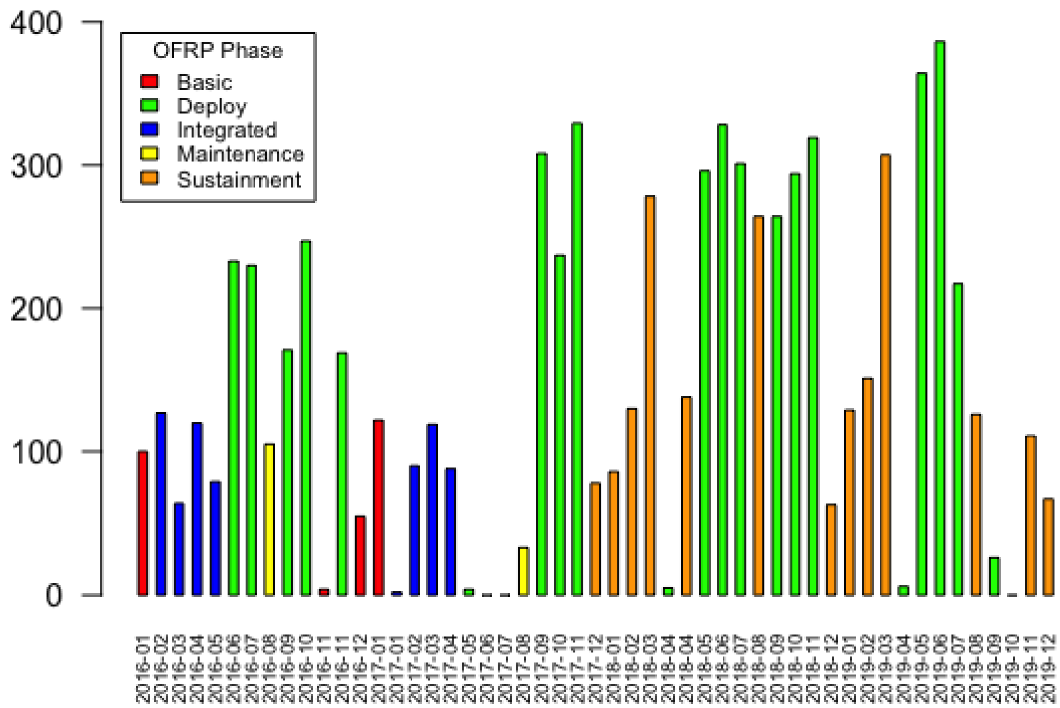

Figure 3.3. VFA-102 (F/A-18F) Monthly OFRP Assignment.

While it is easy to see in the case of VFA-102 that OFRP phases are mis-assigned, it is not as straightforward to determine, in general, whether OFRP assignments in the sortie data are accurate. This is because non-FDNF squadrons follow the typical OFRP cycle. However, in the Dashboard data, non-FDNF squadrons are routinely assigned multiple OFRP phases within a single month, appear to conduct phases out of sequence, or are missing phases altogether. This calls into question the reliability of the "master_ofrp" field generally (Barnhill et al. 2020).

We replace the "master_ofrp" field with the "Deployed" field where each sortie is assigned a "Deploy" or "Not Deployed" label. Whether a squadron is deployed or not is determined by the sortie launch date and comparing it to known CVW deployment dates. If the launch date took place between the beginning and end of a deployment date, the sortie is considered a deployed sortie. Unfortunately, this does not capture all OFRP phases, but it does allow us to determine which sorties are deployed since we are relying on known deployment date ranges instead of assuming deployment labels are correct in the original OFRP field.

Route: The new "Route" field works to fix inconsistencies in launch and land locations while still using original "master_route" terminology. It gives greater weight to the SHARP rather than the NAVFLIR record when present but defaults to the NAVFLIR when the SHARP is not available. The new field re-categorizes 9,219 sorties.

Time of Day (TOD): This field determines the time of day the sortie takes place. NAVFLIR data possesses fields for landing codes. Using these codes we can determine at what time of day the landings took place. While it is possible for sorties to cross over from daytime to nighttime (or vice versa), the field is considered correct enough for analysis purposes. Table 3.4 exhibits landing code assignments based on time of day.

Table 3.4. Sortie Time of Day Assignment Based on NAVFLIR Codes.

Landing Code	Time of Day
1,2,3,4,5,6,7	Day
A,B,C,D,E,F,G,N,P,Q	Night

TMR: Sortie TMR assignment is originally determined by a TMR code in either the NAVFLIR or SHARP data. If both are present, the Dashboard prioritizes the NAVFLIR TMR assignment. However, when both records are present and valid (i.e., a three-character TMR code is assigned in each record) 26% of TMR codes are mismatched between the two records. The SHARP record is more important to the aircrew because it populates the pilot logbook, thus, we believe the SHARP log is more trustworthy than the NAVFLIR. The new field assigns a TMR code based first on SHARP if both are present. If neither record is present or records are present but there is no TMR recorded, the TMR is assigned as "Unknown."

In addition to prioritizing the SHARP TMR record, the TMR assignment is truncated to two characters. This reduces the number of TMR categories so the categories can be used more effectively in predictive modeling. Specifically, our implementation of random forests can only accommodate a maximum of 53 levels in any categorical variable.

Mission: This field is constructed to capture mission assignment for any particular sortie in the data set. This entry only exists in the SHARP record so if the SHARP record is not present, it is given an "Unknown" value.

Configuration: Configuration is defined as the type of external stores carried by an aircraft on one or more of eleven pylons on the F/A-18E/F or EA-18G. The data exhibit this configuration by showing what each pylon is carrying. In Section 3.1.3 we show there are 10,200 configurations for F/A-18E aircraft, 6,988 configurations for F/A-18F aircraft, and 1,580 configurations for EA-18G. In addition to the original "configuration" field, there is a unique configuration identifier, or "unique_config_id," field. This field assigns a numeric label to a particular configuration. This identifier indicates a category where several different configurations are assigned. Nonetheless, there are still over 12,000 separate identifiers across all TMS.

Instead of differentiating among specific configurations, a new "configuration" field is constructed to generalize over the data set. It shows the number of external stores that an aircraft requires for a sortie reducing the number of configurations from over

14,000 to 12. This mechanism assumes that the only thing affecting fuel consumption (with regard to external stores) is the number of stores carried. This study does not suggest that this is exactly the case, but we are prepared to assume that the predominant effect on fuel consumption is increased weight rather than the drag associated with different externally carried stores.

3.2.2 Variable Selection

The original data set consists of 79 variables from the three different data sets. In Section 3.2.1 we showed construction of new variables that amalgamate information from the three data sources, increase the usefulness of non-MU data, and incorporate information from outside sources to provide detail in the data. This section will identify variables seen as important to both predictive modeling and general results.

In the Fuel Conservation Analytics Dashboard, MU data is considered "truth," so all variable choices are meant to either drive results related to, or predictions about, MU fuel consumption. Variables have been constructed to merge data from the three separate data sets. In addition, redundancies and mismatches in the original data set are reduced. This allows us to discard some original variables while still being able to glean meaningful information about fuel consumption.

Where SHARP and NAVFLIR variables are redundant SHARP is given greater weight due to its importance to the aircrew. However, NAVFLIR information is useful in generating variables that are not present in the original data set, like the time of day the flight occurred (i.e., "TOD"). All MU variables related to fuel consumption were considered to be response variables. Table 3.5 shows 20 predictor variables used for this analysis.

Table 3.5. Original and Generated Predictor Variables.

Variable	**Source**
TMS	Generated
Airwing	Generated
Squadron	Generated
Route	Generated
Duration	Generated
Mission	SHARP
TMR	Generated
nav_approachcode	NAVFLIR
sharp_fuelburned	SHARP
sharp_fuelburnrate	SHARP
sharp_sortieduration	SHARP
sharp_fuelstart	SHARP
sharp_fuelend	SHARP
sharp_fuelused	SHARP
sharp_fuelgiven	SHARP
sharp_fueltaken	SHARP
sharp_fueldumped	SHARP
TOD	Generated
Mon	Generated
configuration	Generated

3.3 Methodology

In this section we will discuss the various methods used to correct missing and mis-assigned data for deployed sorties. Additionally, this section discusses how data prediction models are constructed and compared. All analysis for this study was conducted on a MacBook Pro with Intel i5 Processor (4 x 2.4 GHz) and 8 GB 2133 MHz LPDDR3 RAM.

3.3.1 Assessment of Missing Data

Of primary concern for this project is the amount of missing MU data from the original data set. A number of hypothesis tests are utilized to determine any sort of pattern to the missing data. Additionally, the study considers ways to potentially compare deployments via non-parametric hypothesis testing.

3.3.2 Deployment Sorties

This study only considers deployed sorties in the analysis. While the data set assigns a “Deploy” label based on OFRP phase, OFRP assignments are determined to be incorrect. Of the approximately 102,000 deployed sorties identified using the “master_ofrp” field, only 73,533 are captured in the Dashboard analysis. To correct for this, this study uses the newly constructed “Deployed” field which identifies 105,197 sorties and imputed any missing numeric values to maximize the number of sorties included. Ultimately, the prediction models utilize 105,192 of the total sorties identified once erroneous observations were removed.

3.3.3 Comparisons and Trend Analysis using Data with Missing Values

One recommendation of the study conducted by Barnhill et al. (2020), was to assess fuel consumption by deployment destination or AOR. The unique sortie data set does not indicate the destination of deploying units. This study addressed this by using sources outside of the data set to determine AOR destination. Categorical variables were constructed to indicate destination and basic statistical techniques are used to determine not only if there is apparent correlation between deployment destination and fuel consumption but what type of sorties seem to cause greater fuel consumption in those locations.

3.3.4 Predicting MU Fuel Consumption

While each MU fuel variable provides some insight, our primary concern is “mu_fuelused” because of its comprehensive definition. Several methods are employed to predict “mu_fuelused” in order to assess the predictive power of the original and constructed variables despite missing data. In this section we describe different methods used to construct prediction models. This portion of the study first focuses on complete-case analysis of

SHARP and MU sorties and then considers methods of imputation to increase the number of observations that may be used in predicting fuel consumption.

For prediction, we utilized the random forest algorithm provided in the "ranger" package in the R programming language (Wright and Ziegler 2017). The random forest algorithm provides a reliable method for prediction for large data sets with many predictor variables. A random forest model is an ensemble model where the predicted values are average predicted values taken over an ensemble (forest) of trees. See (James et al. 2017) for a detailed description of random forests and their use.

Complete-Case Analysis

Complete-case analysis involves using only those observations with no missing data. For this analysis, this included the 45,824 observations from all TMS aircraft with complete SHARP and MU information. Using this data set, we utilize SHARP fields to predict "mu_fuelused" fitting two different models.

The first model involves using only "sharp_fuelused" to determine how well it predicted "mu_fuelused." The second model uses 13 SHARP fields again predicting "mu_fuelused."

Notably, since we only use those observations with complete observations (or cases) using the results of these model fits to fill in missing "mu_fuelused" for those observations with just SHARP information runs the risk of bias. This is because we cannot assume the "complete-case" data subset serves as a representative sample of the larger sortie population.

Mean Value Imputation

A third model for predicting "mu_fuelused" is constructed by imputing missing numeric variable values (for both predictor and response variables) for deployed sorties with the mean of each respective variable's values taken over sorties where those values are present. Mean imputation uses information from observed sorties, regardless of matching label, to fill in missing fields and thereby increase the number of useable observations to 105,192 across TMS. While mean value imputation allows us to include observations not used in the complete-case models, we run the risk of underestimating variance of the output. This is because mean imputation assumes that missing values are similar to those that are observed which discounts other information like sortie duration which affects fuel consumption the most. Unlike complete-case analysis, we use the 20 constructed and original predictors shown in Table 3.5.

Multiple Imputation

MI offers the opportunity to impute values for both categorical and numeric variables for use in statistical inference and predictive modeling. While there are many MI algorithms, we use the “missForest” package in R (Stekhoven and Buhlmann 2011).

A benefit of MI is that not only does it populate missing values of the response variables, but it will also populate values in the predictor variables, as well. This allows the analysis to increase the number of observations used to train any prediction model developed. In order to test the efficacy of MI on the unique sortie data set the random forest model is employed.

Prediction Models

This study constructed and analyzed four random forest models using a combination of original and generated predictors to assess how well MU fuel consumption can be predicted despite missing information. Two of the models used a data set constructed of complete observations where both SHARP and MU values were present and only employed original SHARP variables for use as predictors. The remaining two used data sets constructed of the same original SHARP variables but included generated variables as well (see Section 3.5). In addition, these two models employed mean value and MI, respectively, to fill missing numeric values to increase the number of available observations. The four models are described in detail below.

1. **Model 1:** MU fuel used predicted by SHARP fuel used (“sharp_fuelused”) based on complete cases only.
2. **Model 2:** MU fuel used predicted by complete cases for 13 SHARP-specific categories.
3. **Model 3:** MU fuel used predicted by 20 predictors, combining generated, SHARP, and NAVFLIR variables and based on all deployed sorties. Missing numeric values were imputed using variable mean values.

4. **Model 4:** MU fuel used predicted by 20 predictors based on all deployed sorties where missing values in numeric variables were imputed using MI, specifically, the "missForest" algorithm.

To evaluate the models, we divided the data into training and test sets. Models 1 and 2 used a train/test split of 90%/10% while models 3 and 4 used an 85%/15% split. For models 2, 3, and 4 we used five-fold cross validation to select the "MTRY" random forest hyper-parameter. The "MTRY" hyper-parameter is the number of predictor variables to examine for each split as the random forest algorithm builds each tree. All models utilized 300 trees to build the associated random forest. Table 3.6 shows the type of predictors, sample sizes for each model, training and test set sizes, and "MTRY" hyper-parameter value.

Table 3.6. Prediction Model Construct.

Model	**Predictors**	**Observations Used**	**Train Set**	**Test Set**	**MTRY**
Model 1	SHARP Fuel Used	45,824	41,242	4,582	1
Model 2	Multiple SHARP	45,824	41,242	4,582	5
Model 3	SHARP, NAVFLIR, and Generated	105,192	89,413	15,779	13
Model 4	SHARP, NAVFLIR, and Generated	105,192	89,413	15,779	12

Of these models, model 4 ("multiple imputation") exhibited the best performance overall though at the expense of computational time. MI pre-processing increased available observations for model training approximately 130% relative to non-imputed models.

CHAPTER 4:
Analysis

4.1 Introduction

In this chapter we assess different elements of the unique sortie data set based on the preparation and cleaning outlined in Chapter 3. Primary focus is given to analyzing consistency of missing data for deployed sorties across aggregation categories and the effects of missing information on inference. Deployment destination (or AOR) is then examined to determine effects of location on fuel consumption along with flight type. Finally, four random forest models are assessed by their ability to predict "mu_fuelused" based on a combination of original and/or generated variables and imputation methods to increase the number of observations available for analysis.

4.2 Missing Data Overview

This study considers missing data in the unique sortie data set. Missing information in the data set presents itself in three ways: 1) Sorties without MU information; 2) Sorties missing in their entirety; 3) Aggregation fields with missing values (Barnhill et al. 2020).

4.2.1 Missing Data Trends

The categories of missing data are defined in Section 2.1.4. Determining the pattern of missing data requires knowing something about where the data originated. While we might be able to ascertain general trends in missing data (e.g., a lower proportion of MU data from 2017 than in the other years in the time frame) we cannot definitively know why or how this occurred. Further, if missing information is inconsistent or exists at only certain aggregation levels, determining a broad pattern of missing data other than MNAR is very difficult. For these reasons, the unique sortie data set will be considered MNAR, meaning that the probability of data being missing is based on both observed and unobserved data. This means that the missing data, regardless of the type of missing information, is considered relevant and cannot simply be ignored in statistical analysis.

4.2.2 Sorties without MU data

Previous analysis focused primarily on operational deployed sorties as those were determined to be of most interest to NAOEP and in the best interest of the NAE. These sorties were aggregated by platform, CVW, and squadron. Deployment was initially determined by an assignment of “Deploy” in the “master_ofrp” field but OFRP assignment exhibited substantial issues. Because of these problems, a “Deploy” status is determined using the newly generated category based on date of deployment for a particular CVW (defined in Section 3.2.1).

As mentioned, the Dashboard uses only MU information to calculate the output metric. Although approximately 1/3 of observed sorties are missing MU data overall, the proportion missing within squadrons or CVWs varies over time. Further, when filtering sorties to levels below “Airwing” and “Squadron” like the category of “TMR,” the proportion of sorties missing MU information or TMR assignment (from SHARP) is in some instances even larger (Barnhill et al. 2020).

To analyze this phenomenon, Barnhill et al. (2020) examined two deployments by different CVWs. CVW deployments are chosen because this is the highest level of aggregation thought to be useful to any potential stakeholder. The two deployments examined were the 2016 CVW-3 deployment and the 2017 CVW-8 deployment. Figures 4.1 and 4.2 display the total sorties flown regardless of platform. It further breaks down those sorties into those with associated MU data and those without.

