DEVELOPING AN AGGREGATE MARGINAL COST PER FLYING HOUR

MODEL FOR THE U.S. AIR FORCE'S F-15 FIGHTER AIRCRAFT

THESIS

Patrick D. Armstrong, Captain, USAF

AFIT/GCA/ENV/06M-01

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GCA/ENV/06M-01

DEVELOPING AN AGGREGATE MARGINAL COST PER FLYING HOUR MODEL FOR THE U.S. AIR FORCE'S F-15 FIGHTER AIRCRAFT

THESIS

Presented to the Faculty

Department of Systems Engineering and Management

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Cost Analysis

Patrick D. Armstrong, BS

Captain, USAF

March 2006

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

AFIT/GCA/ENV/06M-01

DEVELOPING AN AGGREGATE MARGINAL COST PER FLYING HOUR MODEL FOR THE U.S. AIR FORCE'S F-15 FIGHTER AIRCRAFT

Patrick D. Armstrong, BS

Captain, USAF

Approved:

____/signed/__________________________	21 March 2006
Michael J. Hicks (Chairman)	Date
____/signed/__________________________	21 March 2006
Curtis G. Tenney (Member)	Date
____/signed/__________________________	21 March 2006
William K. Stockman (Member)	Date

AFIT/GCA/ENV/06M-01

Abstract

Flying operations comprise 49% of the US Air Force readiness budget. Current forecast models of the Cost Per Flying Hour (CPFH) program suffer significant errors. These errors are as high as 25% of total annual cost, which is equivalent to the entire US Air Force space budget. These forecast errors place considerable budgetary and operational readiness risk on the US Air Force.

This research presents a new forecasting method for high frequency cost estimation of base level Cost Per Flying Hour. Using a balanced panel of base level, monthly data on Depot Level Reparables and Consumables for all active duty F-15s and their variants, this thesis presents a stochastic forecast, simulation and analytical model. This model is a fixed effect, time series cross sectional model with seasonal autoregressive elements (monthly binary variables) and a standard white-noise error term.

This model incorporates factors identified as prime contributors to CPFH. These include base/month mean temperature spread (with a salinity control included in the base fixed effect coefficient), programmatic and policy changes, economic estimates of cost changes embodied in the producer price index and aviation fuel costs. I also include a wartime variable (permitting forecast simulation over alternative deployment schedules), mean flight time duration (both combat and training operations) and average aircraft age at each installation.

While the results of these estimates are important contributions to our understanding of the dynamics of the CPFH program, the major contribution of this

research is in the dramatic improvement over existing models. The root mean squared errors from the out of sample forecast period generated by the models presented in this research improve upon the existing models from 77% to 99%.

AFIT/GCA/ENV/06M-01

To my wife and two daughters, the lovely ladies in my life.

You three are my reason for being.

Acknowledgments

First and foremost, I would like to thank God for giving me the wisdom and perseverance to complete this enormous task. I need to be especially thankful for the most precious things in my life; my wife and two daughters. Their undying devotion, understanding, and encouragement kept me focused on the goal. I also need to express my sincere appreciation for the support and encouragement my family and friends provided throughout the last eighteen months.

I would like to express my sincere appreciation to my faculty advisor, Dr Hicks. His insight, experience, patience and tutelage were certainly appreciated and much needed. Thank you for pushing me to do greater and better things. I would also like to thank my readers, LtCol Tenney and Dr Stockman, for their feedback and support; it is nice to have a perspective from outside the "weeds", when the "weeds" is all you can see.

There are several individuals and their respective agencies I need to thank for their efforts. First, Tom Lies, Larry Klapper, and Gary McNutt, from the Air Force Cost Analysis Agency, for their endeavor in getting me the monthly data I needed and sponsoring this thesis. Next, I am very grateful to Technical Sergeant Kevin Wendt from the Air Force Combat Climatology Center. He is a true testament and steward of the Noncommissioned Officer Corps. Lastly, I need to thank Mark Gossett from Battelle Corporation for his drive and initiative in coordinating the data gathering process. Without Mr. Gossett, the endeavor could have been more daunting.

My final thanks go to my classmates—thank you for always being there for me when I needed you and picking me up when I fell down.

Patrick D. Armstrong

.

Table of Contents

Page

Page

List of Figures

List of Tables

Page

DEVELOPING AN AGGREGATE MARGINAL COST PER FLYING HOUR MODEL FOR THE U.S. AIR FORCE'S F-15 FIGHTER AIRCRAFT

I. Introduction

"It is better to be vaguely right than precisely wrong."

John Maynard Keynes

Since the Civil War, the U.S. Government has tried accurately to predict the cost of war, and in every instance, predictions have fallen short of actual expenditures (Nordhaus, 2002:2). For example, U.S. Government estimates of the federal expenditures for the Civil War were estimated to be $240 million, when in fact; costs exceeded $3,200 million (Nordhaus, 2002:2). Similarly, early estimates for the Vietnam War were under estimated by approximately 90% (Nordhaus, 2002:2). In addition to inaccurate forecasts of conflict costs, the Department of Defense (DoD) also faces issues with forecasting steady state requirements. As an example, the U.S. Air Force (USAF) under estimated the cost per flying hour program (CPFH) Program by an aggregate of $850M in 1997 and 1998 (GAO, 1999:3). As a result, the USAF had to solicit the U.S. Congress for additional funds to maintain aircraft and pilot mission capability; otherwise, U.S. war fighting capability and air dominance were at risk. These issues of inaccurate forecasting of conflict costs estimates and steady state requirements are further compounded by a seemingly convoluted budgeting process, as evidenced in the following excerpt:

> The flying hour requirement in the budget does not include flying in support of contingency operations…However, hours flown in support of contingency operations are counted against the programmed hours already funded in the President's budget up to the number of hours an aircraft would have flown at its home station. For additional hours flown, the Air Force receives additional funding from a centrally managed Department of Defense (DoD) contingency account. (GAO, 1999:4)—

The ability to forecast accurately starts at the lowest level possible; this is the wing/base level in the USAF. If these low-level estimates are inaccurate, then the associated error rates of aggregated estimates will increase as the initial estimates have to go through additional layers of "forecasts" at the MAJCOM and the Air Force Cost Analysis Improvement Group (AFCAIG). This is especially evident if subsequent echelons use similarly poor forecasting models. Therefore, it is paramount that the analysts at the lowest organizational levels have the necessary tools to perform the robust analyses needed to provide accurate forecasts.

Background

In recent years, the Operating and Maintenance (O&M) portion of the President's budget has been growing at about 4% per year, while the number of aircraft, number of hours flown, and number of maintenance personnel have been decreasing (GAO, 2000:1). A significant portion of the O&M budget is the CPFH Program. The CPFH program is 6.4% of the FY07 USAF budget (Faykes, 2006:22), as depicted in Figure 1. The CPFH Program is comprised of three major cost drivers or factors: depot level Reparable (DLR), consumables (CONS), and Aviation fuel (AVFUEL), with DLRs being the most significant cost driver. The DLR and CONS portions of the CPFH program, as found by the Air Force Cost Analysis Improvement Group (AFCAIG), increased by over 9.7% from FY96 through FY00 (Kammerer, 2002:19). This large increase caused several MAJCOMs to request supplemental funding to maintain their wartime readiness (Kammerer, 2002:19). The AFCAIG is the agency responsible, with inputs from the Major Commands (MAJCOMS), for the development of the CPFH rates used. These rates are developed for each Mission Design Series (MDS) by MAJCOM. As a result, each aircraft type (i.e. F-16CD, F-15CD, F-15E) has a unique set of CPFH rates for each MAJCOM, creating difficulties in trying to forecast a CPFH rate for an aggregate MDS.

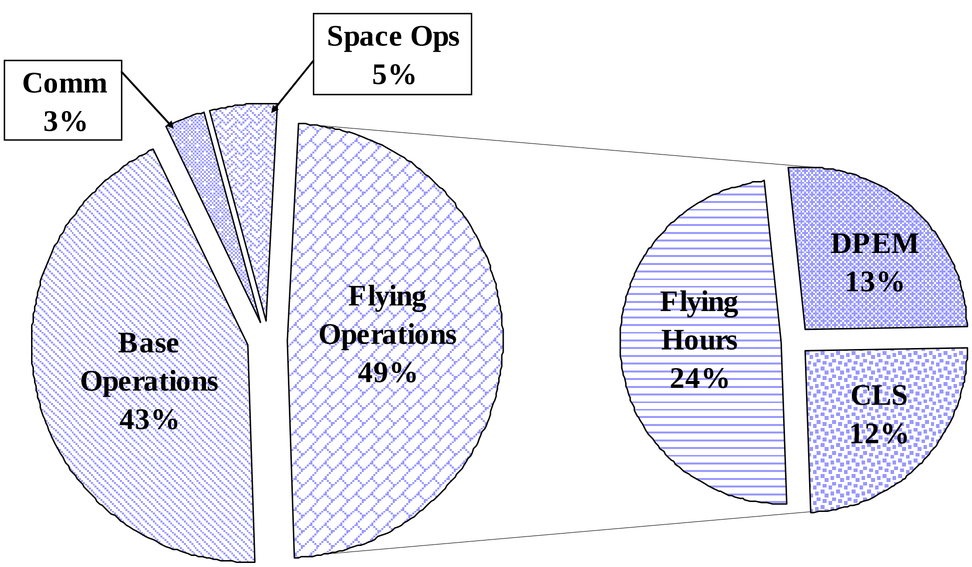


Figure 1. FY07 Air Force Budget

Source: FY07 Air Force Budget, PowerPoint Presentation, Major General Frank Faykes, Director AF Budget, 2006

Problem Statement

The goal of this research, sponsored by the AFCAIG, was to find a "marginal CPFH" rate such that if a Command flies in excess of its programmed baseline (PB) direct hours, the additional funding to pay for contingencies etc. is commensurate with the additional (marginal) cost for the extra hours flown, not the full value of a flying hour for that weapon system. This research sought to develop this "marginal" rate using an aggregate modeling method—panel data. To meet this goal, specific research questions were developed and are presented in the following section.

Research Questions/Objectives

The following objectives and research questions were addressed in the body of this thesis:

1. Primary Objective:

    - To develop an accurate, flexible, defensible, and easily used forecast model for the marginal CPFH of the F-15 fleet for all USAF active duty bases, MAJCOMs, and AFCAIG to use.

2. Research Questions:

    1. Can an aggregate model be developed for the entire F-15 fleet by Mission Design Series, or are the predictors base specific?
    2. Is a seasonal trend/business cycle in the CPFH rates for the F-15 fleet?
    3. Do the monthly average temperatures and salinity at a location influence the F-15 fleet CPFH rates?

4. Does the average age of the aircraft have an effect on the F-15 fleet CPFH rates?
5. Does the average sortie duration have an effect on the F-15 fleet CPFH rates?

Purpose

"Each of our Communities shares a common goal: to produce credible and defendable estimates to keep our aircraft flying." (Kammerer, 2002:19). With the flying hour program comprising a major portion of a base's budget, it is vital that these estimates be "accurate and defendable." By providing a model that can accurately estimate the depot level Reparable and consumable portions of the CPFH program, this research provides the base or wing commander an indispensable tool for budget management. As previously indicated, accuracy improvements at the lowest level should carry forward to the MAJCOM and Air Force levels. Therefore, the development of this model can benefit the entire USAF.

Research Focus

This research analyzed cost per flying hour (CPFH) data from all the USAF's F-15C, D, and E bases. The monthly data was aggregated from many different sources, to include Air Force Total Ownership Costs (AFTOC); Reliability and Maintainability Information System (REMIS); and the Air Force Combat Climatology database (AFCCC). These databases contain economic, operational, climatic, and programmatic data for all Air Force MDSs from 1998-2006 (Hawkes, 2005:6). The time frame being analyzed was FY01 through FY05. In addition, the development of the CPFH model was limited to the depot level Reparable (DLR) and consumable (CONS) portions of the CPFH rate. The models were developing using panel model techniques which allows for temporal and cross-sectional data analyses.

II. Literature Review

Chapter Overview

The purpose of this chapter is to describe the programmatic detail and existing research on cost per flying hour. First, a brief summary of the evolution of the F-15 fighter aircraft through its major Mission Design Series (MDS) changes will be provided followed by a discussion of the DoD's Planning, Programming, Budgeting and Execution (PPBE) System with an emphasis on the development of the CPFH factors. A literature review related to research on the prediction of CPFHs and O&M costs will then be presented. Finally, previous research relevant to the selection of the additional independent variables to include, temperature, salinity, and retail aviation fuel prices, used in this research will be offered.

F-15 History

Beginning

In response to a growing threat from the Soviet Union's development of the MiG-25 Foxbat fighter, the USAF needed to design a new aircraft to counter this superior threat, leading to the birth of the F-15 Eagle (King & Massey, 1997:10). On 23 December 1969, the USAF awarded McDonnell Douglas the F-15 contract. The F-15 is still the Air Force's principal air superiority and interdiction platform--it has survived 96 combat "dog fights" without losing a single aircraft (King & Massey, 1997:10).

F-15 A/B Eagle

The initial configuration of the F-15 had its first flight on 27 July 1972. "The F-15A was a single seat model and the F-15B is a two seat model" (King & Massey, 1997:11). There were over 360 F-15 A/B delivered to the Air Force with the 1st Tactical Fighter Wing (TFW) at Langley AFB, Virginia, being the very first operational F-15 combat wing. Today, the Air National Guard units in Florida, Louisiana, Missouri, and Oregon are flying the majority of the remaining F-15A/B models (King & Massey, 1997:11).

F-15 C/D Eagle

In June 1979, the next evolution of the F-15 emerged. The newer model had a larger internal fuel capacity (2,000 lbs. greater) and was capable of carrying conformal fuel tanks. The Multi-Stage Improvement Program (MSIP) phased in additional upgrades from 1985-1997.

These upgrades included, "structural, radar, and electronic warfare upgrades, along with wiring needed to employ the advanced medium range air-to-air missile (AMRAAM)" (King & Massey, 1997:11). A total of 470 F-15 C/Ds (408 F-15C single-seat and 62 F-15D two-seat) were accepted by the USAF. These aircraft are currently based at Eglin AFB, Florida; Elmendorf AFB, Alaska; Kadena AB, Japan; RAF Lakenheath, United Kingdom; Langley AFB, Virginia; and Mountain Home AFB, Idaho (King & Massey, 1997:11).

F-15 E Strike Eagle

The latest version of the F-15 is the E model. This version was built to fulfill the role of the Dual Role Fighter (DRF)—having the ability to perform precision strike missions on its own and air-to-air interdiction. On 11 December 1986, the first F-15E (two-seat) flew its maiden flight. It is very similar to the F-15D except "the aircraft is optimized for air-to-ground missions" (King & Massey, 1997:11). The modifications to achieve this new role included a stronger airframe, usage of conformal tanks, employment of a weapon's systems officer (WSO), and, most importantly, upgraded avionic systems (King & Massey, 1997:11). The upgrades to the avionics were "an improved radar for air-to-ground targeting; a two pod system for high speed, all-weather low level flight and targeting called Low Altitude Navigation and Targeting Infrared for Night (LANTIRN); and enhanced cockpit instrumentation" (King & Massey, 1997:12).

Although the Air Force has accepted the last F-15E it contracted for, the assembly lines have remained intact due to the Saudi Arabian and Israeli governments purchasing Foreign Military Sales (FMS) versions of the F-15E. The 225 F-15Es purchased by the Air Force are currently assigned to Eglin AFB; Elmendorf AFB; RAF Lakenheath; Mountain Home AFB; Nellis AFB, Nevada; and Seymour Johnson AFB, South Carolina (King & Massey, 1997:12).

PPBE System

Overview

"The ultimate objective of the DoD PPBS [PPBE] is to provide the best mix of forces, equipment and support attainable within fiscal constraints" (DoD, 1984:2). This objective is attained through the careful planning and execution of the PPBE process. The key output of the PPBE process is the Future Years Defense Program (FYDP) which summarizes all programs approved by the Secretary of Defense (SECDEF) for the DoD. The FYDP consists of budget and personnel information about the prior year, current year, the biennial budget years, and the following four years. It is the current and biennial budget year with which this research is concerned; as these are the years impacted the most by the CPFH factors developed by the Air Force Cost Analysis Improvement Group (AFCAIG).