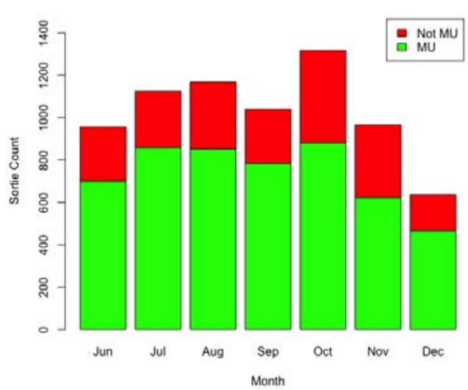

Figure 4.1. CVW-8 2017 Deployment.

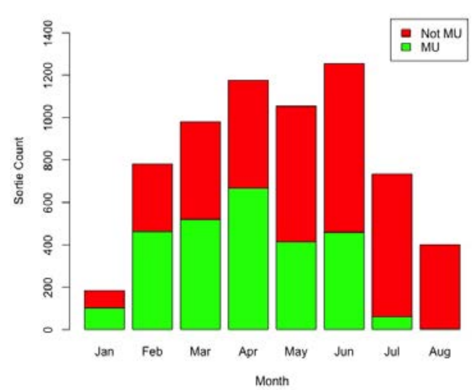

Figure 4.2. CVW-3 2016 Deployment.

These figures show that missing data is inconsistent between CVWs even though both deployed to the same AOR. MU data is present in 72% of available sorties from the CVW-3 deployment and 41% of available sorties from CVW-8. Comparisons of other CVW deployment MU proportions yield similar results suggesting that the rate at which sorties are missing MU is unpredictable (Barnhill et al. 2020).

Figures 4.3, 4.4 and 4.5 generated by Barnhill et al. (2020) show the number of sorties by squadron for each TMS. While these plots ignore the number of times each squadron deployed, discrepancies between squadrons are apparent. This gives further evidence that there is no consistent pattern to missing sorties for squadrons of like TMS.

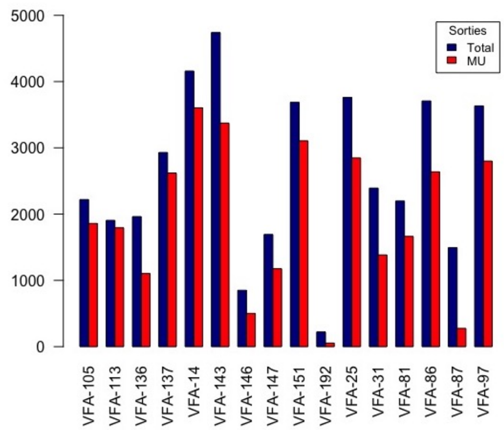

Figure 4.3. MU vs. Non-MU F/A-18E Sorties.

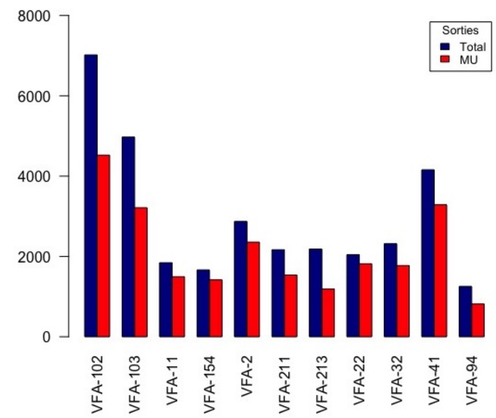

Figure 4.4. MU vs. Non-MU F/A-18F Sorties.

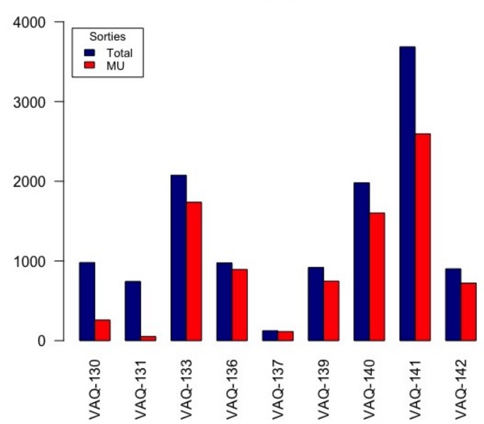

Figure 4.5. MU vs. Non-MU EA-18G Sorties.

Sorties Missing in their Entirety

This section reproduces analysis conducted by Barnhill et al. (2020) for NAOEP and addresses sorties that are completely missing from any of the three data sources. Referencing Figures 4.3, 4.4 and 4.5, there are several examples of what looks like missing sorties regardless of platform. For each platform on between a six- to eight-month deployment, one may expect to see approximately 2,000 sorties for F/A-18E/F squadrons and between 800 and 1,300 sorties for EA-18G squadrons (Barnhill et al. 2020).

Among F/A-18E squadrons in Figure 4.3 VFA-192 shows the oddest pattern. This squadron was deployed for a total of about 12 months during the four-year time frame of the data, so thousands of flights are expected in the data. In fact only 247 deployed sorties are recorded for this squadron. Further, VFA-192 recorded no flights in any data source for the CVW-2 2017 deployment. Its "sister squadron" VFA-137 (F/A-18E) recorded a total of 1,738 sorties giving an indication of how many sorties one would expect to observe. VFA-146 is another F/A-18E squadron suspected of missing sorties. For F/A-18F squadrons shown in Figure 4.4, VFA-94 exhibits a low sortie count while in Figure 4.5 VAQ-137 appears undercounted with only 125 deployed sorties recorded despite approximately seven months of deployment (Barnhill et al. 2020).

Considering Proportions of MU Data

While it is instructive to look at the data in large aggregation levels like those exhibited in Figures 4.3, 4.4, and 4.5, doing so discounts issues that may arise between different deployments or years. Figures 4.6, 4.7, and 4.8 show the proportion of MU data by deployment for the three TMS. Figure labeling uses the year of deployment, the specific number of deployment in that year (e.g., 1st, 2nd), and associated CVW (e.g., "16-1_7" means the first deployment for CVW-7 in 2016). For F/A-18F and EA-18G squadrons, there is typically only one squadron per CVW for each deployment while there are usually three F/A-18E squadrons per deployed CVW. For this reason, F/A-18E squadron distributions of MU proportions are illustrated using boxplots.

Figure 4.6 shows wide variability in MU proportions during many deployments while others appear relatively similar among the two or three F/A-18E squadrons assigned to a CVW for a particular deployment. Of note, we see that the 2019 CVW-1 deployment (i.e., “19-1_1”) shows no MU sorties for the deployment. Other examples of anomalies are clearly observable. The 2017 CVW-2 deployment (i.e., “17-1_2”) exhibits a wide variation in proportion ranging from 87.4% of MU data (of the total) to zero. In this case, CVW-2 consisted of two F/A-18E squadrons and one F/A-18C as shown in Table B.2. As previously identified, one F/A-18E squadron, VFA-192, did not register any sorties during the deployment.

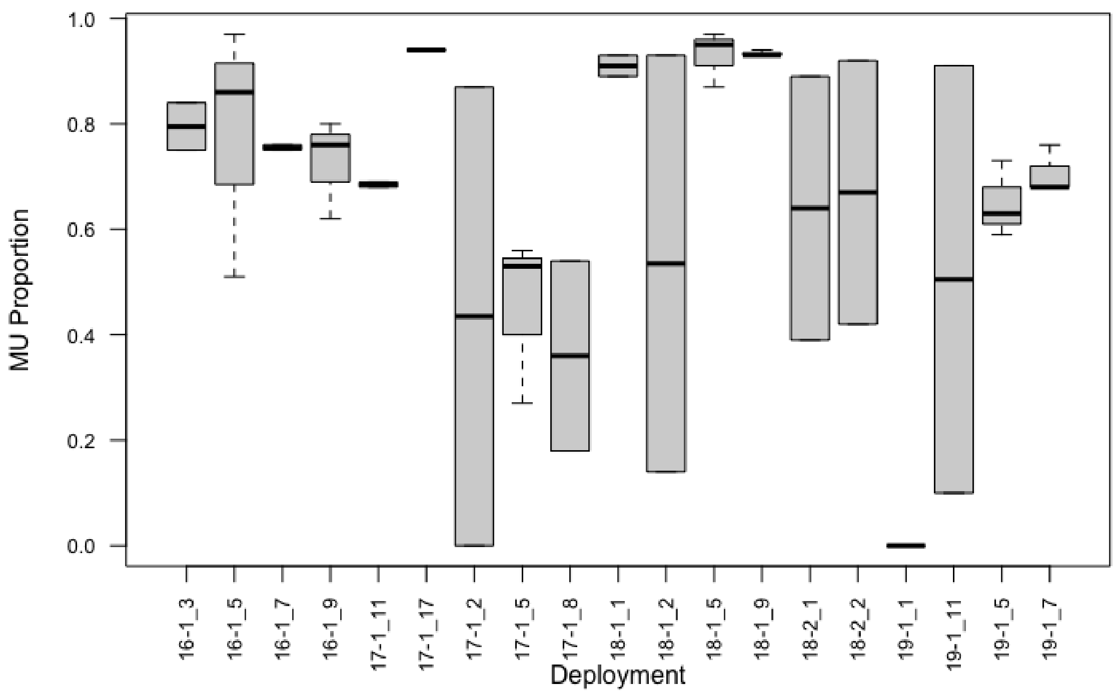

Figure 4.6. Proportion Distribution of F/A-18E MU Sorties by Deployment.

While Figure 4.6 exhibits differences in MU data proportions, it does not provide complete information. In some cases, a high proportion of MU sorties is observed but relatively few sorties were captured. This is pertinent to the CVW-11 2019 deployment where a proportion of 10.1% is present for VFA-146 and 90.9% for VFA-31. While the proportion for VFA-146 is definitely low, only 129 sorties were flown by the squadron. Similarly, for VFA-31, only 241 sorties were captured. This is likely due to the fact that CVW-11 deployed from the end of 2019 into 2020 and there are no sorties in the data set past the end of 2019 (see

Table C.2). Nonetheless, having a high proportion of MU sorties (of the total) does not mean a squadron's data is complete.

While F/A-18E squadrons as a group are missing large amounts of MU data, we have the ability to compare CVW "sister squadrons" which provides insight if some appear to be missing MU sorties. Using the median number of sorties by deployment of 1,691, we gain information on how many sorties are typically observed during a deployment. The median is preferred to account for sorties that are not present in any of the databases since we cannot determine how many of these sorties actually exist. Additionally, using the median lessens the effect of outliers in sortie counts, specifically due to sortie undercount.

F/A-18F squadrons exhibit a higher overall proportion of available MU data than both F/A-18E and EA-18G squadrons at 76%. Figure 4.7 shows the proportion of F/A-18F sorties with MU data by deployment. The median sortie count for each deployment shows approximately 1,500 sorties per deployment. Unlike F/A-18E sorties, only one F/A-18F squadron exists per CVW in most cases (exceptions include are the "17-1_17," "18-1_1," "18-2_1," and "19-1_1" deployments). Similar proportion comparison between F/A-18F squadrons is not considered because there is only one squadron per deployment. Assuming there are similar demands on two separate deployments, even for the same squadron, is unwise.

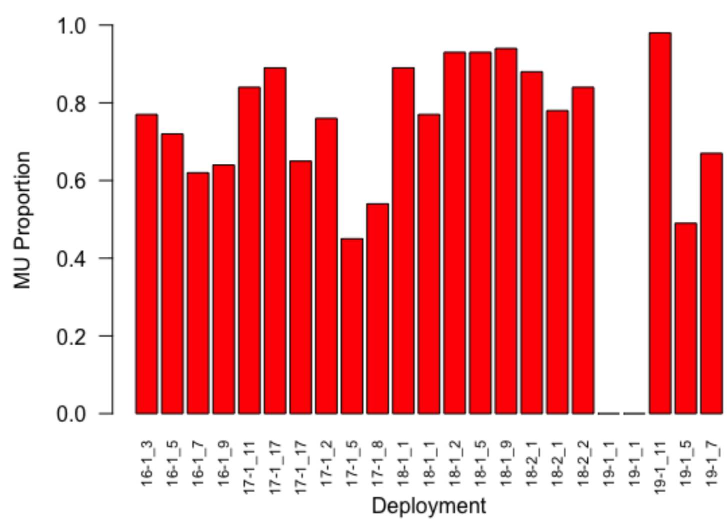

Figure 4.7. Proportion of F/A-18F MU Sorties by Deployment.

EA-18G squadrons suffer the most from missing MU data largely because there are so few squadrons (nine total) compared to the other TMS. During the 19 deployments, the median sortie count for a squadron was 739 sorties. The average proportion of MU sorties was approximately 71%, with a standard deviation of 28% indicating wide variation in MU proportions. Figure 4.8 illustrates proportions of EA-18G MU sorties by deployment.

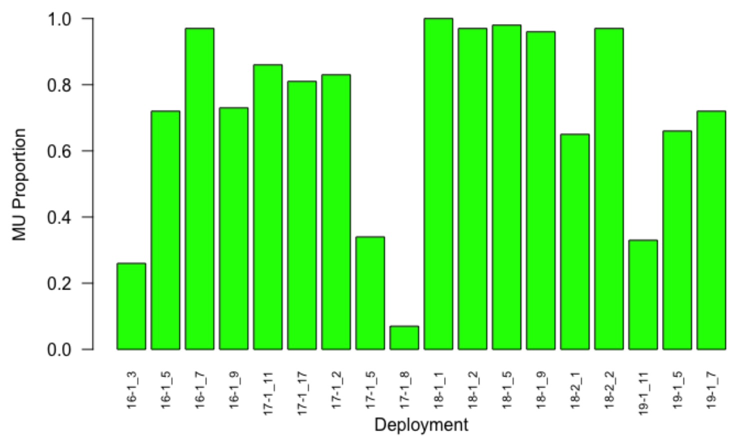

Figure 4.8. Proportion of EA-18G MU Sorties by Deployment.

Several EA-18G squadrons indicate large amounts of missing MU data. Among these squadrons, VAQ-130 and VAQ-131 show a comparatively high number of total deployed sorties when compared to the TMS median (739 sorties per deployment) but an exceptionally low proportion of MU sorties for these specific squadrons, 26% for VAQ-130 and 7% for VAQ-131. VAQ-141 also exhibits lower proportions but these vary, depending on the year of deployment. The squadron's proportion varies from a low of 34% in 2017 to a high of 98% in 2018 with 2016 and 2019 showing 72% and 66% respectively. The reasons for the proportion differences, especially for deployments involving the same squadron, are unclear. Table 4.1 shows the number of total and MU sorties by deployment year and squadron.

Table 4.1. EA-18G Sortie Count.

Squadron	Year	Total Sorties	MU Sorties	Proportion
VAQ-130	2016	980	258	0.26
VAQ-131	2017	742	52	0.07
VAQ-137	2018	125	114	.91
VAQ-141	2016	724	524	0.72
VAQ-141	2017	735	250	0.34
VAQ-141	2018	1,098	1,071	0.98
VAQ-141	2019	1,129	750	0.66

Missing data manifests itself in a different way with VAQ-137 where a large number of sorties are not captured at all. The squadron deployed twice in 2018 each for approximately four months and cumulatively, the data only shows 125 total sorties. One would expect approximately 700 sorties for the entirety given the amount of time the squadron spent deployed.

Comparing MU and Non-MU Sorties

In this section and elsewhere we perform hypothesis tests as if the observed data were like a random sample from a (hypothetical) distribution of data. To assess differences in the rates of missing MU data, this study employed the χ^2 test. The χ^2 test allows us to compare squadrons by the amount of MU data present while deployed. In this case the null hypothesis, H_0, is that the distributions of missing and present MU data are not dependent on the squadron within a CVW with the alternative hypothesis, H_a, being that they are. The entirety of the CVW is first considered and then the F/A-18E squadrons within that CVW. The reason F/A-18E squadrons are considered is because the study assumes the number of sorties executed within a specific CVW should be similar between squadrons of like TMS. In both cases, this may give an indication as to whether there are issues with particular CVWs on specific deployments.

To first compare all squadrons on a particular deployment, we use CVW-5's 2018 deployment to the western Pacific AOR because it shows a relatively high proportion of captured sorties for all squadrons. Figure 4.9 was constructed by Barnhill et al. (2020). It shows a representation of MU vs. non-MU sorties by month from 2016 to 2019 specifically detailing deployed sorties. Non-MU sorties are those in light blue while dark blue indicates MU sorties. Gray bars show non-deployed sorties, which we ignore for this analysis. Table 4.2 shows the MU and non-MU sorties for each squadron during the deployment.

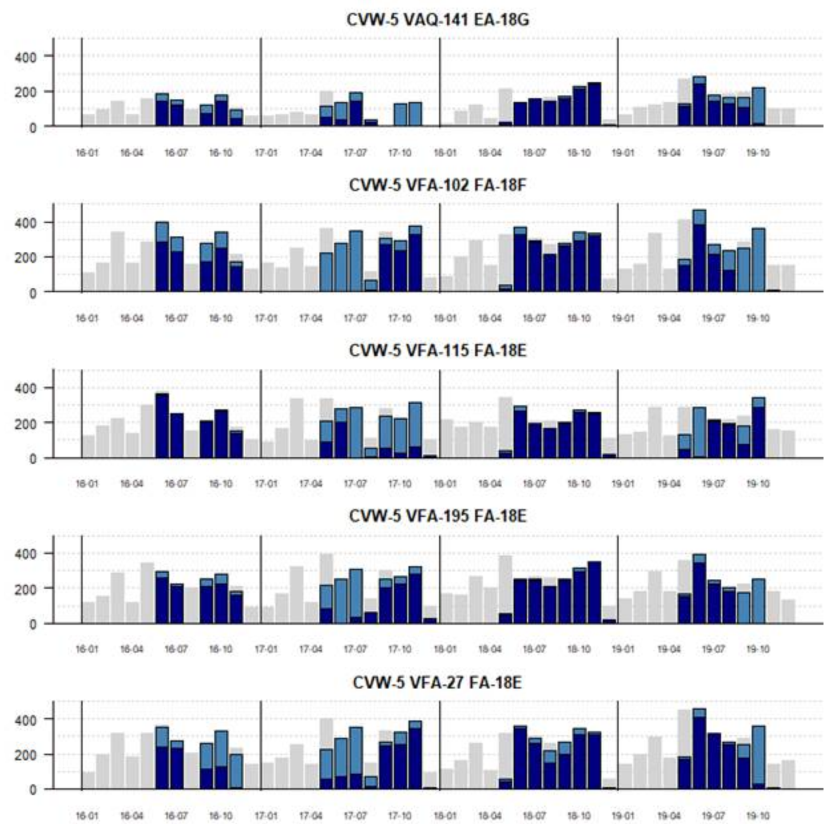

Figure 4.9. CVW-5 MU and Non-MU Sorties by Month 2016 to 2019.
Source: Barnhill et al. (2020).

Table 4.2. CVW-5 2018 MU vs. Non-MU Sorties by Squadron.

TMS	Squadron	MU	Not MU
F/A-18E	VFA-115	1,365	65
F/A-18E	VFA-195	1,651	55
F/A-18E	VFA-27	1,619	235
F/A-18F	VFA-102	1,724	130
E/A-18G	VAQ-141	1,071	27

Conducting the χ^2 test provides χ^2 statistic of 187.14 on 4 degrees of freedom with p-value ≈ 0 leading us to reject H_0. This suggests that the rate of MU sorties in this deployment is dependent on the squadron.

Now we consider only the F/A-18E squadrons since there are three of them. Our null hypothesis remains the same, that available MU data and squadrons are independent. Conducting the χ^2 test results in a test statistic of 140.12 on two degrees of freedom and a p-value $\approx$ 0. This gives further indication that available MU sorties is dependent on the squadron. Similar tests conducted on other deployed CVWs yielded similar results, namely, that the availability of MU sorties is dependent on the squadron.

Aggregation Fields with Missing Values

Refining aggregation to more detailed categories revealed more problems associated with the data fields themselves. Of interest are the TMR and mission assignments because together they indicate the type of sortie being conducted. For many sorties, these values are left blank or they do not exist because there is no NAVFLIR or SHARP record. The "TMR" field can be problematic even when present, since in many cases sorties are not assigned a legitimate code (i.e., a three-character alphanumeric identifier). Moreover many TMRs are associated with only a few sorties. While each platform exhibits fewer than 15% unassigned TMRs, most TMRs have no more than 30 sorties assigned. Out of 103 F/A-18E TMR categories, 62% have fewer than three sorties assigned. For F/A-18F sorties, 50% of

TMR codes have three or fewer sorties assigned. EA-18G MU TMRs show 52% of sortie codes with fewer than 30 sorties.

Issues with mission assignment are particularly apparent for EA-18G squadrons. Only 14% of these sorties are assigned a mission. For the F/A-18E squadrons, 79% of sorties have an assigned mission and 87% have an assigned TMR. In F/A-18F sorties 82% of sorties are assigned a mission and 86% possess TMR assignments. Table 4.3 shows the numbers of deployed sorties with MU data and the numbers of those sorties with unassigned TMR or mission.

Table 4.3. Deployed MU Sorties with Total and Proportion of Unassigned TMR or Mission, by Platform.

Platform	**Total**	**Unassigned TMR**		**Unassigned Mission**	
F/A-18E	43,881	5,717	.13	7,683	.18
F/A-18F	23,461	3,199	.14	4,879	.21
E/A-18G	8,718	879	.10	7,506	.86

Given the number of missing values in the TMR and mission fields, care must be taken when aggregating fuel consumption at these levels. Further, because there are so many TMR options for each TMS and a low number of observations per TMR, output metrics based on TMR aggregation are only meaningful for certain TMR assignments.

4.3 Conclusions Based on Data with Missing Values

Though care must be taken in drawing conclusions from data sets with missing observations, relevant conclusions are still possible though they are limited to broad aggregation levels and may not be comparable to like groups (i.e., deployment to deployment comparison). Because the F/A-18E squadrons experienced a TMS transition from the F/A-18C to the F/A-18E platform, this section addresses the effect of differing TMS compositions in a CVW with regard to deployment comparison. Additionally, we examine how deployment AOR affects fuel consumption for deployments from 2016 to 2019.

4.3.1 Implications of CVWs with Differing TMS

Deploying CVWs usually consist of three F/A-18E squadrons, one F/A-18F squadron, and one EA-18G squadron. However, in several cases, this structure differs as shown in Tables B.2 and B.3. Further, some CVW TMS structures change between deployments. In total, we observe seven deviations out of 19 deployments from the typical CVW.

In most cases, structure change occurs with regard to the number of F/A-18E squadrons. These structural deviations occur in four forms. The first occurs when squadrons began the transition from F/A-18C to F/A-18E aircraft. Any squadron sorties conducted in F/A-18C aircraft are not present in the data. The second case is when U.S. Marine Corps (USMC) F/A-18C squadrons are assigned to deployed CVWs to fill out the aviation complement. As with USN F/A-18C squadrons, these sorties are not included in the data set. The third case is the absence of a third F/A-18E squadron for some deployments. Lastly, we see that instead of one F/A-18F squadron and three F/A-18E squadrons, one CVW is made up of two of each.

This does not affect the analysis for a specific CVW, but it makes it harder to compare two deployments with different structures. An example of this occurs with CVW-7. In 2016 CVW-7 consisted of one F/A-18F, one EA-18G, two F/A-18E squadrons, and one F/A-18C squadron. In 2019, CVW-7 deployed with three F/A-18E squadrons, one F/A-18F, and one EA-18G squadron. Table 4.4 gives a comparison of sortie counts by F/A-18E squadron.

Table 4.4. CVW-7 F/A-18E Sortie Count and Average Fuel Used by Deployment Year.

Year	Squadron	Total MU Sorties	Avg. Fuel Used
2016	VFA-143	1,486	19,696.9
	VFA-25	1,193	
2019	VFA-143	1,887	13,159.9
	VFA-25	1,654	
	VFA-86	1,367	

The two CVW-7 deployments exhibit similar percentages of MU data (2016: 73.1%; 2019:

69.7%). However, in 2016 CVW-7 only possessed two F/A-18E squadrons instead of three. VFA-83 flew the F/A-18C and filled the role of the third "single-seat" squadron on that deployment. Conversely, three F/A-18E squadrons were present for the 2019 deployment. We cannot definitively compare the same CVW between deployments given the lack of sorties from VFA-83. Comparing the entirety of the CVW between deployments and resultant fuel output omits any fuel consumed by VFA-83 in 2016. Using the average of sorties from "sister squadrons" in the CVW as a guide, this means approximately 1,800 sorties are missing of which 75% would be expected to contain MU information.

Considering only the F/A-18E TMS, we may compare central tendency measures between the two deployments. Figure 4.10 shows the distribution of fuel used for MU sorties conducted between each deployment for that TMS.

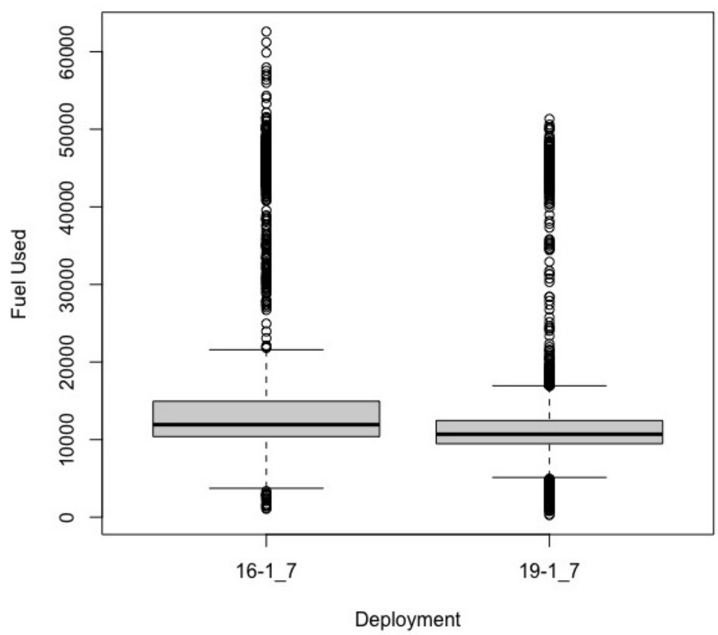

Figure 4.10. Fuel Used by F/A-18E Sorties from the CVW-7 2016 and 2019 Deployments.

Because neither distribution allows for a normality assumption, this study uses the non-parametric Wilcoxon-Mann-Whitney (WMW) test for comparison of medians giving an indication of similarity of distributions between two independent samples from the same population (Sprent and Smeeton 2007). The null hypothesis, H_0, is that medians are equal whereas the alternative, H_a, is that the medians are unequal. Using a level of significance, $\alpha = 0.05$, we reject H_0 in the comparison ($w > 8,400,000$, p-value < 2e-16). Keeping the proportion of missing sorties in mind, this result gives indication that CVWs should not be assumed to have similar fuel consumption distributions across deployments.

We also examine deployments conducted by CVW-5 in 2016, 2017, 2018, and 2019. CVW-5 is assigned to FDNF in Japan and is consistently scheduled to execute similar deployment operations from year to year in the western Pacific Ocean. Squadrons assigned to CVW-5 are much less likely to change as compared with CVWs based in the continental U.S. (CONUS). Figure 4.11 shows the distribution of fuel used for the four deployments for F/A-18E TMS.

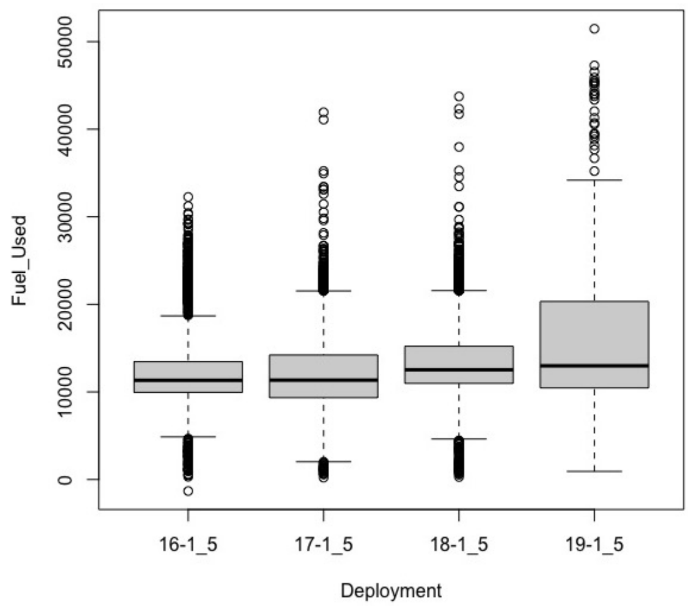

Figure 4.11. CVW-5 2016, 2017, 2018, and 2019 Deployment F/A-18E Fuel Used.

To compare the median fuel consumption across deployment years, the Kruskal-Wallis (KW) test for equality of medians is employed. This is similar to the WMW test though it is specifically designed for more than two comparative categories. In this case, the comparison is the median fuel used for each deployment. For this test H_0 is that medians are identical between the populations (deployment sorties) and the H_a is that they are not (Sprent and Smeeton 2007). Conducting the test we find that at least one median differs from the rest (test statistic > 424 on three degrees of freedom, p-value < 2e-16).

Because the KW test is an omnibus test, we do not know which population median is different (Sprent and Smeeton 2007). To figure this out, we use pairwise comparisons (Daxue Consulting 2009). Using a level of significance, $\alpha = .05$, we see that the only deployments for which the null hypothesis of equal medians cannot be rejected are the 2016 and 2017 deployments (p-value = 0.086). All other pairs show evidence of differences. Though comparison of the 2016 and 2017 CVW-5 deployments suggest no difference, it is important to consider that 2017 exhibited the highest proportion of unincorporated MU sorties (63.2% of total sorties or 4,327 unincorporated sorties). We do not know how these sorties affect the overall fuel usage, making any sort of inference suspect.

4.3.2 Fuel Consumption by Deployment AOR

Recommendations from Barnhill et al. (2020) included examining fuel consumption by deployment AOR. The original data does not capture location information with any aggregation field. However, using known historic deployment destinations, a new category was constructed.

When CVWs deploy aboard aircraft carriers, they are usually bound for the Middle East (5th and 6th Fleet), the western Pacific (7th Fleet) AORs, or some combination of the two. These AORs are shown in Figure B.1. To reach the Middle East, aircraft carriers must transit 7th Fleet if they deploy from the west coast of the CONUS or 6th Fleet if they depart the east coast of the CONUS. Oftentimes, CVWs will be required to fly sorties in support of the numbered fleet commander or simply to maintain readiness during transit. These requirements demand flights in multiple AORs throughout a deployment. Table B.4 outlines numbered fleet destinations for each CVW deployment though it is not specifically captured in the data set. For the purposes of this thesis, Middle East deployments include those that at some point spent time in 5th or 6th Fleet. Western Pacific deployments are those that exclusively deployed to 7th Fleet.

To evaluate differences between deployment destinations we constructed a new, binary "Location" category where a sortie is assigned a 0 if they only went to the western Pacific and a 1 otherwise. Based on the available MU data, higher fuel consumption occurs on deployments ultimately destined for the Middle East AOR. Additionally, the presence of combat missions in that AOR correlate with higher fuel consumption, as well. Figure 4.12

shows the distribution of fuel used for all deployments. Figures 4.13 and 4.14 show fuel consumption separated by deployment destination. Table 4.5 provides the number of sorties by TMS in each AOR.

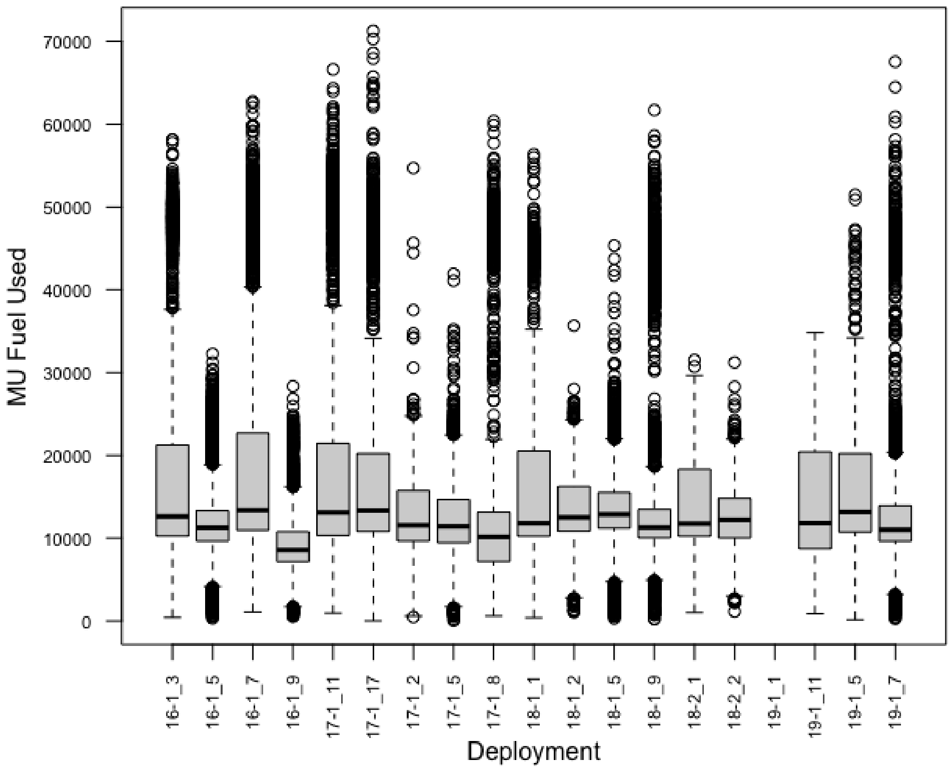

Figure 4.12. Distribution of MU Fuel Used by Deployment.