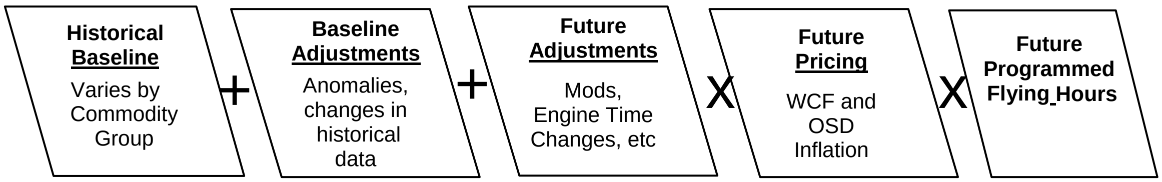


Figure 2. Graphical Representation of CPFH Factors, (SAF/FMC, 2005:18)

Source: SAF/FMC, "FY07 APOM Action Officer Cost Per Flying Hour (CPFH) Air Force Cost Analysis Improvement Group Tutorial", Electronic Message, Jan 05

Development of CPFH Factors

The AFCAIG develops the following variable direct flying hour cost factors for each MAJCOM and each aircraft type:

1. Reparable flying spares—Material Support Division (MSD)/Depot Level Reparable (DLR)
2. Consumable supplies—General Support Division (GSD)
3. Consumable supplies outside GSD—Government Purchase Card purchases (GPC)
4. Aviation fuel—AVFUEL

These four factors combined provide the total CPFH rate. Next is a brief description of each factor and how the AFCAIG, with the MAJCOMS input, develops them.

DLR

DLRs are aircraft parts, when removed, that are sent to depot to be repaired; however, the home unit's maintenance facility has the capability to repair a few of these (Rose, 1997:5). Generally, these are expensive parts (approximately 64% of the CPFH rate) and sometimes referred to as "black boxes." The Spares Requirements Board (SRRB) uses eight quarters of historical data to develop the DLR factor sent to each MAJCOM. The MAJCOMs take this factor and adjust it for expected future changes. AFCAIG reviews the MAJCOMs adjusted DLR factors before applying inflation adjustments (SAF/FMC, 2005:18).

GSD

GSD items are parts/supplies that have no authorized repair procedures (e.g., they are disposable parts or supplies) (Rose, 1997:4). The MAJCOMs develop the GSD factor using prior year obligations divided by actual flying hours. Again, adjustments are made to the factor for known changes (e.g. warranty expirations, modifications, time compliance technical orders, etc.) AFCAIG reviews the MAJCOMs adjusted GSD factors before applying inflation adjustments. Due to using obligations from three years prior to develop this factor, the factor will experience an adjustment one year prior to the money being obligated (SAF/FMC, 2005:20).

GPC

GPC items are the same as GSD items except GPC part/supplies are not purchased through government channels (e.g. local hardware store purchase, cleaning supplies, etc.) GPC factors are developed using the same method as GSD parts/supplies.

AVFUEL

AVFUEL is defined as fuel (JP-4, JP-8, off-station fuel and in-flight refueling) used during flight (Rose, 1997:5). The AVFUEL baseline is a rolling average of the previous three-years actual consumptions stated in terms of gallons per flight hour. This estimation, based on DoD estimated prices, has been relatively accurate and has limited problems (SAF/FMC, 2005:16). Therefore, this research will not investigate this component of the CPFH model.

Related Research

Much of the research on developing CPFH factors/models has centered around the Component (AF, Army, Navy) level and/or CPFH factors for other than fighter aircraft. As with this research, previous analysis was based on a large, macro level picture. Alternatively, this research will investigate the capability of building an aggregate model for the F-15 fleet by MDS that can also be applicable to a base level program. The previous research identified numerous deterministic/causational variables that this research will use in the analysis and development of the F-15 CPFH models. This research will add economic, climatic, and seasonal variables to further the research into obtaining valid predictor models of the CPFH rate. The following paragraphs will summarize the previous research that has been done on CPFH factors and will also present this summary in a table. The first research to be summarized is the thesis written by Hawkes (2005).

Hawkes (2005) used both programmatic and operational explanatory variables to predict the DLR rates for the F-16 C/D. The basis behind only looking at the DLR costs stems from approximately 65% of the total CPFH rate is attributable to DLRs. Hawkes evaluated nine variables, to include aircraft age, average sortie duration (ASD), MAJCOM, base location, utilization rate, percent engine type, percent block modification, percent deployed, and the previous year's CPFH rate. Of these nine variables, only percent engine type, percent block modification, percent deployed, and the previous year's CPFH rate, had not been investigated by previous research. The sample set for the thesis was all active duty and Air National Guard bases that flew the F-16 C/D. The data for this research was obtained from the Air Force Total Ownership Costs database and Air Force Knowledge System (AFKS) as is much of the data for this research.

In the initial analysis of the data and the correlation of the independent variables, Hawkes (2005) concluded that three variables, average sortie duration, utilization rate, and percent deployed, were significantly correlated. The scatterplot of these three variables, along with the correlation matrix, are displayed in Figure 6. Hinkle, Wiersma, and Jurs (1982) provided the following framework to interpret the correlation between variables: (a) very high (0.90-1.00); (b) high (0.70-0.90); (c) moderate (0.50-0.70); (d) low (0.30-0.50); and (e) little if any correlation (0.00-0.30) (Hinkle et al, 1982:100). As depicted in Figure 6, high correlation between these three variables, first through the correlation coefficient being greater than 0.50 for each pair of variables, indicating each of these variables is correlated (Hinkle et al, 1982:100). Second, the scatterplots in Figure 6 indicate each pair of the three variables has a linear relationship, suggesting correlation. These findings motivate the use of average sortie duration and its components of average training sortie duration and average combat sortie duration.

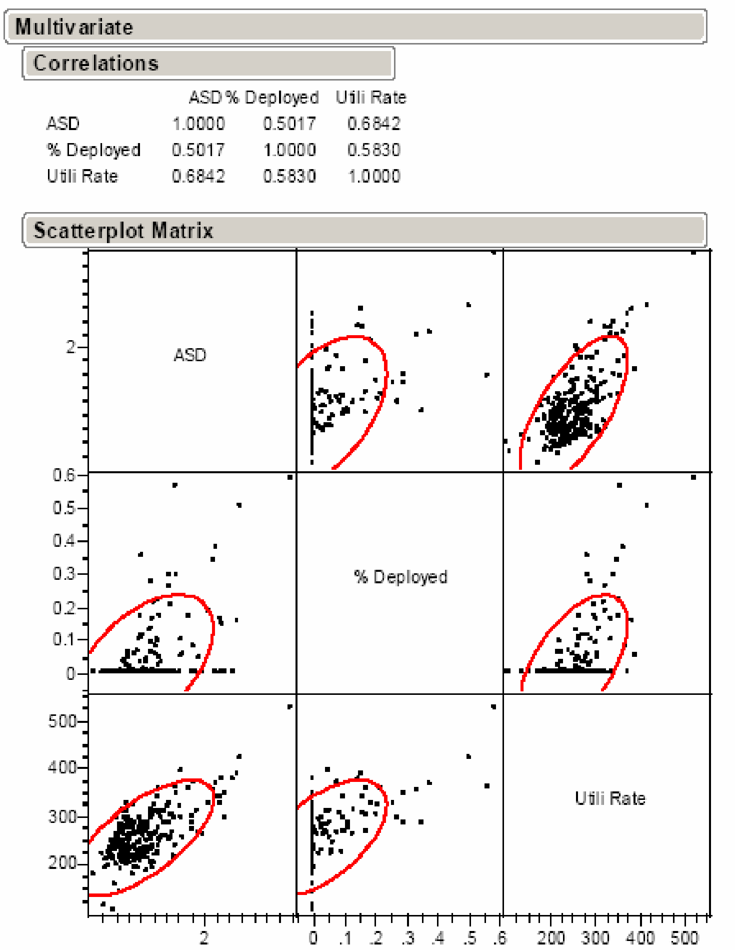


Figure 3. Scatterplot Matrix of Correlated Variables (Hawkes, 2005:40)

One-way analysis of variances (ANOVA) on all of its explanatory variables were then computed. From the ANOVA analysis it was determined that the most significant variables were lag 1 CPFH, average sortie duration (ASD), engine type, block, MAJCOM, and base. Although this test does not take into account the interactions of the variables, it represents an initial starting point for the analysis (Hawkes, 2005:44). In lieu of using stepwise regression in building his models, Hawkes preferred to analyze the "individual leverage plots and by plotting the residuals of various models against each explanatory variable" (Hawkes, 2005:46). It was determined very quickly that the Air National Guard (ANG) bases behaved quite differently than the active duty bases. Therefore, separate models were built for each; ANG and active bases. The significant variables found in the finalized ANG model included utilization rate, % block (30), bases NJ and Ellington, and lag 1 CPFH. For the final active duty model, the significant variables included utilization rate, % block (50), average age, and bases Nellis and Alaska. During the process of his research, Hawkes removed one active duty base, Mountain Home AFB, due to unexplainable outlying values from the other data that may have been a result of invalid data.