Table 4.5. MU Sorties by AOR.

AOR	F/A-18E	F/A-18F	EA-18G
Western Pacific	20,028	8,262	4,279
Middle East	23,843	15,196	4,437

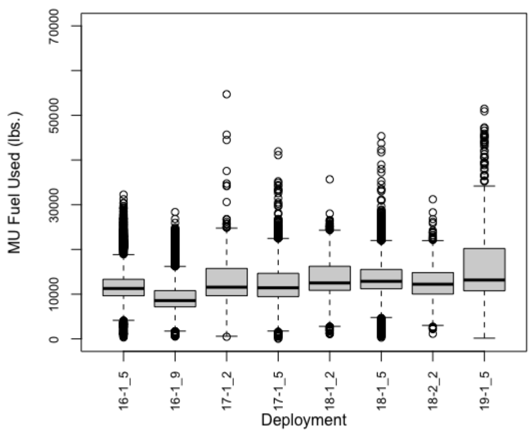

Figure 4.13. Western Pacific Deployment Fuel Distribution.

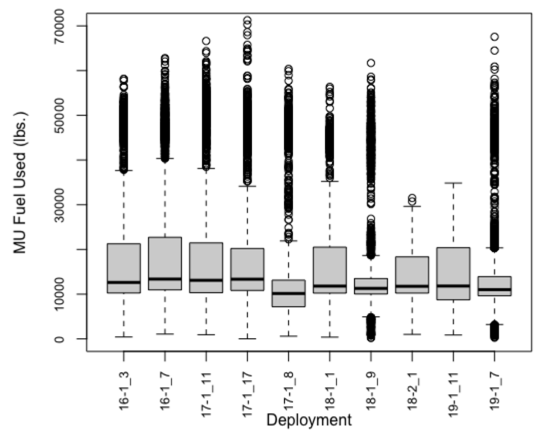

Figure 4.14. Middle East Deployment Fuel Distribution.

Of specific interest are the differing distributions between deployments. Figure 4.13 shows that those CVWs that deployed only to the western Pacific exhibited few sorties that resulted in greater than 40,000 pounds of fuel used per flight and thereby a narrower distribution. For those CVWs that deployed to the Middle East numbered fleets, more flights routinely exhibited fuel consumption of greater than 40,000 lbs.

To address differences in fuel consumption sortie duration is first considered because flight time is the primary factor in fuel consumption. Figure 4.15 shows sortie duration by location using all TMS. Given available MU data, longer sorties appear more likely to occur on deployments going to the Middle East rather than the western Pacific.

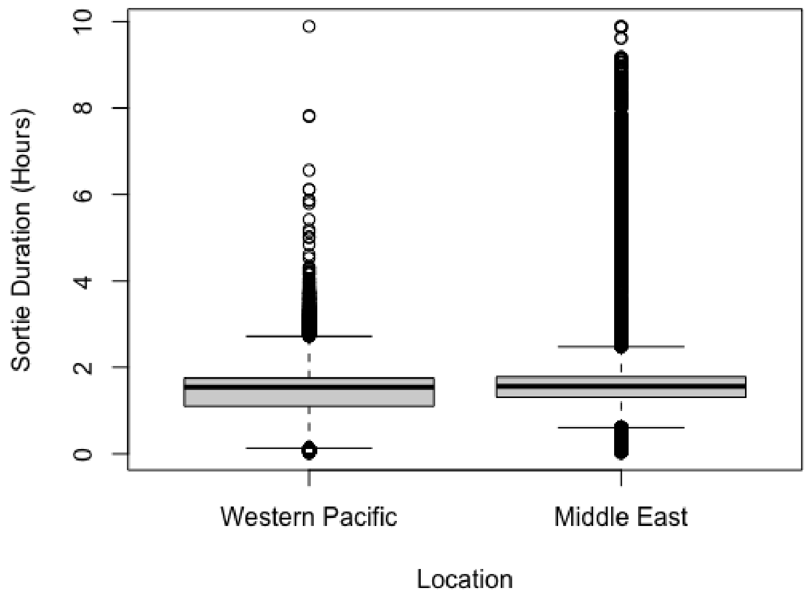

Figure 4.15. Sortie Duration for
Western Pacific and Middle East Deployments.

Fuel consumption is highly correlated with sortie duration so, overall, higher consumption rates occur on deployments that ultimately deploy to the Middle East. The study examines this in a three ways: 1) overall fuel used by TMS; 2) fuel taken by an aircraft during in-flight refueling (IFR) by TMS; and 3) fuel dumped by TMS and location.

Fuel Consumed by Deployment AOR

As posited in the previous section, given the available fuel data, higher sortie fuel consumption occurs on deployments that do not deploy exclusively to the western Pacific. Table 4.6 shows average fuel used per sortie by deployment for all TMS using available MU sorties. Table 4.7 shows average fuel used for each TMS by location. Figures 4.16, 4.17, and 4.18 show that sorties consuming more fuel occurred more often on deployments that deployed to the Middle East for all TMS.

Table 4.6. Avg. Fuel Used (in lbs.) by Deployment and AOR.

Year	CVW	AOR	Sorties	Avg. Fuel Used
2016	CVW-3	5th/6th Fleet	5,161	18,325.7
2016	CVW-5	7th Fleet	4,578	12,305.5
2016	CVW-7	5th/6th Fleet	4,798	20,959.6
2016	CVW-9	7th Fleet	6,424	9,221.4
2017	CVW-2	7th Fleet	3,119	13,241.1
2017	CVW-5	7th Fleet	3,518	12,745.1
2017	CVW-8	5th/6th Fleet	2,681	15221.1
2017	CVW-11	7th/5th Fleet	3,630	20,534.7
2017	CVW-17	7th/5th Fleet	5,173	18,671.3
2018	CVW-1	6th Fleet	3,254	15,960.4
2018	CVW-1	6th Fleet	2,663	13,597.7
2018	CVW-2	7th Fleet	1,916	13,893.3
2018	CVW-2	3rd Fleet	922	12,654.9
2018	CVW-5	7th Fleet	7,430	13,954.9
2018	CVW-9	7th/5th Fleet	8,108	13,465.9
2019	CVW-1	5th/6th Fleet	Unknown	Unknown
2019	CVW-5	7th Fleet	4,664	14,929.3
2019	CVW-7	5th/6th/7th Fleet	7,601	13,590.4
2019	CVW-11	3rd Fleet	404	13,550.6

Table 4.7. Average Fuel Used (in lbs.) by TMS and AOR.

F/A-18E Avg. Fuel Used by AOR	
AOR	**Avg. Fuel Used**
Western Pacific	12,681.5
Middle East	16,067.5

F/A-18F Avg. Fuel Used by AOR	
AOR	**Avg. Fuel Used**
Western Pacific	13,396.9
Middle East	16,769.4

EA-18G Avg. Fuel Used by AOR	
AOR	**Avg. Fuel Used**
Western Pacific	11,361.7
Middle East	16,970.5

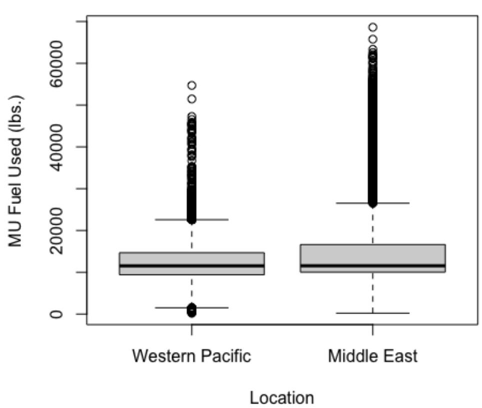

Figure 4.16. F/A-18E Fuel Used by AOR.

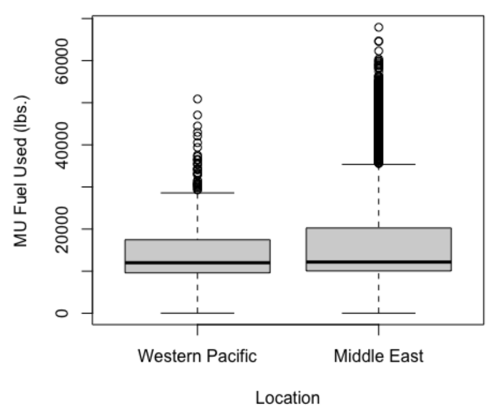

Figure 4.17. F/A-18F Fuel Used by AOR.

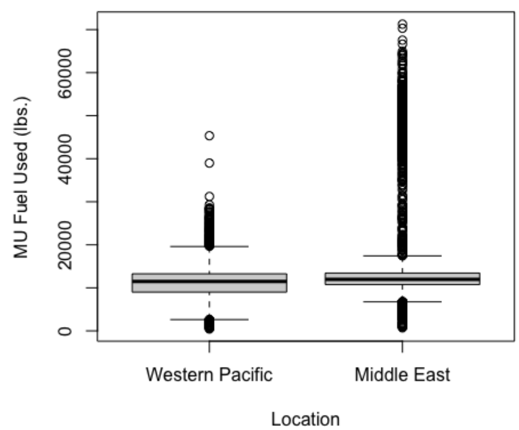

Figure 4.18. EA-18G Fuel Used by AOR.

In-flight-Refueling Fuel Taken by Deployment AOR

Longer sorties require more refueling opportunities during the flight. In addition to overall fuel used per sortie, the study examines fuel taken during tanking events for each sortie. Table 4.9 shows average fuel taken during tanking events by deployment as well as the number of tanking events that occurred. Table 4.8 shows the number of tanking events by TMS and deployment destination. Figures 4.19, 4.20, and 4.21 show that Middle East deployments possess higher "fuel taken" values across all TMS than western Pacific deployments. Smaller values exist in all types of deployments but in no case does one see values greater than approximately 38,000 pounds for any TMS only deploying to the western Pacific. Tanking events for EA-18G sorties appear to occur much more infrequently than for other TMS. However, between each AOR, EA-18G sorties exhibit a comparable number of tanking events though a considerable difference in average fuel taken during a tanking event. There is no clear indication why there is such a large difference for EA-18G sorties occurring in the Middle East AOR.

Table 4.8. Tanking Events by Platform and AOR.

	TMS		
AOR	F/A-18E	F/A-18F	EA-18G
Western Pacific	8,978	4,291	1,271
Middle East	10,583	8,507	1,302

Table 4.9. Avg. Fuel Taken (in lbs.) by Deployment and AOR.

Year	**CVW**	**AOR**	**Tank Sorties**	**Avg. Taken**
2016	CVW-3	5th/6th Fleet	3,042	9,822.3
2016	CVW-5	7th Fleet	2,471	2,552.3
2016	CVW-7	5th/6th Fleet	2,961	14,109.2
2016	CVW-9	7th Fleet	1,921	1,815.5
2017	CVW-2	7th Fleet	1,399	1,903.1
2017	CVW-5	7th Fleet	1,811	2,276.1
2017	CVW-8	5th/6th Fleet	1,022	15,748.2
2017	CVW-11	7th/5th Fleet	1,699	19,346.4
2017	CVW-17	7th/5th Fleet	2,521	13,322.3
2018	CVW-1	6th Fleet	1,968	6,803.9
2018	CVW-1	6th Fleet	1,595	2,459.4
2018	CVW-2	7th Fleet	951	2,253.1
2018	CVW-2	3rd Fleet	366	3,613.6
2018	CVW-5	7th Fleet	3,264	2,330.7
2018	CVW-9	7th/5th Fleet	2,676	7,265.5
2019	CVW-1	5th/6th Fleet	Unknown	Unknown
2019	CVW-5	7th Fleet	2,357	3,701.0
2019	CVW-7	5th/6th/7th Fleet	2,743	5,793.0
2019	CVW-11	3rd Fleet	167	5,560.1

Table 4.10. Average Fuel Taken (in lbs.) by TMS and AOR.

F/A-18E Avg. Fuel Taken by AOR	
AOR	**Avg. Fuel Taken**
Western Pacific	1,145.0
Middle East	4,555.4

F/A-18F Avg. Fuel Taken by AOR	
AOR	**Avg. Fuel Taken**
Western Pacific	2,346.1
Middle East	8,587.4

EA-18G Avg. Fuel Taken by AOR	
AOR	**Avg. Fuel Taken**
Western Pacific	2,650.4
Middle East	20,065.5

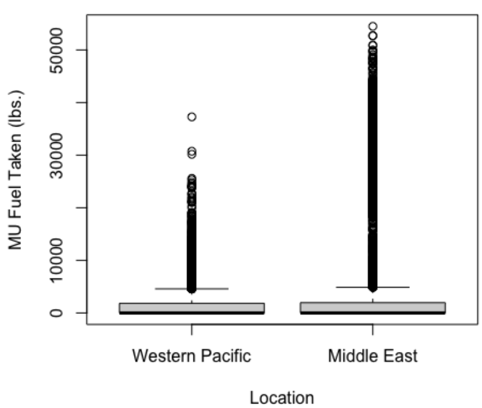

Figure 4.19. F/A-18E Fuel Taken by AOR.

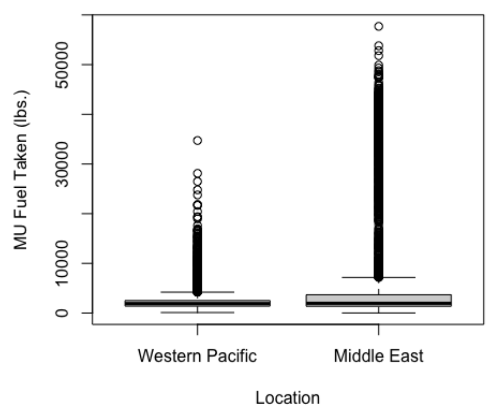

Figure 4.20. F/A-18F Fuel Taken by AOR.

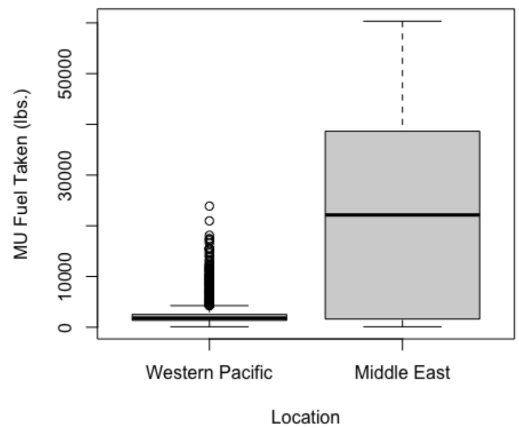

Figure 4.21. EA-18G Fuel Taken by AOR.

Fuel Dumped by Deployment AOR

The amount of fuel dumped by aircraft is particularly important to fuel conservation because it indicates how much fuel goes completely unused. Pilots dump fuel so they can reach an acceptable landing weight aboard ship. Analysis suggests that the median amount of fuel dumped is similar between the AORs. Table 4.11 shows the number of sorties with a dump event separated by TMS and platform destination. Overall, fewer dump events occur in sorties occurring in the western Pacific AOR than the Middle East AOR.

Table 4.11. Dumping Events by Platform and AOR.

	TMS		
AOR	F/A-18E	F/A-18F	EA-18G
Western Pacific	4,945	2,723	931
Middle East	6,936	6,345	1,307

Table 4.12 indicates the number of sorties with dump events for each deployment and average fuel dumped per sortie for each sortie that had a dump event. Figure 4.22 gives the distribution of the pounds of fuel dumped by deployment. Nothing about the distribution

gives an indication that location affects how much fuel is dumped during a sortie. Figures 4.23, 4.24, and 4.25 illustrate the fuel dumped by TMS for each AOR. Regardless of TMS and location, median fuel dumped values are approximately equal and distributions appear similar.

Table 4.12. Avg. Fuel Dumped (in lbs.) by Deployment and AOR.