Lastly, Hawkes (2005) identified a lurking variable which he called the "year effect." He summarized the "year effect" as being a yearly trend in the data that was much more prominent in the ANG model than the active duty model. Hawkes offered three conclusions concerning the "year effect". First, he suggested the interaction of the explanatory variables was not causing this effect. Second, the year effect was greater for the ANG than for active duty. Lastly, he narrowed the effect down to either the change in mission profiles due to the terrorist activities of September 11, 2001, or modifications to the F-16 fleet. The Hawkes (2005) research represents the most current analysis of this question.

Laubacher (2004) evaluated various methods to forecast the CPFH rates for USAF helicopters. The analysis started with comparing the actual CPFH figures with those submitted in the Program Objectives Memorandum (POM). The actual cost data was obtained from the AFTOC database, while the budgeted numbers were obtained from the POM submissions. Laubacher calculated the percent error between the POM submission and actual expenditures to help determine if there were any apparent trends by MAJCOM in terms of either over or under budgeting.

Next, Laubacher (2004) utilized three separate forecasting techniques to analyze each MAJCOM's actual CPFH figures. The three techniques used in this analysis were the 3-year moving average (MA), single exponential smoothing (SES) method, and the Holt's linear method (Laubacher, 2004:62). The first two forecasting techniques cannot account for trends in the data (Makridakis, 2003:143,161), where as the Holt's method can (Makridakis, 2003:155). Neither of these methods; however, can account for any type of seasonality that may be present in the data. This is an important observation for this research as one of the research questions involves detecting seasonality within the CPFH program. Laubacher based the robustness of these three methods on four common forecasting accuracy measures; Mean Error (ME), Mean Absolute Error (MAE), Mean Percent Error (MPE), and Mean Absolute Percent Error (MAPE) (Laubacher, 2004:64). The method of determining if the forecasted CPFH costs were more accurate than the budgeted forecasts was to compare the variances of each with respect to the actual costs. The first rotary aircraft Laubacher investigated was the MH-53J "Pave Low".

The MH-53J “Pave Low” helicopter, which is flown in two MAJCOMs: Air Education and Training Command (AETC) and Air Force Special Operations Command (AFSOC), had a significant trend component in the CPFH costs. Therefore, it was intuitive that the Holt’s method performed the best for this MDS. Holt’s method performed well in respect to the actual versus forecasted cost versus budgeted costs. The forecasted variances for the three years examined were better than the budgeted variances in two of the three years (Laubacher, 2004:72). Lastly, Laubacher used the forecasting equation developed earlier to forecast one period ahead and compared this number with the actual. “By adding this single data point, the MAPE improved by more than two percent” (Laubacher, 2004:72).

The next MDS examined was the UH-1N “Huey” helicopter at AETC, Air Mobility Command (AMC), and Pacific Air Forces Command (PACAF). The same procedure used for the previous MDS was employed to analyze each MAJCOM. For AETC, Holt’s method significantly outperformed the other two forecasting methods. In the comparison of the actual versus budgeted and actual versus forecasted, the forecasted variances again were much better than the budgeted variances. The analysis of the AMC data provided almost the same answer as for AETC. Holt’s was the best method, as the forecasted figures outperformed the budgeted, and the forecast model accurately forecasted the next period. Lastly, PACAF was examined. The PACAF data were very unstable in that were many large fluctuations in the 7-year span. This led Laubacher to select the SES method. He concluded Holt’s model could not effectively estimate large fluctuations. The SES method provided good forecasts that still outperformed the budgeted, but not as well as in previous MAJCOM analysis. Lastly, the SES model was used to forecast one period ahead, and this was compared to the actual for that period. The results were a 3% increase in the MAPE, believed to be attributable to the instability of the data (Laubacher, 2004:82-83).

The last MDS examined was the HH-60G “Pave Hawk” helicopter. The MAJCOMS that fly this helicopter include AETC, Air Combat Command (ACC), Air Force Reserve Command (AFRC), and PACAF. AETC was first to be examined. The best forecasting method for AETC and this MDS was Holt’s, again. The analysis of the forecasted figures and the budgeted figures indicated that the forecasted figures outperformed the budgeted figures in all the years except for the first year. This was due a large increase in CPFH costs between the first and second year. As for the forecast of one period ahead, this model did not perform as well, increasing the MAPE by 4.5% (Laubacher, 2004:86).

ACC was analyzed next and, again, the Holt method outperformed the other two methods. The analysis of the forecasted figures versus budgeted figures resulted in significantly better variances for the forecasted figures. The forecast ahead of one period did not perform well at all. It increased the MAPE by 9.7%, possibly as a result of a three-year decrease in CPFH costs and then a sharp increase in the last year (Laubacher, 2004:89). As for AFRC, the Holt method was optimum. As with some of the other MAJCOMs, a sudden increase in the data caused the forecasted figures to be better than the budgeted figures two out of the three periods. The one period ahead forecast did not have any significant change in the accuracy measures. The last MAJCOM to be analyzed was PACAF. The best method for PACAF proved to be the SES method, similar to the previous MDS and PACAF. Due to a large increase in the later years of the data set, the forecasted variances were much better than the budgeted variances over the last two time periods. Finally, the forecast ahead was very close to the actual amount, but only decreased the MAPE by 1% due to the large fluctuations in the data.

The Physics Based Alternative to Cost-Per-Flying-Hour Models of Aircraft Consumption study, commissioned by the Assistant Secretary of the Air Force for Cost and Economics and performed by the Logistics Management Institute (LMI), centered on trying to develop a better model in predicting CPFH rates (Wallace, 2000: iii). The reason being, the proportional models that were used to predict the consumption during OPERATION DESERT STORM over predicted by more than 200% (Wallace, 2000: iii). "These proportional models predict the maintenance needs of a fleet of aircraft on the basis of a simple scaling method" (Wallace, 2000: 1-1). The proportional models failed because these models were based on flying patterns that did not change drastically. However, during times of conflict, for example OPERATION DESERT STORM, Kosovo, etc, the flying patterns of the aircraft did change significantly. During these conflicts, the number of landings remained relatively the same, but the number of flight hours drastically increased. This, in turn, reduced the amount of "idle" time the aircraft spent on the ground. The measures of number of landings, time on ground, and sortie duration, were what the proportional models used in their respective predictions.

Therefore, a better model needs to be built to account for changes in the flying patterns of a fleet of aircraft.

There was a current physics based model, developed by Dr. David Lee, which LMI used as a foundation for their model. The Lee model used the variables:

- take/off landing cycles
- ground hours
- flying hours (Wallace, 2000: 2-1).

This original physics-based model outperformed the proportional based models, but the researchers thought improvements were possible. They made two changes to the model after further research:

- Changed the input distribution for ground time from binomial to a Poisson process to account for a more constant stress on the aircraft from temperature and humidity (Wallace, 2000: 2-2).
- Separated landings by type of landing—cold cycles for initial take-off and final landing and warm cycles for touch and goes (cause more stress to the aircraft than the cold starts, thus higher maintenance costs). Fighter aircraft rarely perform touch-and-go maneuvers; therefore, this segregation does not apply to F-15s (Wallace, 2000: 2-2).