Year	CVW	AOR	Dump Sorties	Avg. Dumped
2016	CVW-3	5th/6th Fleet	1,848	2,138.3
2016	CVW-5	7th Fleet	402	2,165.3
2016	CVW-7	5th/6th Fleet	2,125	2,450.1
2016	CVW-9	7th Fleet	1,003	1,966.1
2017	CVW-2	7th Fleet	1,686	2,112.4
2017	CVW-5	7th Fleet	961	2,358.2
2017	CVW-8	5th/6th Fleet	2,362	1,874.5
2017	CVW-11	7th/5th Fleet	963	1700.4
2017	CVW-17	7th/5th Fleet	1,172	1,763.1
2018	CVW-1	6th Fleet	539	1,952.8
2018	CVW-1	6th Fleet	638	2,390.2
2018	CVW-2	7th Fleet	931	1,893.5
2018	CVW-2	3rd Fleet	1,698	1,737.1
2018	CVW-5	7th Fleet	2,836	2,207.2
2018	CVW-9	7th/5th Fleet	238	1,462.7
2019	CVW-1	5th/6th Fleet	Unknown	Unknown
2019	CVW-5	7th Fleet	1,956	2,355.4
2019	CVW-7	5th/6th/7th Fleet	1,657	1,919.1
2019	CVW-11	3rd Fleet	167	2,775.2

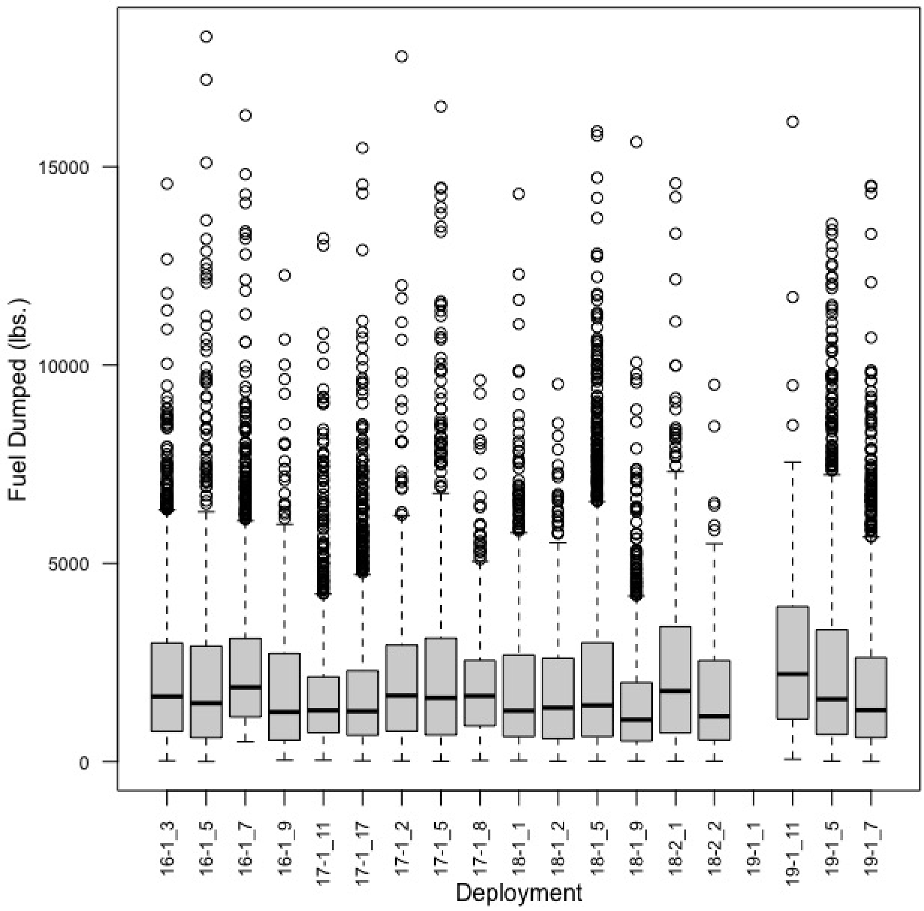

Figure 4.22. Fuel Dumped by Deployment.

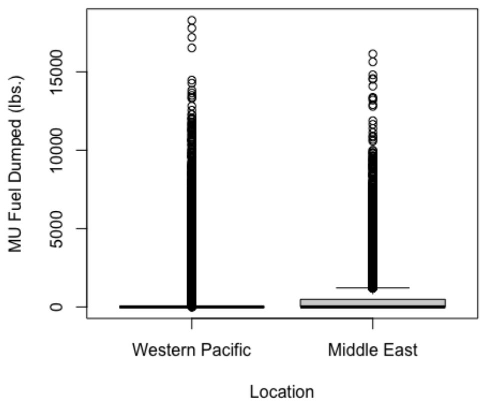

Figure 4.23. F/A-18E Fuel Dumped by AOR.

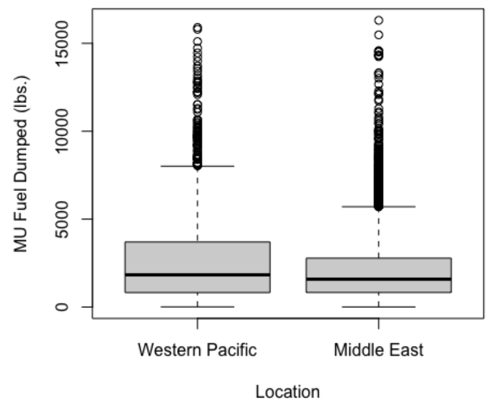

Figure 4.24. F/A-18F Fuel Dumped by AOR.

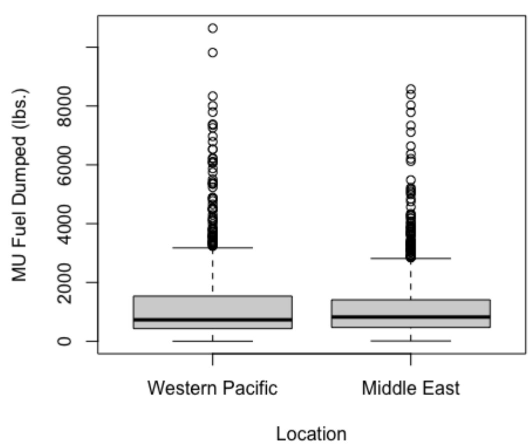

Figure 4.25. EA-18G Fuel Dumped by AOR.

Table 4.13. Average Fuel Dumped (in lbs.) by TMS and AOR.

F/A-18E Avg. Fuel Dumped by AOR	
AOR	**Avg. Fuel Dumped**
Western Pacific	530.7
Middle East	570.6

F/A-18F Avg. Fuel Dumped by AOR	
AOR	**Avg. Fuel Dumped**
Western Pacific	2,603.8
Middle East	2,092.1

EA-18G Avg. Fuel Dumped by AOR	
AOR	**Avg. Fuel Dumped**
Western Pacific	1,221.4
Middle East	1,127.5

Fuel Used by Mission Type and AOR

Analysis in previous sections suggests that fuel consumption is lower for those sorties occurring on deployments exclusively to the western Pacific versus those executed on deployments to the Middle East. In this section the study considers the effect of sortie mission type on fuel consumption based on deployment AOR. We speculate that increased fuel consumption is due to the presence and demands of combat missions in the Middle East as compared to the western Pacific.

The sortie mission type is determined from the TMR code. Using the first character, or flight purpose code (FPC), in the three-character alphanumeric code, the type of mission is derived for each sortie. Table 4.14 defines FPCs for the NAE and Tables 4.15, 4.16, and 4.17 show the numbers of sorties with each FPC by TMS and deployment AOR. In practice, a flight will always be assigned a TMR; however, in this analysis many sorties do not possess a TMR due to the lack of NAVFLIR and SHARP data. Flights lacking a TMR have been assigned a "U" for "Unassigned." Figures 4.26, 4.27, 4.28, 4.29, 4.30, and 4.31 show boxplots for all TMS FPCs differentiated by location. Only FPCs with greater than 100 sorties are shown.

Table 4.14. Flight Purpose Codes (FPC). Adapted from: CNAF (2016).

FPC	Flight Purpose
1	Training
2	Support Services
3	Operational Tasking
5	Contingency Operations
6	Combat Operations
7	Military Exercises
U	Unassigned

Table 4.15. F/A-18E FPCs by AOR.

AOR	1	2	3	5	6	7	U
Western Pacific	12,491	5,430	511	267	347	2	980
Middle East	12,053	6,062	643	132	3,524	0	1,429

Table 4.16. F/A-18F FPCs by AOR.

AOR	1	2	3	5	6	7	U
Western Pacific	4,695	2,814	335	0	29	2	383
Middle East	6,496	5,322	658	1	1,774	2	942

Table 4.17. EA-18G FPCs by AOR.

AOR	1	2	3	5	6	7	U
Western Pacific	3478	359	0	0	0	0	442
Middle East	2,727	196	137	0	954	0	423

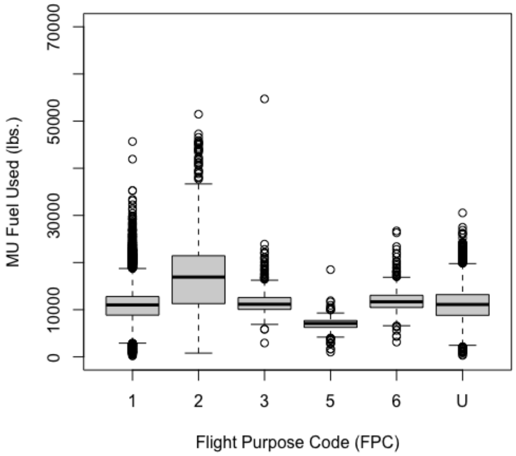

Figure 4.26. F/A-18E Western Pacific Fuel Used by FPC.

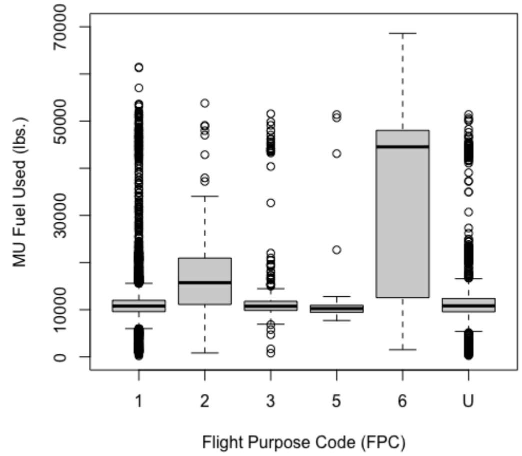

Figure 4.27. F/A-18E Middle East Fuel Used by FPC.

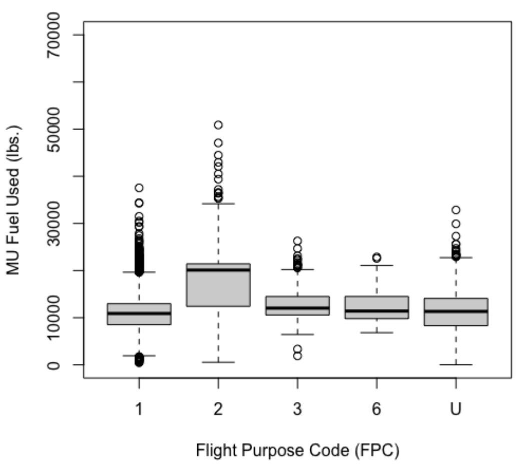

Figure 4.28. F/A-18F Western Pacific Fuel Used by FPC.

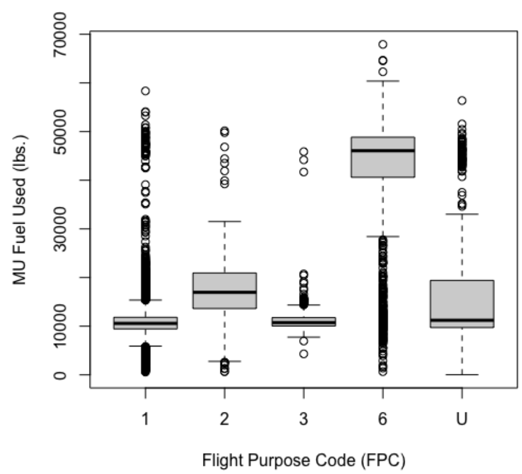

Figure 4.29. F/A-18F Middle East Fuel Used by FPC.

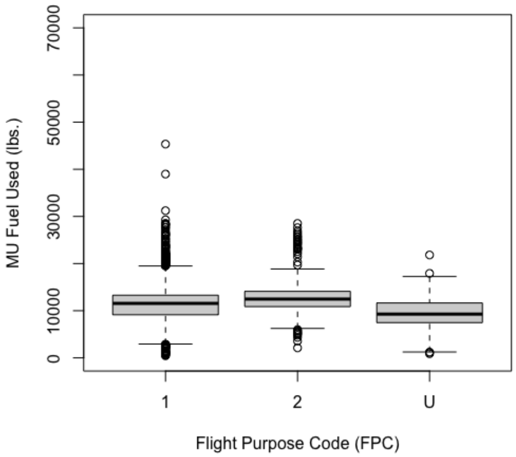

Figure 4.30. EA-18G Western Pacific Fuel Used by FPC.

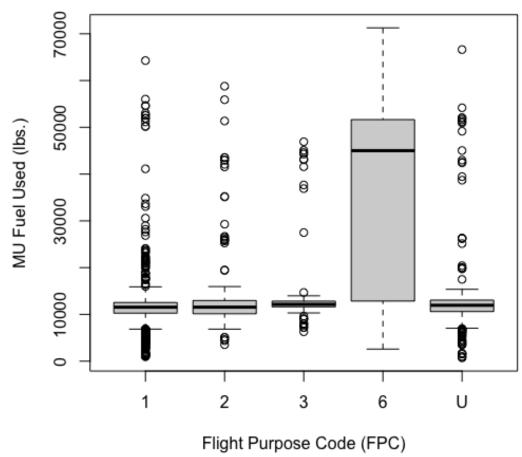

Figure 4.31. EA-18G Middle East Fuel Used by FPC.

For all TMS, combat flights (denoted by a "6") occurring in the Middle East are associated with higher fuel consumption metrics. While median fuel consumption is similar for other FPCs between AOR, the western Pacific median combat fuel consumption is far lower (when they exist) than for combat sorties in the Middle East. These findings give an indication that combat operations drive an increase in tanking operations (due to increased sortie duration) and thereby fuel consumption. However, the exact reason why sortie durations are higher for Middle East combat operations is not immediately clear.

4.3.3 Conclusions

This section examined results from statistical techniques used to derive broad conclusions about the fuel consumption for deployed F/A-18E/F and EA-18G sorties for the time frame of 2016 to 2019. While the effect of missing sorties cannot be underestimated, especially because it cannot be known, broad inference is still possible. The study determined that direct comparison of fuel consumption between deployments is not reliable between different CVWs or between deployments for the same CVW due to differing sortie distributions, even if deployments were to exhibit similar proportions of missing MU data.

While comparing deployments is not tenable, one can draw broad conclusions about deployment destination and associated fuel consumption by TMS. The data show that Middle East deployment sorties consumed more fuel than those sorties deployed to the western Pacific, presumably due to the nature of combat mission requirements in the Middle East though the mission requirements are not immediately apparent in the data. Identifying trends such as these, though, could allow decision makers and operational staffs to better plan deployment fuel requirements based on ongoing events in deployment AORs.

4.4 Prediction Analysis

This study analyzed four random forest models using a combination of original and generated predictors to assess how well MU fuel consumption in available observations could be predicted. Models also included various methods of imputation to increase the number of available observations in the data set.

Imputation may be used to populate missing data values in a data set. Increased data inclusion helps reduce bias though the analyst must be cognizant of the pattern of missing data. If the pattern of missing data is considered MNAR, analysis and inference is at risk of bias because the effect of missing data may be substantial even though it is not known. While imputation can still bolster information regarding the known observations, analysts must be careful to understand that detailed inference is not recommended.

Previous analysis conducted by Barnhill et al. (2020) suggested that the pattern of data was MNAR. This meant that while broad generalizations are possible, specific output metrics desired by NAOEP cannot be reliably computed. Determining whether computation of metrics is reliable or not requires knowing the proportion of missing data and whether it is in the form of missing value fields or suspected missing observations for a given aggregation level.

Barnhill et al. (2020) confirm that SHARP fuel values are different enough from MU fuel values (see Figure 4.32) that they cannot be used to predict missing MU fuel values. Instead, the authors used random forest algorithms to predict fuel consumption specifically related to MU data. Using a combination of predictors, Barnhill et al. (2020) built random forest models to predict “mu_fuelused” in order to assess whether predictions could prove effective in imputing missing MU fuel values.

In this section, we continue the research by comparing several random forest models to examine which attained the best predictions as determined by a combination of R^2, mean absolute difference (MAD), and root mean squared error (RMSE). Models described in Section 3.3.4 were used to predict fuel used for sorties for all TMS in a combined data set. Pre-processing methods pared deployed observations from 105,197 observations to those exhibited in Table 3.6.

4.4.1 Model Performance

In this section we discuss model performance metrics for each of the four models. Reference to Section 3.1.3 for predictor definitions.

Table 3.6 provides the sample sizes for each model. In models 1 and 2, the number of observations was determined by available complete-case observations for SHARP data that also contained MU information. Models 3 and 4 represent the total number of deployed sorties as determined by pre-processing methodologies outlined in Chapter 3 with removal of five observations with negative fuel values. Table 4.18 shows the overall performance metrics and hyper-parameter choice after cross-validation. All test sets consisted of observations with MU data that were present regardless of whether the data set construction employed imputation.

Table 4.18. Prediction Model Metrics.

Model	**MTRY**	R^2	**MAD (in pounds of fuel)**	**RMSE**
Model 1	1	0.511	3,584.5	6,960.8
Model 2	5	0.926	1,590.4	2,882.9
Model 3	13	0.961	1,046.3	1,958.6
Model 4	12	0.985	631.1	1,172.4

Increasing the number of predictors will always increase R^2 or percentage of explained variance. We note that R^2 increases greatly between those models only using SHARP predictors and those using a combination of original and generated predictors. Further the MAD and RMSE metrics decrease with each subsequent model as we increase the number of predictors and observations. Ultimately, model 4 displays the best ability to predict MU fuel used. The resulting MAD represents approximately three minutes of flight time, that is, the error associated with estimating "mu_fuelused" from this model averages about three minutes per sortie using a standard fuel burn rate of 9,000 pounds per hour (Deloitte Analytics Team 2020). In the following sections we examine each model in detail.

Model 1: SHARP Fuel Used

This section describes model 1 which only uses the "sharp_fuelused" category to predict MU fuel used ("mu_fuelused"). Figure 4.32 shows the relationship between SHARP and MU fuel used values. We examine the relationship between the two because ideally, SHARP fuel entries could be used to impute missing MU fuel entries (and vice versa) because they should be the same. The red line in Figure 4.32 indicates the relationship expected between the two categories if SHARP matched MU for every observation.

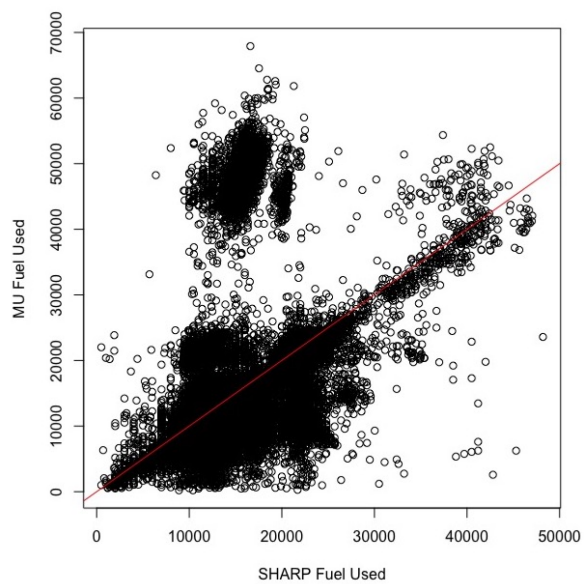

Figure 4.32. SHARP Fuel Used vs. MU Fuel Used.

While observations generally trend together, we see large clusters suggesting a discrepancy between the two "fuel used" values in the same observation. This alone calls into question the accuracy of both fuel values. However, because MU is direct aircraft output and SHARP values are human inputs, MU is considered correct while SHARP appears to be subject to human error. Thus, we cannot simply impute missing MU values with SHARP values with high accuracy.

While all SHARP categories are subject to human error, they can still be used for modeling. Because only one predictor is used in model 1 there is no need to assess different values for the “MTRY” hyper-parameter. Using the 90%/10% train/test split we constructed a model of 300 trees. Figure 4.33 shows the residual plot for Model 1 where we want to see residuals close to zero.

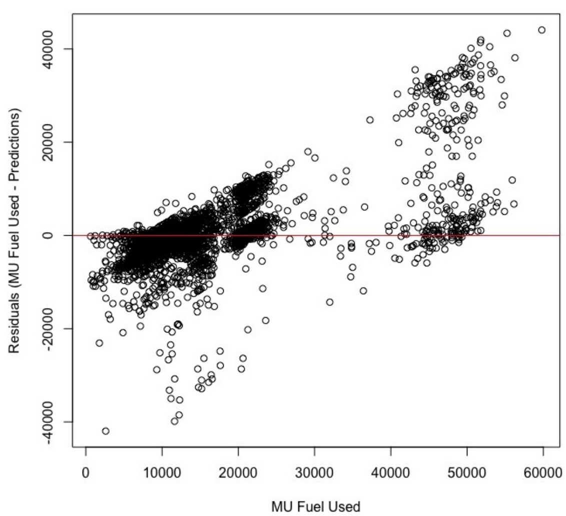

Figure 4.33. Residual Plot for Model 1.

We see that SHARP overestimates MU fuel for low values and underestimates MU values for larger values. The R^2 is 0.511 meaning approximately 50% of the variance in the response is explained in this model.

Model 2: Multiple SHARP Categories

This section describes Model 2 which uses 13 SHARP categories to predict MU fuel used values. Table 4.19 shows the predictors used. Using the 90%/10% train/test split, a random forest model was constructed using 300 trees. The MAD and RMSE performance metrics decrease by 55% and 58% respectively compared with the same metrics for model 1.

Figure 4.34 shows the residual plot for model 2. Model 2, using multiple SHARP categories more effectively predicts MU fuel used as indicated, visually, by residual proximity to zero. While residuals still indicate overestimates at small values and underestimates at higher values, errors are much less pronounced than in model 1.

Table 4.19. Model 2 Predictors.

Predictor
sharp_sortieduration
sharp_tms
sharp_squadron
sharp_fuelstart
sharp_fuelend
sharp_fuelused
sharp_fuelburned
sharp_fuelburnrate
sharp_fueldumped
sharp_fuelgiven
sharp_fueltaken
sharp_tmr
sharp_tmr1

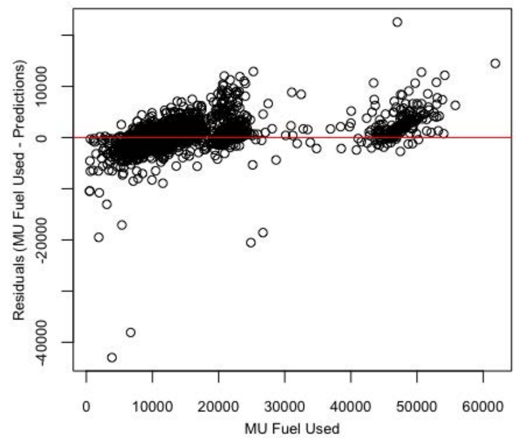

Figure 4.34. Residual Plot for Model 2.

Model 3: Imputing Predictors with Mean Values

This section describes the third model which uses 20 variables to predict MU fuel used. To increase the number of usable observations, means of non-missing individual numeric variable values are used to impute missing values for each respective numeric variable. This increases the number of usable observations from 45,824 to 105,192 across all three TMS. Table 4.20 shows the predictors used in the model, dividing them between predictors generated by the authors and predictors extracted from the original data set. This model uses nine of the original thirteen SHARP predictors used in models 1 and 2 because several of the generated predictors capture information from the SHARP predictors as described in Section 3.2.1 (i.e., "TMS" for "sharp_tms," "Squadron" for "sharp_squadron," "TMR" for "sharp_tmr," and "Mission" for "sharp_tmr1").

Given the number of available observations, an 85%/15% train/test split was employed and a random forest model built using 300 trees with the "MTRY" hyper-parameter equal to 13. Figure 4.35 shows that the residuals cluster around zero with smaller spread. R^2 indicates that model 3 explains almost 3% more variability in fuel consumption than model 2. While R^2 is expected to increase because model 3 uses more predictor variables, a 3% increase in R^2 is considerable. The MAD and RMSE decrease from model 2 by 34% and 32%

respectively and from the original model 1 by 71% each, further indicating that model 3 more accurately predicts MU fuel used than previous models.

Table 4.20. Model 3 Predictors Separated by Generated and Original Variables.

Generated Predictors	**Data Set Predictors**
TMS	nav_approachcode
Squadron	sharp_sortieduration
Airwing	sharp_fuelstart
Route	sharp_fuelend
Duration	sharp_fuelused
Mission	sharp_fuelburned
TMR	sharp_fuelburnrate
TOD	sharp_fueldumped
configuration	sharp_fuelgiven
Mon	sharp_fueltaken

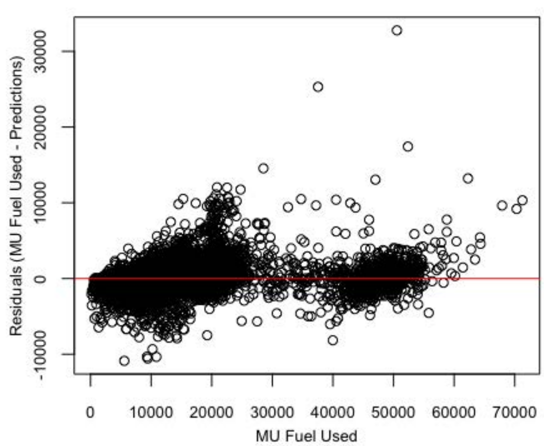

Figure 4.35. Residual Plot for Model 3.

Model 4: Multiple Imputation Model

This section describes the fourth model in which missing numeric values are imputed using the "missForest" MI algorithm. Model 4 uses the same predictors as model 3 with the same number of observations.

To execute MI on the 105,192 observations, the data set was divided by TMS and the "missForest" algorithm was used to impute each sub-data set. While "missForest" is useful because it can impute both continuous and categorical predictors, it is time-intensive. Imputing the F/A-18E sub-data set of approximately 60,000 observations with 21 variables required 84 hours, uninterrupted. The F/A-18F sub-data set required approximately 18 hours and the EA-18G sub-data set required approximately three hours. Once the imputation completed, the three TMS sub-data sets were merged to re-form a new data set of 105,192 observations of complete data.

Figure 4.36 exhibits the model 4 residuals. We observe that model 4 residuals are closer to zero with fewer aberrations than Model 3. Model 4 possesses a R^2 value of 0.985 with MAD and RMSE equal to 631.1 and 1,172.4, respectively. This far outperformed other models, indicating that MI is effective in filling missing values and the resulting data set predicts comparatively well.

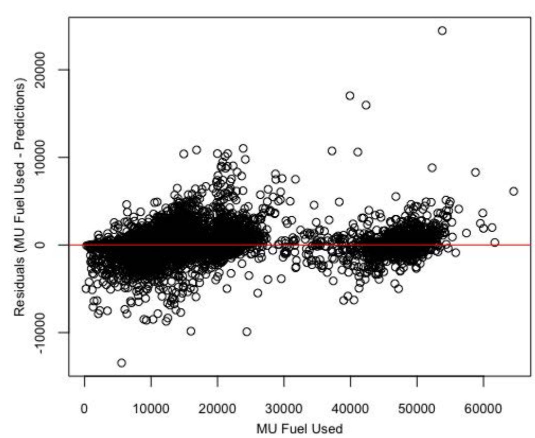

Figure 4.36. Residual Plot for Model 4.

4.4.2 Prediction Analysis Conclusion

This section assessed four separate models of increasing complexity to determine how well MU fuel consumption can be predicted using original data variables alone and combined with generated predictors. The study showed that imputing MU data directly with "SHARP fuel used" values is not recommended because of the discrepancy between the MU and SHARP values. Of the four models, model 4 which used a data set imputed using MI techniques, resulted in the model with the best performance in both MAD and RMSE. However, even for models 2 and 3 we observe that MU fuel data may effectively predict without the computational obstacles of MI.

Regardless of the effectiveness of MI, statistical inference or prediction on this data are problematic. Specifically, these concerns arise from the amount of missing data, the assumed MNAR pattern, and the unknown effect of the missing data on metric or prediction output.

CHAPTER 5:
Conclusion

5.1 Conclusions

In this study we assessed the usefulness of F/A-18 fuel data underpinning NAOEP's Fuel Conservation Analytics Dashboard. Despite the inherent value and richness of the amount and sources of data, the analysis confirms that the amount and unpredictability of missing data observations, field values, or MU data makes derived output metrics unreliable. The study exhibited these issues through comparison of proportions of missing data across TMS by deployment, deployments conducted by the same CVW and associated squadrons to the same AOR, and individual aggregation fields resident in the data set.

Regardless of the missing data, the study highlights several trends that warrant further exploration. Specifically, assessing fuel consumption by deployment location indicates higher MU fuel consumption for CVWs ultimately deploying to the Middle East AOR versus those that only went to the western Pacific, regardless of TMS. The study also gives evidence that the prevalence of combat missions as determined from TMR data drive higher MU fuel consumption in the Middle East AOR. While there is no claim that location and associated mission type are the only considerations for fuel consumption, the combination of the two give reason to consider these factors in future analysis especially since location is absent in the original data.

This study also considers how well predictor variables resident in the data and generated in pre-processing predict MU fuel consumption. Analysis suggests that increasing the number of available observations by various methods of imputation enhances predictive ability in random forest models. While models using only predictors resident in the data set exhibit relatively high R^2 (.9260), models using imputation indicate overall improvement as indicated by R^2, MAD, and RMSE. Of the two imputation models considered (mean and multiple imputation), model 4 which employed MI to fill missing values exhibited the best predictive capability as indicated by metrics provided in Table 4.18 as well as by visual inspection in Figure 4.36.

5.2 Recommendations

5.2.1 Missing Data Mitigation

Missing information in the data set inserts unintended bias into any statistical output resulting in questionable conclusions that might result. It is recommended that NAOEP not only endeavor to capture as high a proportion of data as possible from relevant databases but to assess reasons why observations are missing in the first place. This will ensure data completeness and reduced bias, allowing for tractable analysis.

5.2.2 Variable Inclusion

While NAVAIR incorporates rich information in the merged data set, it is recommended that it expand the type of data it includes to capture nuances of naval aviation. Including AOR-specific location information will allow for more detailed inference regarding fuel consumption especially since 1) no deployment destination information is present in the current data set and 2) operations differ between deployments even between different CVWs making comparison difficult. NAOEP can easily access historic deployment destinations through CNAF to add detail to the data set.

Additionally, obtaining correct OFRP information will allow for detailed analysis beyond simple deployment or non-deployment. OFRP information can expand analysis to focus on readiness across aggregation levels serving to inform decision-makers on fuel requirements to prepare units for upcoming deployments.

Lastly, aircraft output parameters can be leveraged to add detail to the data set. Specifically, the MU records airspeed and altitude which can be incorporated into the data adding detail for further analysis.

5.2.3 Prediction Analysis

This study showed that available data can be used to predict fuel consumption metrics of interest despite high proportions of missing data by way of imputation. However, because so much data is missing, predictions must still be qualified with the amount of missing information. Assuming NAOEP is able to correct deficiencies in missing data and incorporate recommended resident and generated predictors as defined in this study and by Barnhill et al. (2020), the Fuel Conservation Analytics Dashboard may be extended to incorporate prediction functionality for USN staff and unit use. This will go far to help units plan for future operations and give administrative staffs the ability to detail fuel distribution for readiness and deployment purposes.

5.3 Areas for Future Research

5.3.1 CNAF Flight Hour Program (FHP)

Flight hours (which are analogous to fuel) are distributed to squadrons quarterly through the CNAF FHP. Distribution is based on expected readiness needs based on OFRP phase and Required Operational Capabilities/Predicted Operational Environment (ROC/POE) requirements (CNAP 2018). Assuming the underlying data is corrected for missing values, the Fuel Conservation Analytics Dashboard data could be used to refine readiness requirements through flight hour and fuel distribution in order to support non-operation DoN and NAE energy conservation goals.