Using this model will not provide a CPFH as simply as the proportional models; however, calculating the costs were very straight forward as seen here (Wallace, 2000: 2-2).

$$\left(FH\right)\times\left(\frac{removals}{FH}\right)\times\left(\frac{Cost}{removal}\right)=\left(FH\right)\times\left(\frac{Cost}{FH}\right)=Cost$$

The methods the researchers used to evaluate the robustness and predictability of the “new” physics-based model was three fold. First, they calibrated their model using the C-5B data from OPERATION DESERT STORM and compared it to the proportional model’s estimates. Then they used three different airframes to test the model against the C-17, F-16C, and KC-10. They chose these airframes because they represented the three major groups of aircraft, transports, fighters, and tankers, and, each of these airframes had been used in a recent conflict. The researchers divided their data into four different sets; the first three were for calibration and testing of the model. The last set was used for final testing because this data set contained time frames for prior to the conflict and during the conflict. This was the best set to use to test the physics-based model since it was built to better model changes in flying patterns. The data were obtained from the AFTOC and REMIS databases. Lastly, relative error and root mean squared were used as the measures to evaluate each models performance. (Wallace, 2000: 4-1).

The researchers concluded, for the initial calibration, the physics-based model significantly outperformed the proportional model hands down; it was more robust and provided more accurate forecasts. However, the data for the C-5B aircraft during OPERATION DESERT STORM was suspect due to its age and possible inaccuracy. Therefore, the researchers used C-17, F-15C, and KC-10 data from Kosovo. For the C-17 analysis, the physics-based model again, outperformed the proportional model. The researchers attributed some of this success to the physics-based models’ parameters and their ability to "…react to the data better than the single parameter of the proportional model" (Wallace, 2000: 4-15).

Next, the researchers examined the KC-10 tanker data during Kosovo. The researchers identified the data did not indicate a discernible change in flying patterns; however, they did feel the data had enough change to test the models. They found the physics-based model outperformed the proportional model only for the small surge time frame. Otherwise, there was no notable improvement over the proportional model for the remaining time frames. Lastly, both models over predicted the costs with the physics-based model to a lesser magnitude. The last airframe tested was the F-16C. The researchers had to limit the data used for the F-16 to one base, Aviano AB, Italy, because of the large size of the F-16C fleet aggregated showed no noticeable flying pattern changes. Aviano was selected because their F-16Cs had flown numerous missions in Kosovo. After the re-setting of the data, results indicated that the physics-based model performed better than the proportional model. However, the proportional model performed well also.

Based on the above study, sortie duration will be used as one of the independent variables in this research project. To further investigate the effect of changing flying patterns, this research will divide the sortie duration into combat sortie duration and training sortie duration.

Wu (2000) estimated the Operation and Support (O&S) costs of USAF aircraft fleets and developed his own model based on operations tempo and physical characteristics. The model built from this research is used in the acquisition process of an aircraft fleet. Although this research is not centered on the estimation of direct flying hours, it has applicability to this research as it examines many of the same underlying questions. Additionally, like the first portion of this research, it looked at the aggregate aircraft fleet to develop its model. The independent variables used in this research included flying hours per aircraft, total aircraft inventory (TAI), flyaway costs per aircraft, average Mission Design (MD) age, and Mission Design Series type. Four different models were then tested using different combinations of the above variables. It was determined the optimum model was developed using the following significant variables: average flying hours; the number of aircraft; flyaway costs; and MD fleet age (Wu, 2000: 49). As this research stated in its conclusion, O&S costs are of a major concern for today's decision makers, as O&S costs represent a significant portion of an aircraft's acquisition and development costs. Furthermore, of total O&S costs, CPFH represented the major cost component. Therefore, the above thesis supports this research's use of average aircraft age as an independent variable to build its models.

Variables Selected for Investigation and Model Building

This research will use the variables based on previous research with the addition of a few additional variables. Table 1 represents variables, based on previous research, which will be incorporated into this project.

Table 1. Variables Identified from Previous Research

	Variables from Previous Research								
Research Article	Avg Sortie Duration	Avg A/C Age	% Deployed	% engine type	MAJCOM	Location	Ute rate	Mission type	CPFH Lag
Hawkes	**X***	**X***	**X**	**X**	**X**	**X***	**X**		**X***
Laubacher									**X***
LMI	**X***					**X***		**X***	
Ming-Cheng		**X***							

* Variables will be utilized in this research

Additional Research Supporting Additional Independent Variables

Additional research has been performed to support the use of climatology variables and jet fuel prices in predicting CPFH factors. Also, the justification for using *program change* and *war* as dummy variables will be explained.

Climatology

In a 1983 article, by Major Larry G. McCourry, in the *Air Force Journal of Logistics*, General Bryce Poe, former Air Force Logistics Command Commander, was quoted as saying, "…he could use the billion dollars spent every year in fighting corrosion to fund one-third of the Air Force's shortfall in aircraft replenishment spares for a fiscal year" (McCourry, 1983:5). Also stated in this article was "…that approximately 28% of the costs for the C-130 fleet and 23% of the costs of the C-141 fleet maintenance are due to corrosion" (McCourry, 1983:6). There are significant resources that could be directly allocated to the CPFH accounts of consumable and even DLRs. The corrosion not only affects the airframe, but also the components that are inside. Any component that is not hermetically sealed can encounter corrosion. In the study *Effect of Environmental Factors on the Corrosion of 2024T3 Aluminum Alloy* (Guo, 2004), the authors conducted laboratory tests on commissioned Naval aircraft to determine the main factors affecting the corrosion of this alloy. They found:

> Among the four factors representing oceanic atmosphere environment, concentration of Cl^- and SO_4^{2-} and temperature have great effect on the corrosion of 2024T3 aluminum alloy while humidity contributes less to it. (Guo, 2004:433)

Additionally, it was found that in cold climates, the numbers of hydraulic leaks were greater than in warm climates. Also, the wetter the climate, the aircraft's avionics system risk a greater chance of failure (Tetmeyer, 1982: IV124). Based on these articles, this research is including the variables *temperature* and *salinity* in its analysis of CPFH rates.

Jet Fuel Prices

The consumer jet fuel prices are being used as a proxy variable to account for the fluctuations and impact the petroleum industry has on the aerospace industry. Oil price fluctuations not only affect the cost of aviation fuel, but also the cost of acquiring other goods such as aircraft parts (Hicks, 2005). This impact is mainly seen in the transportation and manufacturing costs of end items used in aircraft from consumables to depot level Reparable. As a result, this variable will be investigated as to its impact on the CPFH rates of the F-15CD and F-15E fleets.

Program Change Dummy

On 1 October 2003, the USAF announced a change to the types of items that could be allocated to the CPFH program. Previously, these items were allocated to the Base Operating and Support account; therefore, it was determined to be a zero-based transfer (ZBT)--no addition or subtraction of the bases money, just in the method of accounting and allocating the costs. The ZBT statement of intent from the Deputy Assistant Secretary, Cost and Economics, is as follows:

> All consumable items directly related to aircraft, aircraft maintenance and the production of sorties and/or flying operations are CPFH expenses. Additionally, aircrew gear/equipment (other than uniforms and personal items) are CPFH expenses. All items that fall in these categories, whether they are on the aircraft or stored off the aircraft are CPFH expenses. Further, some Non-Fly Aviation Fuel (AVFUEL) used in support of flying operations is a CPFH expense. (SAF/FMC, 2003:1-2)

The program change dummy was used to ascertain if there was an impact to the CPFH program when the Zero Based Transfer (ZBT) program change was initiated. The variable will be coded as binary and will start on 1 October 2003, when the ZBT was initiated.

War Dummy

This variable represents the start of OPERATION IRAQI FREEDOM and its continuance through today. The logic behind choosing a war dummy was two-fold. First, during a war/conflict flying practices are very different—more and longer sorties will be flown during these times than in peace-time. This is another, exogenous, way of determining if sortie duration has an impact on the CPFH. Second, war potentially has an enormous impact on the economy as a whole. In the Department of Defense, more money is made available to carry out the mission at hand. This means more money is also available to allocate towards the CPFH program to help maintain a higher mission capability rate than in peace-time. Also, the impact on the economy could drive prices up for those items needed to fly aircraft, such as fuel, spare parts, and consumables. This war dummy will help identify, if it is significant, whether wars/contingencies have an impact on the CPFH rate and can be accounted for by the analyst at a base. This variable may also prove a potent tool for simulation of war time cost changes at the MAJCOM or higher echelons.

Summary

This chapter summarized the previous research done on CPFH development, the Air Force budget process, and the evolution of the F-15 “Eagle” aircraft. In addition, it outlined the variables chosen for the estimates of the models and the motivation in choosing them for the estimation. The following chapter, “Data and Methods”, will describe the databases where the data was obtained from and the form of the data. Lastly, it provided an overview of the techniques to be used to analyze the data and build the models.

III. Data and Methods

Chapter Overview

The purpose of this chapter is to describe the data and methods used to answer the research questions in Chapter 1. First, the chapter will discuss the development of the database by describing where the data were obtained from and its form. Lastly, this chapter will briefly explain the method used, panel models, to analyze the data and build the models summarized in Chapter 4.

Database Development

The main components of this researches database were obtained from the Air Force Total Ownership Cost (AFTOC) database, Air Force Reliability and Maintainability Information System (REMIS), or the Air Force Combat Climatology Center (AFCCC) database. The AFTOC database is an unclassified repository of Air Force weapons systems' operation and support (O&S) costs. The data were compiled from numerous other databases to include: the Fuels Automated Management System (FAMS) which provides fuel data, the Command On-Line Accounting & Reporting System (COARS) which provided actual expenditures and the Military Personnel Data System E300Z report that provides the military personnel expenditures (Laubacher, 2004:58-59).

The AFTOC data were provided by the Air Force Cost Analysis Agency and it contained the depot level Reparable (MSD) and consumables (GSD) portions of the CPFH rate for each base. The data were provided in then year (TY) dollars for each base by fiscal year, fiscal month, and MSD. An example of this data is shown in Table 2.

Table 2. Example of Cost Data from AFTOC Database

FY	FY_Month	Command_CPFH	Base	Data_Type	MD_CAIG	Net_Cost
2001	01	ACC	EGLIN AFB (FL)	GSD	F-15C/D	$802,746.25
2001	02	ACC	EGLIN AFB (FL)	GSD	F-15C/D	$904,619.43
2001	03	ACC	EGLIN AFB (FL)	GSD	F-15C/D	$538,797.04
2001	01	ACC	EGLIN AFB (FL)	MSD	F-15C/D	$2,475,965.71
2001	02	ACC	EGLIN AFB (FL)	MSD	F-15C/D	$4,433,262.13
2001	03	ACC	EGLIN AFB (FL)	MSD	F-15C/D	$3,904,812.09

The Air Force Combat Climatology Center is a repository of climatology observations for over 10,000 individual locations throughout the world. Included within the database were the surface observations for individual stations (e.g., Eglin AFB, Elmendorf AFB), which was what this research is using. The data received from the center provided all of the climatology data (mean temp, mean max temp, mean min temp, max temp, and min temp), except for the independent variable, salinity. Temperature was represented in degrees Celsius, and salinity was determined by proximity to salt water and was coded as binary; “1” for being close to salt water and “0” for not. An example of this data is shown in Table 3.

Table 3. Example of Data provided by AFCCC

Year	Month	Meanmaxtmp °C	Meanmintemp °C	MeanTemp °C	Max °C	Min °C
2001	1	12.40	1.27	6.79	24	-8
2001	2	16.11	4.79	10.25	28	-3
2001	3	16.61	5.77	11.38	26	-2
2001	4	24.13	11.00	18.15	33	0
2001	5	27.68	15.32	21.99	34	9

The last database used for obtaining data was Air Force Reliability and Maintainability Information System (REMIS). Like the Air Force Total Ownership Cost (AFTOC) database, REMIS is a repository of multiple other data sources. The main purpose of this database is to provide maintenance and logistic data for all Air Force weapon systems. The average age of the aircraft data was extracted from this database and was provided by SAF/XP. An example of this data is shown in Table 4.

Table 4. Example of Data provided by REMIS

F-15 C/D					
BASE	**2001**	**2002**	**2003**	**2004**	**2005**
eglin air force base	225.16	237.16	249.11	260.72	272.72
elmendorf air force base	169.01	180.78	192.78	204.78	216.78
kadena air base	249.53	261.53	273.63	285.63	297.73
langley air force base	207.37	219.58	232.01	243.30	254.56

In addition, the total number of hours flown, number of training hours flown, number of combat hours flown and the number of sorties flown for training and combat by base and MDS were provided by the Air Force Cost Analysis Agency in a separate worksheet. An example of this data is shown in Table 5.

Table 5. Example of Sortie Data from REMIS Database

Fscl_Per	Month	MD_Rollup	Mission	AFTOC_Installation	Sum of FH	Sum of Sorties	Avg sortie duration	Mission	Sum of FH	Sum of Sorties	Avg sortie duration	TOTAL FH	TOTAL Sum of Sorties	TOTAL AVG Srt Dur
2001Q1	1	F-15C/D	Combat	EGLIN AFB (FL)	0.0	0.0	0.0	Training	1,303.7	933.0	1.4	1,303.7	933.0	1.40
2001Q1	2	F-15C/D	Combat	EGLIN AFB (FL)	5.8	3.0	1.9	Training	860.6	623.0	1.4	866.4	626.0	1.38
2001Q1	3	F-15C/D	Combat	EGLIN AFB (FL)	643.4	187.0	3.4	Training	656.5	341.0	1.9	1,299.9	528.0	2.46
2001Q2	4	F-15C/D	Combat	EGLIN AFB (FL)	789.3	213.0	3.7	Training	641.8	533.0	1.2	1,431.1	746.0	1.92
2001Q2	5	F-15C/D	Combat	EGLIN AFB (FL)	717.9	172.0	4.2	Training	685.3	538.0	1.3	1,403.2	710.0	1.98
2001Q2	6	F-15C/D	Combat	EGLIN AFB (FL)	359.3	92.0	3.9	Training	734.6	459.0	1.6	1,093.9	551.0	1.99

The remaining independent variables are as follows:

Jet Fuel: The historical data for jet fuel for resale was obtained from the *October 2005 Monthly Energy Review* (Energy, 2006) from the Energy Information Administration.

Program Change Dummy Variable (DV): This binary variable represented the date, 1 October 2003, which the ZBT CPFH program change was initiated.

War DV: This binary variable represented the start of OPERATION IRAQI FREEDOM and its continuance through this date. OPERATION ENDURING FREEDOM does not have a separate variable assigned because it spans the entire range of the data being used.

Seasonal DVs: These binary variables represented the months of the year, except for October which is the base month, and they will measure the seasonality within the data.

Methods

Panel Model

The first method used in the analysis of the data was the Panel Model. Panel models are used to examine cross-sectional time-series data and help in determining the relationship a set of time-series variables have across a different set of individual observations. In other words, this method analyzed an independent variable across "groups" (sites, locations, bases, cities, countries, etc.) with respect to multiple time periods. According to Peter Kennedy in "A Guide to Econometrics", there are numerous appealing features of the panel model, of which the following four are most prominent:

1. The model is stochastic and not deterministic
2. Panel data provides the ability to deal with omitted explanatory variables in both the cross-section and time-series when they are looked at individually. The omission of these variables leads to biased estimations (Kennedy, 2003:302).
3. Panel data leads to a more efficient estimation because panel data increases the variability. The combining of the data, time-series with cross-sectional, in essence combines the variability of both data sets. This helps reduce the multicollinearity problems associated with the data sets individually (Kennedy, 2003:302). Additionally, in the traditional cross-sectional regression model, the variation between "groups" is incorporated into the error term and cannot be ascertained. Panel modeling enables the ability to account for such variation.
4. The use of panel data allows researchers the ability to analyze issues that cannot be studied by using time-series or cross-sectional data alone (Kennedy, 2003:302).

5. “Panel data allow better analysis of dynamic adjustment.” (Kennedy, 2003:302). Simply put, it allows the researcher to investigate the interactions of variables across a range of individuals, cities, bases, etc.
6. Increases the number of observations available for testing(degrees of freedom-out of sample testing)
7. Potentially isolates temporal/spatial specific variations

There are two main types of panel data analysis, fixed effects and random effects (Kennedy, 2003:304). Fixed effects panel data assumes there are minimal time-series effects on the dependent variable, but more cross-sectional influences. That is the, intercept of the regression is specific to the “group” effect and not the time effect. The second main type of panel model is the random effects model. This model assumes there is a random constant term that is attributed to a random error specific to a particular observation. Random effects models can accommodate variables that are time-invariant (don’t vary within the individual “group”) where as fixed effects omit these variables.