5.3.2 NAVAIR Digital Initiatives

The Fuel Conservation Analytics Dashboard is but one effort to spearhead energy conservation efforts. However, several energy initiatives utilize the same flight databases to analyze conservation efforts. Missing data assessment from this study and prior research conducted by Barnhill et al. (2020) should be used to assess the effects of missing data in these parallel initiatives.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A:
Unique Sortie Data Set

A.1 Unique Sortie Data

This section defines the categories of the unique sortie data set.

master_matchstatus : Indicates which of seven categories the sortie falls into: "Triple Match" (227,682 sorties), "MU-NAVFLIRS" (55,241 sorties), "MU-SHARP" (20,207 sorties), "NAVFLIRS-SHARP" (105,916 sorties), "MU Only" (10,149 sorties), "NAVFLIR Only" (26,737 sorties), or "SHARP Only" (20,469 sorties).

master_launchicao: Determined launch International Civil Aviation Organization (ICAO) designator determined by examining NAVFLIR and SHARP records. In the absence of either of these, it is left blank. If both are available but are different, NAVFLIR is used.

master_landicao: Determined landing ICAO designator determined by examining NAVFLIR and SHARP records. In the absence of either of these, it is left blank. If both are available but are different, NAVFLIR is used.

master_route: Type of route chosen from four categories: "Land" (322,791 sorties), "Land to Ship" (3,715 sorties), "Ship" (125,966 sorties), and "Ship to Land" (3,780 sorties). If only MU data is available, the field is left blank (10,149 sorties).

master_airwing: The CVW to which the squadron flying the sortie is assigned. There are nine CVWs: "CVW-1," "CVW-2," "CVW-3," "CVW-5," "CVW-7," "CVW-8," "CVW-9," "CVW-11," or "CVW-17."

master_command: The TYCOM to which the squadron flying the sortie is assigned. These categories are "CNARF NAVY," "LANT NAVY," "NASC FS," and "PAC NAVY." If no TYCOM is assigned, then a "NULL" value is assigned.

master_ofrp : The OFRP cycle phase during which the sortie takes place. These phases are "Maintenance," "Basic," "Integrated," "Deploy," and "Sustainment." If OFRP phase is unassigned or only MU data is present, the field is blank. Only one phase should be assigned for each month of sorties for a particular squadron.

mu_buno: The unique BUNO belonging to the aircraft used for the sortie as recorded by the MU.

mu_squadron: The squadron responsible for the aircraft as recorded by the MU data.

mu_launchdate: The launch date and time recorded by the MU. All MU times are GMT.

mu_landdate: The land date and time recorded by the MU. All MU times are GMT.

mu_fuelstart: The starting fuel load recorded by the MU.

mu_fuelend: The ending fuel load recorded by the MU.

mu_fueldumped: The fuel dumped as recorded by the MU.

mu_fuelgiven: The cumulative fuel dispensed to other aircraft during tanking evolutions as recorded by the MU.

mu_fueltaken: The fuel received by an aircraft from tanker aircraft as recorded by the MU.

mu_fuelburned: The total fuel burned by an aircraft as recorded by the MU.

mu_fuelused: The total fuel used by an aircraft as recorded by the MU. Fuel used is the difference between the beginning fuel value and the ending fuel value and the sum of fuel given, fuel taken, and fuel dumped.

mu_fuelburnrate: The fuel burn rate recorded by the MU in pounds per hour.

mu_tms: The MU-recorded TMS. Assigns either F/A-18E, F/A-18F, or EA-18G.

configuration: Shows the ordnance or stores carried on each aircraft pylon. There are 11 pylons on the three TMS aircraft. Configuration fields show specific designations of the stores carried by the aircraft on the sortie. There are 10,200 configurations for F/A-18E aircraft, 6,988 configurations for F/A-18F aircraft, and 1,580 configurations for EA-18G.

unique_config_id: A numeric code assigned to the aircraft configuration. The code serves as a grouping code for a set of configurations as opposed to a unique identifier for a particular aircraft configuration. There are 8,647 unique identifiers for F/A-18E aircraft, 5,924 unique identifiers for F/A-18F aircraft, and 1,098 unique identifiers for EA-18G aircraft.

arresting_hook_flag: Indicates whether an arresting hook is deployed. This value is supposed to be binary, taking a "0" or "1" and is derived from the MU.

canopy_ladder_flag: Indicates whether a canopy or ladder is deployed. This value is supposed to be binary, taking a "0" or "1" and is derived from the MU.

dumpinfirst_flag: Indicates whether an aircraft dumped in the first half of the sortie. This value is supposed to be binary, taking a "0" or "1" and is derived from the MU.

dump_flag: Indicates whether an aircraft dumped during a sortie. This value is supposed to be binary, taking a "0" or "1" and is derived from the MU.

give_flag: Indicates whether an aircraft dispensed fuel to another aircraft during the sortie. This value is supposed to be binary, taking a "0" or "1" and is derived from the MU.

take_flag: Indicates whether an aircraft received fuel from another aircraft during the sortie. This value is supposed to be binary, taking a "0" or "1" and is derived from the MU.

single_engine_flag: Indicates whether an aircraft operates with one engine inoperative during the sortie. This value is supposed to be binary, taking a “0” or “1” and is derived from the MU.

nav_buno: Indicates the BUNO recorded on the NAVFLIR for the sortie.

nav_sqd: Indicates the squadron responsible for the specific aircraft maintenance recorded on the NAVFLIR for the sortie.

nav_tms: Indicates the TMS recorded on the NAVFLIR for the sortie.

nav_depart_date: Indicates the departure date and time as recorded on the NAVFLIR for the sortie. Typically, this is recorded in local time.

nav_arrival_date: Indicates the arrival date and time as recorded on the NAVFLIR for the sortie. Typically, this is recorded in local time.

nav_duration: The length of the sortie as computed by the recorded departure and arrival dates on the NAVFLIR for the associated sortie.

nav_depart_icao: The departure location ICAO designator of the sortie as recorded on the NAVFLIR for the associated sortie.

nav_arrival_icao: The arrival location ICAO designator of the sortie as recorded on the NAVFLIR for the associated sortie.

nav_tmr: The TMR code assigned to the sortie as recorded on the NAVFLIR. The Dashboard defaults to NAVFLIR if both SHARP and NAVFLIR are present.

nav_approaches: The number of instrument approaches conducted during a sortie as recorded on the NAVFLIR.

nav_approachcode: The type of instrument approach(es) conducted by the aircrew as recorded on the NAVFLIR. These codes are alphanumeric in nature taking the values 1-4 and A-C depending on whether the approach was actual (numeric) or simulated (alphabetic). Simulated approaches are those conducted for practice.

nav_landings: The number of landings conducted during the sortie as recorded on the associated NAVFLIR.

nav_landingcode: The type of landing conducted during the sortie as recorded on the associated NAVFLIR. These codes are numeric or alphabetic in nature depending on whether it is a daytime or nighttime landing. These codes take the values 1-7 for daytime and A-Z for nighttime. There are 52,259 missing or "NULL" codes.

sharp_buno: Indicates the BUNO recorded on the SHARP for the sortie.

sharp_launchdate: Indicates the departure date and time as recorded on the SHARP record for the sortie. Typically, this is recorded in local time.

sharp_landdate: Indicates the arrival date and time as recorded on the SHARP record for the sortie. Typically, this is recorded in local time.

sharp_launchicao: The departure location ICAO designator of the sortie as recorded on the SHARP record for the associated sortie.

sharp_landicao: The arrival location ICAO designator of the sortie as recorded on the SHARP record for the associated sortie.

sharp_sortieduration: The length of the sortie as computed by the recorded departure and arrival dates on the SHARP record for the associated sortie.

sharp_tms: Indicates the TMS recorded on the SHARP record for the sortie.

sharp_squadron: Indicates the squadron flying the associated aircraft recorded on the SHARP record for the sortie. In some cases, the squadron flying the aircraft will be different than the squadron responsible for the aircraft maintenance. This can lead to a difference between the NAVFLIR and SHARP records for the associated sortie.

sharp_fuelstart: Indicates the starting fuel as recorded by the pilot on the SHARP record.

sharp_fuelend: Indicates the ending fuel as recorded by the pilot on the SHARP record.

sharp_fuelused: The total fuel used by an aircraft as logged by the pilot on the SHARP record. Fuel used is the difference between the fuel start and end values and includes fuel given, fuel taken, and fuel dumped.

sharp_fuelburned: The total fuel burned by an aircraft as recorded by the pilot on the SHARP record.

sharp_fuelburnrate: The fuel burn rate as computed from the fuel burned over the duration of the sortie.

sharp_fueldumped: The amount of fuel dumped as recorded by the pilot on the SHARP record.

sharp_fuelgiven: The cumulative fuel dispensed to other aircraft during tanking evolutions as recorded by the pilot on the SHARP record.

sharp_fueltaken: The fuel received by an aircraft from tanker aircraft as recorded by the pilot on the SHARP record.

sharp_tmr: The TMR code assigned to the sortie as recorded on the SHARP record. The Dashboard defaults to NAVFLIR if both SHARP and NAVFLIR are present.

sharp_tmr1: The specific mission assigned to the sortie as recorded by the pilot on the SHARP record. This is the only place in the data where a mission is recorded.

There are 33 unique missions for F/A-18E aircraft, 34 unique missions for F/A-18F aircraft, and 23 unique missions for EA-18G aircraft. This includes the possibility of no assignment. There are 17,202 F/A-18E sorties with no mission assignment (8.2% of F/A-18E sorties), 13,312 F/A-18F sorties with no mission assignment (8.7% of F/A-18F sorties), and 74,771 EA-18G sorties with no mission assignment (78% of EA-18G sorties).

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B:
NAE Organization

B.1 NAE Organization 2016 to 2019

The sortie data provided by NAOEP spans from 2016 to 2019. This appendix describes the operational organization for F/A-18E/F and EA-18G squadrons during this time frame.

While fuel consumption analysis is of interest in all types of F/A-18 and EA-18G squadrons, fuel consumption for deploying, or operational, squadrons assigned to CVWs is given primacy of importance. Focus on operational squadrons is warranted for several reasons. First, flight hours are allocated to each squadron based on their readiness month as determined by the ROC/POE and OFRP phase (Department of the Navy 2014). Additionally, squadrons in the same CVW are in the same OFRP phase, allowing comparison between squadrons of like TMS. This also allows comparison of proportion of captured sortie data between squadrons within a CVW based on OFRP phase. Lastly, the squadrons assigned to a CVW deploy together, capturing details associated with a particular deployment, such as mission requirements.

B.1.1 NAE Squadron Types

NAE squadrons consist of three types: operational, reserve, and training. There are currently 34 operational F/A-18E/F squadrons with each squadron being assigned to a CVW. EA-18G operational squadrons are separated into squadrons assigned to CVWs and expeditionary squadrons. Expeditionary squadrons deploy individually to AORs as required by the CCDR. There are nine EA-18G CVW and three expeditionary squadrons.

Training squadrons are those squadrons that support the manning and training of aviation personnel. Typically, FRSs provide initial training to naval aviators who then proceed on to an operational squadron. There are three F/A-18E/F FRSs and one EA-18G FRS.

The last major type of squadron is the reserve squadron. These squadrons exist in the event more combat-ready aircraft and pilots are necessary to support operational requirements.

There is one reserve F/A-18E/F squadron and one EA-18G squadron in the NAE. Operational squadrons are assigned to CVWs and are the focus of this thesis, regardless of TMS.

B.1.2 Tactical Aircraft (TACAIR) Squadron Aircraft Inventory

This thesis focuses on USN TACAIR squadrons which fall into two categories: Strike Fighter (VFA) squadrons and Electronic Attack (VAQ) squadrons. The number of aircraft assigned to VFA squadrons belonging to a CVW is dependent on the particular TMS. VFA squadrons that fly the F/A-18E are made up of 10 to 12 single-seat aircraft. VFA squadrons that fly the F/A-18F are responsible for 12 two-seat aircraft. CVW VAQ squadrons are assigned four to six EA-18G aircraft (CNAF 2020).

B.1.3 Carrier Air Wing (CVW)

The largest operational aviation entity is the CVW. Each CVW is assigned to a nuclear powered aircraft carrier (CVN) for deployment, but over time, CVWs may switch to other CVNs depending on deployment and maintenance cycles (CNAF 2020).

The number of operational CVWs fluctuates depending on operational demand. Currently, there are nine CVWs constructed using multiple squadrons of several different aircraft platforms. For this data, the focus is on F/A-18E/F and EA-18G platforms. Table B.1 outlines the base of each CVW by coast. The coast assignment determines whether a CVW will be deployed on a CVN based on the east or west coast.

Table B.1. CVW by Coast Assignment.

CVW	Coast Assignment
CVW-1	East
CVW-2	West
CVW-3	East
CVW-5	Japan
CVW-7	East
CVW-8	East
CVW-9	West
CVW-11	West
CVW-17	West

Typically, a CVW will be made up of five squadrons of tactical aircraft (i.e., F/A-18E/F and EA-18G). Usually, there will be one F/A-18F squadron, one EA-18G squadron, and three F/A-18E squadrons. From 2016 to 2019, there were several squadrons that had not transitioned from the F/A-18C to the F/A-18E. Additionally, in some cases, CVWs were augmented with a USMC F/A-18C squadron, replacing one F/A-18E squadron. Sorties conducted in F/A-18C aircraft are not included in the unique sortie data set. Additionally, in very rare instances, two F/A-18F squadrons were assigned to a CVW to fill out the five-squadron requirement. So in most cases regarding F/A-18F and EA-18G squadrons, assessment of the TMS for that deployment suffices for assessment of the squadron itself. Tables B.2 and B.3 provide the inventory of squadrons by CVW deployment and year for the 2016 to 2019 time frame.

Table B.2. CVW Deployment Squadron Assignments 2016 to 2019.

Platform	**CVW-1**	**CVW-2**	**CVW-3**	**CVW-5**	**CVW-7 2016**	**CVW-7 2019**
F/A-18F	VFA-11	VFA-2	VFA-32	VFA-102	VFA-103	VFA-103
F/A-18E/F+	VFA-211+	VFA-137	VFA-86	VFA-27	VFA-143	VFA-25
F/A-18E	VFA-136	VFA-192	VFA-105	VFA-115	VFA-25	VFA-143
F/A-18E/C*	VFA-81	VFA-34*	VFA-131*	VFA-195	VFA-83*	VFA-86
EA-18G	VAQ-137	VAQ-136	VAQ-130	VAQ-141	VAQ-140	VAQ-140
+F/A-18F squadron. *F/A-18C squadrons not captured in the data set.						

Table B.3. CVW Deployment Squadron Assignments 2016 to 2019.

Platform	**CVW-8**	**CVW-9**	**CVW-11 2017**	**CVW-11 2019**	**CVW-17**
F/A-18F	VFA-213	VFA-41	VFA-154	VFA-154	VFA-22
F/A-18E	VFA-31	VFA-151	VFA-146	VFA-31	VFA-94
F/A-18E	VFA-87	VFA-97	VFA-147	VFA-146	VFA-113
F/A-18E/C	VFA-37	VFA-14	VMFA-323	—-	VMFA-312
EA-18G	VAQ-131	VAQ-133	VAQ-142	VAQ-142	VAQ-139
VMFA are U.S. Marine Corps F/A-18C squadrons not captured in the data.					

B.1.4 CVW Deployments

Most fuel consumption analysis through Air ENCON has focused on technological upgrades or process improvement. Little analysis has been conducted focusing on deployed CVWs and their associated squadrons. A CVW and its member squadrons will deploy to numbered fleet AORs. Figure B.1 shows the geographic headquarters of each numbered fleet.

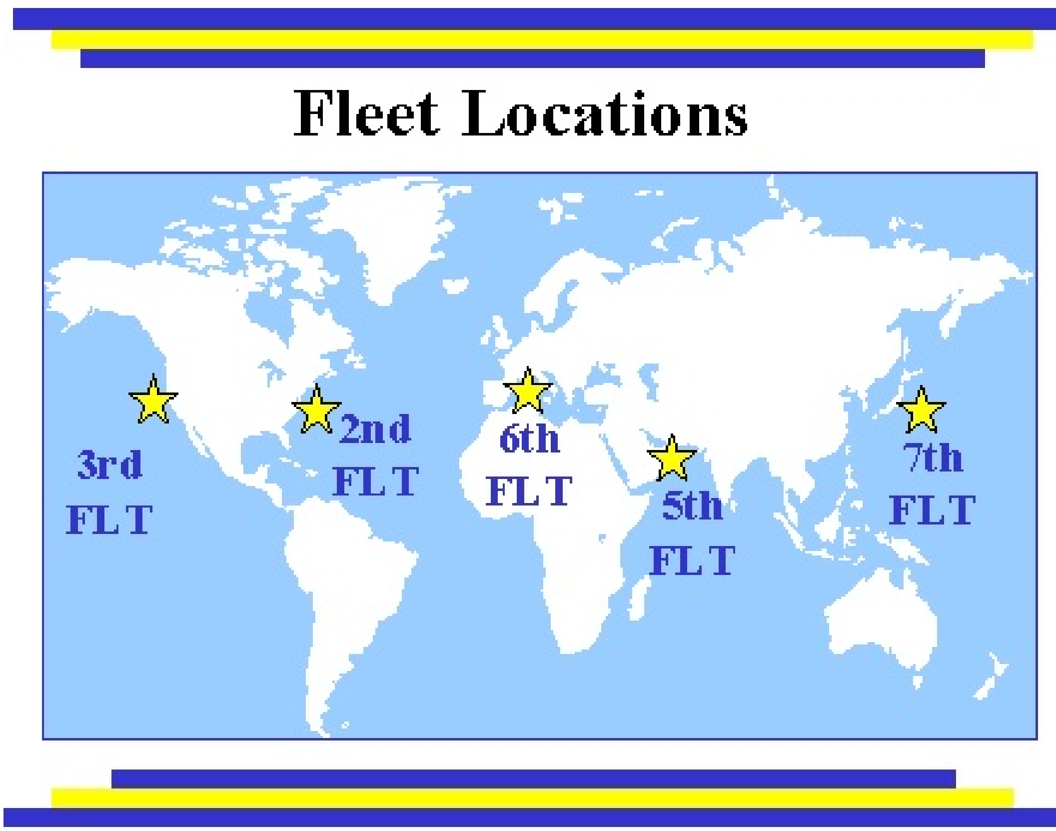

Figure B.1. U.S. Navy Numbered Fleet Areas of Responsibility (AOR). Source: Numbered Fleets (2020).

Typically, deployments will pass through multiple AORs. From 2016 to 2019 there were 19 aircraft carrier deployments with their associated CVWs. CVW deployment destinations are outlined in Table B.4, displaying deployments by year and location (Vasiljevic 2020).

Table B.4. CVW Deployments by Year and AOR.