The determination of which model, fixed effects or random-effects, best fits the data being described can be tested using the Hausman Specification Test. The Hausman test is based on the null hypothesis that the two models, fixed and random, are not different. The alternative hypothesis is the two models are different. The predominant method in use is fixed effects. This research will use the fixed effects model to analyze the F15 fleet data. Regarding fixed effects models, Kennedy, states, “If the data exhaust the population, then the fixed effects approach, which produces results conditional on the cross-section units in the data sets, seems appropriate because these are cross-sectional units under analysis" (Kennedy, 2003:312).

A few assumptions with panel data need to be addressed, to include model specification, stationarity of the dependent variable, heteroskedasticity, normality of the residuals, and multicollinearity. These assumptions and the tests to identify them will be specifically addressed in Chapter 4.

Summary

This chapter explained how the data were obtained from each of the repository databases; Air Force Total Ownership Cost (AFTOC) database, Air Force Reliability and Maintainability Information System (REMIS), and the Air Force Combat Climatology Center (AFCCC) database. It also described where the data was obtained for each of the variables not found in the three databases. Next, the foundation for the methods used to analyze the data and develop models that answer the research questions identified in Chapter 1 was provided. It briefly described the panel data model and the two different types of panel models, fixed effects and random effects. In addition, this chapter also discussed some of the assumptions that have to met and verified, through various tests, for the methods being used. The steps and results of the methods used will be described in the next chapter.

IV. Analysis and Results

Chapter Overview

The purpose of this chapter is to explain the processes that were undertaken to analyze the developed database using the knowledge gained from the literature. The first items discussed will be the *a priori* models developed that focused the analysis of the data. Next, the assumptions that need to be addressed for time-series data prior to modeling it is explained along with the determination of the lag structure. The individual model results will be thoroughly discussed to include the post-estimation tests that need to be performed. Lastly, each of the models will be measured as to how accurate they perform and how well they compare to previous models. First to be discussed is the theoretical model specification.

Theoretical Model

The first step in building any model is to start with a theoretical model of all the variables that the research considers. The equations below provide the foundation for the building of the *a priori* model that follows.

$$CPFH_{Rate} = f(DLR_{Rate} + Consummable_{Rate} + AVFUEL_{Rate}{}^{1}) \qquad (1)$$

---

[1] For this research, $AVFUEL_{Rate}$ is being considered a constant due to its stability and accuracy of prediction (Rose, 1997:8)

Where:

DLR_{Rate} = f(Consumable Rate + Total Average Sortie Duration
+ Average Training Sortie Duration + Average Combat Sortie Duration
+ Program Change DV + War DV + Jet Fuel + Average Temperature
+Average Temperature Difference + Seasonal DVs
+ Producer Price Index Aerospace Industry) (2)

$Consummable_{Rate}$ = f(Total Average Sortie Duration + Average Training Sortie Duration
+ Average Combat Sortie Duration + Program Change DV
+ War DV + Jet Fuel + Average Temperature
+ Average Temperature Difference + Seasonal DVs
+ Producer Price Index Aerospace Industry (3)

With the model specified in general terms, this research looked at the correlation matrices for each Mission Design Series (MDS) to determine if there were independent variables that were correlated with each other—a correlation coefficient greater than the 0.50 in absolute value (Hinkle et al, 1982:100). These correlation matrices can be found in Table 6 for the F-15CD fleet and Table 7 for the F-15E fleet. The scatterplot matrix graphs identified several variables that were correlated with each other. First, Total Average Sortie Duration was highly correlated with Average Training Sortie Duration and Average Combat Sortie Duration. Therefore, Total Average Sortie Duration was selected because this variable was believed to best address the investigative question in Chapter 1. The next variables that were found to be correlated were Mean Temperature Difference and Mean Temperature. Based on the research summarized in Chapter 2, the Mean Temperature Difference as the measure of temperature was used. Lastly, there were four other variables that had correlation coefficients that exceeded the 0.50 in absolute value threshold (Hinkle et al, 1982:100), but they were determined to be spurious relationships and definitely held no causal relationships between them. Later in the analysis, correlations were evaluated based on the Variation Inflation Factors (VIF) calculated after the regressions were computed.

Table 6. Correlation Matrix for F-15CD Fleet all Bases

Correlation Matrix-F-15CD all Bases

	DLR Rate	CONS Rate	PPI Aerospace	Total Avg Sortie Duration	Avg Combat Sortie Duration	Avg Training Sortie Duration	ZBT Program Change	War Dummy	Commercial Jet Fuel for Resale	Avg Aircraft Age	Salinity Dummy	Avg Mean Temp	Mean Temp Difference
DLR Rate	**1.000**												
CONS Rate	0.264	**1.000**											
PPI Aerospace	0.425	0.257	**1.000**										
Total Avg Sortie Duration	-0.202	-0.079	-0.027	**1.000**									
Avg Combat Sortie Duration	0.004	0.060	0.046	0.227	**1.000**								
Avg Training Sortie Duration	-0.170	-0.086	-0.047	**0.949**	0.047	**1.000**							
ZBT Program Change	.	.	.	.	.	.	**.**						
War Dummy	.	.	.	.	.	.	.	**.**					
Commercial Jet Fuel for Resale	0.310	0.230	**0.787**	-0.036	0.063	-0.097	.	.	**1.000**				
Avg Aircraft Age	0.189	0.190	0.088	0.062	0.035	0.060	.	.	0.069	**1.000**			
Salinity Dummy	0.195	-0.034	0.025	0.333	0.334	0.305	.	.	0.026	0.234	**1.000**		
Avg Mean Temp	0.302	0.173	**0.527**	-0.182	0.021	-0.232	.	.	**0.590**	**0.579**	-0.033	**1.000**	
Mean Temp Difference	0.035	0.028	0.250	-0.315	0.019	-0.339	.	.	0.269	-0.271	**-0.671**	**0.574**	**1.000**

Table 7. Correlation Matrix for F-15E Fleet all Bases

Correlation Matrix-F-15E all Bases

	DLR Rate	CONS Rate	PPI Aerospace	Total Avg Sortie Duration	Avg Combat Sortie Duration	Avg Training Sortie Duration	ZBT Program Change	War Dummy	Commercial Jet Fuel for Resale	Avg Aircraft Age	Salinity Dummy	Avg Mean Temp	Mean Temp Difference
DLR Rate	**1.000**												
CONS Rate	0.325	**1.000**											
PPI Aerospace	0.337	0.332	**1.000**										
Total Avg Sortie Duration	-0.320	-0.237	0.001	**1.000**									
Avg Combat Sortie Duration	-0.195	-0.156	-0.125	**0.757**	**1.000**								
Avg Training Sortie Duration	-0.107	-0.054	0.095	**0.506**	0.042	**1.000**							
ZBT Program Change	.	.	.	.	.	.	**.**						
War Dummy	.	.	.	.	.	.	.	**.**					
Commercial Jet Fuel for Resale	0.168	0.315	**0.785**	-0.028	-0.149	-0.029	.	.	**1.000**				
Avg Aircraft Age	0.154	-0.139	0.122	-0.216	-0.168	0.230	.	.	0.099	**1.000**			
Salinity Dummy	-0.258	-0.325	0.027	0.485	0.308	0.433	.	.	0.028	0.146	**1.000**		
Avg Mean Temp	0.229	0.396	**0.602**	-0.255	-0.163	-0.371	.	.	**0.655**	-0.058	-0.231	**1.000**	
Mean Temp Difference	0.271	0.339	0.388	-0.312	-0.296	-0.116	.	.	0.420	0.290	**-0.582**	**0.526**	**1.000**

Therefore, for the panel model, the specified notations for the above equations were as follows (signs represented the theoretical direction the variable was believed to affect the dependent variable). Additionally, there was one of the *a priori* equations below for each MDS, F-15CD and F-15E.

$$DLR_{it} = \alpha_i + \alpha_{i+1}base_i - \beta_1 TotAvgDur_{1it} + \beta_2 ProgramChange_{2it} + \beta_3 War_{3it} + \beta_4 JetFuel_{4it} + \beta_5 AvgAge_{5it} + \beta_6 MeanTempDiff_{6it} + \beta_7 PPI\ Aero_{7it} + \beta_8 Consum\_Rate_{8it} + \beta_{9-19} MonthlyDummies_{9-19it} + \varepsilon_{it} \quad (4)$$

$$CONS_{it} = \alpha_i + \alpha_{i+1}base_i - \beta_1 TotAvgDur_{1it} + \beta_2 ProgramChange_{2it} + \beta_3 War_{3it} + \beta_4 JetFuel_{4it} + \beta_5 AvgAge_{5it} + \beta_6 MeanTempDiff_{6it} + \beta_7 PPI\ Aero_{7it} + \beta_8 Consum\_Rate_{8it} + \beta_{9-19} MonthlyDummies_{9-19it} + \varepsilon_{it} \quad (5)$$

Where *it* is the *i*th base in the *t*th time period and β_9- β_{19} represent the eleven monthly dummy variables with October being the base month.