Year	CVW	AOR
2016	CVW-3	5th/6th Fleet
2016	CVW-5	7th Fleet
2016	CVW-7	5th/6th Fleet
2016	CVW-9	7th Fleet
2017	CVW-2	7th Fleet
2017	CVW-5	7th Fleet
2017	CVW-8	5th/6th Fleet
2017	CVW-11	7th/5th Fleet
2017	CVW-17	7th/5th Fleet
2018	CVW-1	6th Fleet
2018	CVW-1	6th Fleet
2018	CVW-2	7th Fleet
2018	CVW-2	3rd Fleet
2018	CVW-5	7th Fleet
2018	CVW-9	7th/5th Fleet
2019	CVW-1	5th/6th Fleet
2019	CVW-5	7th Fleet
2019	CVW-7	5th/6th/7th Fleet
2019	CVW-11	3rd Fleet

B.1.5 Sortie Type

The underlying data supporting the Analytics Dashboard categorizes aircraft sorties in several ways. Aside from organization by squadron and CVW, the most important method of categorization is by OFRP phase. By using, these phases, an analyst can determine at which point in a squadron's readiness cycle a sortie occurs. Further, by examining sorties in this way, we can analyze fuel consumption by deployment. For the purposes of this section, "sortie type" is defined by OFRP phase. The following section describes the OFRP in detail.

Optimized Fleet Response Plan

In 2014 the USN instituted the OFRP with the aimed goal of maintaining continuous support of the CCDR. The underlying intent was to establish a sustainability cycle to meet CCDR requests without overloading any particular unit and associated Sailors. The OFRP consists of five phases: maintenance, basic, integrated, advanced, and sustainment (Department of the Navy 2014).

Maintenance Phase: The beginning of the OFRP cycle is the maintenance phase. This is the period when deployable units conduct major maintenance functions on assigned equipment. This maintenance may consist of equipment upgrades, modernization, in-depth periodic maintenance, or force reconstitution. Additionally, units may undergo required inspections or certifications necessary to ensure proper maintenance practices are being followed and material readiness achieved.

Basic Phase: The basic phase is the second phase in the OFRP. During this phase, deployable units exercise core competencies that are the underpinning of the more advanced phases. While readiness is higher than the maintenance phase, units continue to operate at lower levels of personnel, equipment, supply, and ordnance readiness than subsequent OFRP phases.

Integrated Phase: The integrated phase aggregates units that are going to be part of a specified, deployable group (e.g., Carrier Strike Group (CSG) or CVW) to complete advanced complex training requirements. The integrated phase may be tailored to specific CCDR requests. Once the integrated phase is complete, the aggregated group is certified to deploy.

Advanced Phase: The advanced phase is specific to independent deploying assets. This paper does not address the advanced phase because it does not apply to the units assigned to a CVW.

Sustainment Phase: The sustainment phase commences following the integrated or advanced phase. For the purposes of this paper, the sustainment phase is assumed to immediately follow the integrated phase. Once units are in the sustainment phase, they are considered deployable. Deployments occur within the sustainment phase. Deployments are followed by the unit remaining in the sustainment phase in the event there is a requirement for the unit to redeploy. In the event units must redeploy, they may be required to execute tailored training events to maintain readiness.

Deployment: Deployments occur within the confines of the sustainment phase. However, for the purposes of this paper, deployment will be considered a separate phase of the OFRP.

For most deployable units, the OFRP structure works cyclically as the name describes. They do not revert backwards to a previous phase. The lone exception are naval forces assigned to FDNF (i.e., CVW-5).

This paper is concerned with sorties conducted by squadrons assigned to deployed CVWs. Deployed sorties in the data set are those that fall within the "Deploy" phase of the OFRP cycle. Chapter 3 gives detail on processing methodologies used to redefine sortie type because of the severe errors in OFRP assignment.

APPENDIX C:
Missing Sortie Proportions

C.1 MU and Non-MU Sorties

Table C.1 provides the percentages of available MU data by deployment. Note that the 2019 CVW-1 deployment is only captured in non-MU data. This deployment occurred at the end of 2019 and continued into 2020. The CVW-11 deployment in 2019 was a one-month deployment for an exercise near Alaska. This appears to be the reason for the low sortie count. Tables C.2, C.3, C.4, and C.5 show the percentage of available MU data by deployment differentiated by TMS.

Table C.1. CVW Sortie Count by Deployment.

Year	CVW	Total Sorties	MU Sorties	% MU Sorties
2016	CVW-3	7,208	5,161	71.6%
2016	CVW-5	6,110	4,578	74.9%
2016	CVW-7	6,572	4,802	73.1%
2016	CVW-9	9,064	6,425	70.9%
2017	CVW-2	3,799	3,119	82.1%
2017	CVW-5	7,845	3,518	44.8%
2017	CVW-8	6,570	2,681	40.8%
2017	CVW-11	4,735	3,632	76.7%
2017	CVW-17	6,118	5,175	84.6%
2018	CVW-1	3,743	3,254	86.9%
2018	CVW-1	3,565	2,663	74.6%
2018	CVW-2	2,169	1,916	88.3%
2018	CVW-2	1,068	922	86.3%
2018	CVW-5	7,942	7,430	93.6%
2018	CVW-9	8,649	8,108	93.7%
2019	CVW-1	990	0	0%
2019	CVW-5	7,532	4,665	61.9%
2019	CVW-7	10,898	7,601	69.7%
2019	CVW-11	614	404	65.8%

Table C.2. F/A-18E Sortie Count by Deployment 2016-2017.

Year	CVW	Squadron	Total Sorties	MU Sorties	% MU Sorties
2016	CVW-3	VFA-105	2,220	1,859	83.7%
2016	CVW-3	VFA-86	1,691	1,271	75.2%
2016	CVW-5	VFA-115	1,242	1,200	96.6%
2016	CVW-5	VFA-195	1,232	1,062	86.0%
2016	CVW-5	VFA-27	1,410	715	50.7%
2016	CVW-7	VFA-143	1,963	1,487	75.8%
2016	CVW-7	VFA-25	1,582	1,193	75.4%
2016	CVW-9	VFA-14	1,956	1,560	79.8%
2016	CVW-9	VFA-151	1,987	1,518	76.4%
2016	CVW-9	VFA-97	1,938	1,206	62.2%
2017	CVW-2	VFA-137	1,738	1,520	87.5%
2017	CVW-2	VFA-192	0	0	0%
2017	CVW-5	VFA-115	1,615	441	27.3%
2017	CVW-5	VFA-195	1,693	905	53.4%
2017	CVW-5	VFA-27	1,921	1,074	55.9%
2017	CVW-8	VFA-31	2,152	1,165	54.1%
2017	CVW-8	VFA-87	1,493	276	18.5%
2017	CVW-11	VFA-146	721	489	67.8%
2017	CVW-11	VFA-147	1,693	1,176	69.5%
2017	CVW-17	VFA-113	1,906	1,794	94.1%

Table C.3. F/A-18E Sortie Count by Deployment 2018-2019.

Year	CVW	Squadron	Total Sorties	MU Sorties	% MU Sorties
2018	CVW-1	VFA-136	856	799	93.3%
2018	CVW-1	VFA-81	934	828	88.7%
2018	CVW-1	VFA-136	774	305	39.4%
2018	CVW-1	VFA-81	939	836	89.0%
2018	CVW-2	VFA-137	816	755	92.5%
2018	CVW-2	VFA-192	142	20	14.1%
2018	CVW-2	VFA-137	376	345	91.7%
2018	CVW-2	VFA-192	80	34	42.5%
2018	CVW-5	VFA-115	1,430	1,365	95.5%
2018	CVW-5	VFA-195	1,706	1,651	96.8%
2018	CVW-5	VFA-27	1,854	1,619	87.3%
2018	CVW-9	VFA-14	2,203	2044	92.8%
2018	CVW-9	VFA-151	1,700	1,588	93.4%
2018	CVW-9	VFA-97	1,695	1,594	94.0%
2019	CVW-1	VFA-136	333	0	0%
2019	CVW-1	VFA-81	325	0	0%
2019	CVW-5	VFA-115	1,352	799	59.1%
2019	CVW-5	VFA-195	1,433	899	62.7%
2019	CVW-5	VFA-27	1,837	1,343	73.1%
2019	CVW-7	VFA-143	2,777	1,887	68.0%
2019	CVW-7	VFA-25	2,178	1,654	75.9%
2019	CVW-7	VFA-86	2,017	1,367	67.8%
2019	CVW-11	VFA-146	129	13	10.1%
2019	CVW-11	VFA-31	241	219	90.9%

Table C.4. F/A-18F Sortie Count by Deployment.

Year	CVW	Squadron	Total Sorties	MU Sorties	% MU Sorties
2016	CVW-3	VFA-32	2,317	1,773	76.5%
2016	CVW-5	VFA-102	1,502	1,077	71.7%
2016	CVW-7	VFA-103	2,333	1,446	70.0%
2016	CVW-9	VFA-41	2,094	1,348	64.4%
2017	CVW-2	VFA-2	1,685	1,286	76.3%
2017	CVW-5	VFA-102	1,881	848	45%
2017	CVW-8	VFA-213	2,183	1,188	54.4%
2017	CVW-11	VFA-154	1,523	1,279	84.0%
2017	CVW-17	VFA-22	2,042	1,818	89.0%
2017	CVW-17	VFA-94	1,252	816	65.2%
2018	CVW-1	VFA-11	853	759	89.0%
2018	CVW-1	VFA-211	1006	774	76.9%
2018	CVW-1	VFA-11	842	737	93.0%
2018	CVW-1	VFA-211	978	764	78.1%
2018	CVW-2	VFA-2	818	761	93.0%
2018	CVW-2	VFA-2	368	309	84.0%
2018	CVW-5	VFA-102	1,854	1,724	93.0%
2018	CVW-9	VFA-41	2,064	1,939	93.9%
2019	CVW-1	VFA-11	149	0	0%
2019	CVW-1	VFA-211	183	0	0%
2019	CVW-5	VFA-102	1,782	874	49.0%
2019	CVW-7	VFA-103	2,640	1,767	66.9%
2019	CVW-11	VFA-154	141	138	97.9%

Table C.5. EA-18G Sortie Count by Deployment.

Year	CVW	Squadron	Total Sorties	MU Sorties	% MU Sorties
2016	CVW-3	VAQ-130	980	258	26.3%
2016	CVW-5	VAQ-141	724	524	72.3%
2016	CVW-7	VAQ-140	694	676	97.4%
2016	CVW-9	VAQ-133	1,089	793	72.8%
2017	CVW-2	VAQ-136	376	313	83.0%
2017	CVW-5	VAQ-141	735	250	34.0%
2017	CVW-8	VAQ-131	742	52	26.3%
2017	CVW-11	VAQ-142	798	688	86.2%
2017	CVW-17	VAQ-139	918	747	81.3%
2018	CVW-1	VAQ-137	94	94	100%
2018	CVW-1	VAQ-137	31	20	64.5%
2018	CVW-2	VAQ-136	393	380	96.6%
2018	CVW-2	VAQ-136	206	199	96.6%
2018	CVW-5	VAQ-141	1,098	1,071	97.5%
2018	CVW-9	VAQ-133	987	943	95.6%
2019	CVW-5	VAQ-141	1,129	750	66.4%
2019	CVW-7	VAQ-140	1,286	926	72.0%
2019	CVW-11	VAQ-142	103	34	33.0%

THIS PAGE INTENTIONALLY LEFT BLANK

List of References

Barnhill D, Morgan B, Buttrey S, Whitaker L, Nussbaum D (2020) Evaluating NAVAIR's fuel dashboard. Unpublished technical report, Naval Postgraduate School, Monterey, CA.

Bennett D (2001) How can I deal with missing data in my study? *Aust N Z J Public Health* 25(5), https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1467-842X.2001.tb00294.x.

Breiman L (2001) Random forests. *Machine Learning* 45(1):5–32, ISSN 0885-6125, http://dx.doi.org/10.1023/A%3A1010933404324.

CNAF (2016) CNAF M-3710.7 Naval air training and operating procedures standardization (NATOPS) program. https://www.secnav.navy.mil/doni/opnav.aspx.

CNAF (2017) OPNAV 4790.2C Naval aviation maintenance program (NAMP). https://www.navair.navy.mil/documents?name=field_document_description_value=NAMP.

CNAF (2020) Squadron CVW assignment 2016-2019, data provided to the author via email, August 30, 2020.

CNAP (2018) CNAPINST 3500.1A Squadron training and readiness. https://timesheet.mysabre.com/elearning/supportfiles/xls/cnapcnal3500/CNAP_CNAL%203500.1A%20_19Dec2018.pdf.

Daxue Consulting (2009) Brief introduction to statistics. Accessed February 27, 2021, https://bookdown.org/thomas_pernet/Tuto/methods-for-describing-a-set-of-data.html.

Deloitte Analytics Team (2019) F/A-18 Super Hornet data analysis. Data provided to NPS, March 30, 2020.

Deloitte Analytics Team (2020) Air ENCON F/A-18 data dictionary, data provided to NPS, January 15,2020.

Department of the Navy (2010a) Department of the Navy's energy program for security and independence. Accessed December 6, 2020, https://www.secnav.navy.mil/eie/Documents/DoNEnergyProgramforSecurityandIndependence.pdf.

Department of the Navy (2010b) A Navy energy vision for the 21st century. Accessed December 8, 2020, https://www.hqmc.marines.mil/Portals/160/Docs/Navy-Energy-Vision-Oct-2010.pdf.

Department of the Navy (2010c) Task Force Energy (TFE) charter. Accessed December 6, 2020, https://navysustainability.dodlive.mil/files/2010/10/TFE_Charter.pdf.

Department of the Navy (2014) OPNAV 3000.15A Optimized fleet response plan (OFRP). Accessed December 8, 2020, https://www.secnav.navy.mil/doni/Directives/03000%20Naval%20Operations%20and%20Readiness/03–00%20General%20Operations%20and%20Readiness%20Support/3000.15A.pdf.

Dong Y, Peng C (2013) Principled missing data methods for researchers. *SpringerPlus* 2(222), https://doi.org/10.1186/2193-1801-2-222.

Eger R, Albright T, Hansen S, Blades B, Bondurant E, Clark M, Donahue L, Farrell L, Rosborough D, Zintac G (2016) Aviation fleet training efficiency metrics and hot pit analysis. Unpublished white paper, Naval Postgraduate School, Monterey, CA.

Jakobsen J, Gluud C, Wettersley J, Winkel J (2017) When and how should multiple imputation be used for handling missing data in randomised clinical trials – A practical guide with flowcharts. *BMC Med Res Methodol* 17(162), https://doi.org/10.1186/s12874-017-0442-1.

James G, Witten D, Hastie T, Tibishrani R (2017) *An Introduction to Statistical Learning,* 7th ed. (Springer, New York, NY).

Kang H (2013) The prevention and handling of the missing data. *Korean Journal of Anesthesiology* 64(5), https://10.4097/kjae.2013.64.5.402.

Leurent B, Gomes M, Faria R, Morris S, Grieve R, Carpenter J (2018) Sensitivity analysis for not-at-random missing data in trial-based cost-effectiveness analysis: A tutorial. *Pharmacoeconomics* 36(8), https://doi.org/10.1007/s40273-018-0650-5.

NAVAIR (2019) Aviation energy dashboard, data provided to the author via email, January 2020.

NAVAIR (2020) Dashboard data sources, data provided to the author via email, August 2020.

Numbered Fleets (2020) *Global Security.* Accessed December 10, 2020, https://www.globalsecurity.org/military/agency/navy/fleet_n.htm.

Olszewski M, Quinn D, Noel B (2012) Navy air energy conservation (Air ENCON) program. NDIA Environment, Energy Security, and Sustainability (E2S2), New Orleans, LA. https://apps.dtic.mil/dtic/tr/fulltext/u2/a565706.pdf.

R Core Team (2017) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/.

Rubin D (1976) Inference and missing data. *Biometrika* 6(3), http://www.jstor.org/stable/2335739.

Salem A, King C, Fox S, Haley R, Klotzbach M (2009) Best practices in the Navy's energy programs : Strategic communication factors operating in the tactical forces. https://calhoun.nps.edu/handle/10945/25357.

Schafer J, Graham J (2002) Missing data: Our view of the state of the art. *Psychological Methods* 7(2), https://doi.org/10.1037/1082-989X.7.2.147.

Sprent P, Smeeton N (2007) *Applied Nonparametric Statistical Methods,* 4th ed. (Kinston-Slalom, ID: Chapman Hall/CRC, Boca Raton, FL).

Stekhoven DJ (2011) Using the "missForest" package. Accessed December 30, 2020, https://stat.ethz.ch/education/semesters/ss2012/ams/paper/missForest_1.2.pdf.

Stekhoven DJ, Buhlmann P (2011) MissForest—Non-parametric missing value imputation for mixed-type data. URL http://dx.doi.org/10.1093/bioinformatics/btr597.

Tabachnick BG, Fidell LS (2013) *Using Multivariate Statistics,* 6th ed. (Pearson Education, Inc., Upper Saddle River, NJ).

Tableau Software, Inc (2020) Tableau Desktop Professional Edition. Http://www.tableau.com.

Van Buuren S (2018) *Flexible Imputation of Missing Data,* 2nd ed. https://stefvanbuuren.name/fimd/.

Vasiljevic B (2020) U.S. Carriers. Accessed December 6, 2020, http://www.uscarriers.net/index.htm.

Wright MN, Ziegler A (2017) ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software* 77(1), ISSN 1548-7660, http://dx.doi.org/10.18637/jss.v077.i01.

THIS PAGE INTENTIONALLY LEFT BLANK

Initial Distribution List

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California