Panel Model Pre-Estimation Assumptions

The first assumption that needs to be met with any time-series data is that of stationarity of the dependent variable. Stationarity is the condition that the data, through time, centers on a constant mean and has a constant variance. The test used to determine if a variable has a unit root, or is stationary, was the Augmented Dickey-Fuller unit root test. This test is based on the null hypothesis that the variable follows a unit-root process (non-stationary); with the alternative hypothesis being the presence of a unit root (stationary). Results of the Augmented Dickey Fuller Unit Root tests are displayed in Table 8. As evident from the tables, the panel data is from a stationary process. This permits estimation in levels to proceed.

Table 8. Augmented Dickey-Fuller Unit Root Test F-15 CD/E all Bases

F-15CD Fleet				
Dickey-Fuller test for Unit Root		Number of obs =		419
		Interpolated Dickey-Fuller		
	Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value
Z(t)-DLR Rate	-16.2330	-3.4460	-2.8730	-2.5700
MacKinnon approximate p-value for Z(t) =			0.0000	
Z(t)-CONS Rate	-16.5800	-3.4460	-2.8730	-2.5700
MacKinnon approximate p-value for Z(t) =			0.0000	

F-15E Fleet				
Dickey-Fuller test for Unit Root		Number of obs =		419
		Interpolated Dickey-Fuller		
	Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value
Z(t)-DLR Rate	-14.5220	-3.4560	-2.8780	-2.5700
MacKinnon approximate p-value for Z(t) =			0.0000	
Z(t)-CONS Rate	-14.2400	-3.4560	-2.8780	-2.5700
MacKinnon approximate p-value for Z(t) =			0.0000	

Panel Model Lag Structure Determination

The first step in performing the model analysis for depot level Reparable (DLR) and consumables (CONS) by Mission Design Series (MDS) was to determine if there was a lag structure within the dependent variable and/or the independent variables. Therefore, each of the dependent variables, DLR and CONS, was regressed against its lags for each MDS. The determination if there was a lag was to be made based on the R^2 and Akaike Information Criteria (AIC) values. According to Kennedy, in *A Guide to Econometrics*, the use of the AIC and R^2 to determine appropriate lag lengths in time-series data is common practice (Kennedy, 2003:88). The optimum lag length is reached when the AIC is minimized and/or R^2 is maximized, or both happen. If, however, R^2 declines and the AIC goes up or down then the optimum lag is not reached. This research first attempted to determine if there was a lag structure for the dependent variables.

As depicted in Table 9, the AIC continually decreased as the lags were increased, and the R^2 fluctuated considerably. Based on the aforementioned criterion, there did not appear to be a discernible lag structure for the DLR of the F-15CD fleet. These results were common for the testing of the F-15 CD CONS, F-15E DLR, and F-15 E CONS. Therefore, results indicated there was no lag structure for the dependent variables. The results of the regressions for the F-15CD Fleet for DLRs are presented in Table 9.

Table 9. F-15CD DLR Lag Structure Results

F-15CD DLR Lag Determination					
Model	**Obs**	**df**	**AIC**	**BIC**	**R^2**
No Lag	84	15	1635.496	1671.958	0.2941
DLR L1	83	16	1605.498	1644.199	0.3762
DLR L2	82	16	1596.195	1634.702	0.3038
DLR L3	81	16	1574.598	1612.909	0.3103
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
DLR L12	72	16	1400.172	1436.599	0.3731

The independent variable lag structure was tested using two different methods. First, each of the independent variables that was believed to have a lag structure (time variant variables: Tot_Avg_Dur, Jet_Fuel, Avg_Age, and Mean_Diff) was regressed against each dependent variable by MDS and by DLR and CONS. Once more, like the testing of the dependent variables, the AIC was continually decreasing with considerable fluctuations in the R^2 value, as represented in Table 10. Results were consistent with all the independent variable tests.

Table 10. F-15CD Tot_Avg_Dur Lag Structure Results vs. DLR only

F-15CD Tot_ Avg_ Dur Lag Determination with DLR only					
Model	**Obs**	**df**	**AIC**	**BIC**	**R^2**
No Lag	84	2	1635.261	1640.123	0.0407
DLR L1	83	2	1612.264	1617.102	0.0516
DLR L2	82	2	1592.679	1597.493	0.0615
DLR L3	81	2	1575.033	1579.822	0.0202
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
DLR L12	72	2	1401.422	1405.975	0.0590

The second method of testing for a lag structure was to change the lags of one of the time variant independent variables while keeping all others constant and then computing the regression. The results of this method were the same as the previous two tests for determining a lag structure; the AICs continually decreased while the R^2s were unstable. Results of lagged Tot_Avg_Dur with the remaining independent variables being held constant are presented in Table 11.

Table 11. F-15CD Tot_Avg_Dur Lag Structure Results vs. DLR only

F-15CD Tot_Avg_Dur Lag Determination vs. all variables					
Model	**Obs**	**df**	**AIC**	**BIC**	**R^2**
No Lag	84	15	1635.496	1671.958	0.2941
DLR L1	83	15	1617.788	1654.071	0.2590
DLR L2	82	15	1598.913	1635.014	0.2625
DLR L3	81	15	1577.074	1612.991	0.2711
DLR L4	80	15	1559.315	1595.045	0.2609
DLR L5	79	15	1541.158	1576.700	0.2600
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
DLR L12	72	15	1396.176	1430.326	0.3903

In view of the lag determination results, no apparent lag structure for the independent variables for this panel data resulted. Next, the results of the panel data models will be discussed.

Panel Model Results

A discussion of the panel data models built from the database developed in Chapter 3 to include the interpretation of the results, post estimation testing, and model validation will be presented next. The common explanation of the post estimation techniques will be discussed first. There were four models built to determine the impact of the independent variables on the dependent variables within a time series and across the fleet by MDS. The four models included F-15CD DLRs, F-15CD CONS, F-15E DLRs, and F-15E CONS. Each of the following models were built as a fixed-effects panel data model using the robust standard error option. Model validation was explored by performing a linear regression with the estimated dependent variables, DLR and

CONS, against the actual historical values. The specific parameters for each model will be presented immediately following the generalized explanation of the specific tests.

Panel Model Post-Estimation Testing

Model Specification

The specification of a panel data model was measured by the performance of a Hausman specification test. The test is based on the H_o: the estimated coefficients of a fixed effects panel regression are not statistically different from the estimated coefficients of a random effects regression. Subsequently, the H_a: the estimated coefficients of the two regressions are different (Stata, 2005:306-307). For the purpose of this research, failing to reject H_o, a large p-value, was the desired outcome; thereby, supporting the use of the fixed effects panel regression. Results of the Hausman Specification Tests are shown in Appendix B.

Normality of Residuals

The normality of a regression's residuals is usually of concern only when performing a hypothesis test, as this is the least restrictive of all the post-estimation tests. The non-normality of residuals has no effect on the coefficients of the independent variables, but it can impact the F- and t-tests and their respective confidence intervals. A histogram plot with a normal plot laid over the top for visual inspection was used. In addition, the Shapiro-Wilk W test for normality was performed on each set of residuals. This hypothesis test has a H_o: the residuals are not discernibly different from a normal distribution with the H_a: the residuals are not normally distributed. Therefore, for the residuals to resemble a normal distribution, failure to reject the null (a large p-value) is the desired outcome. For all but the F-15CD CONS model, the Shapiro-Wilk W test showed the residuals were not normally distributed. However, as stated before, this is only a concern when performing hypothesis tests. Results of the Shapiro-Wilk W test along with the histogram plot of the residuals are located in Appendix C.

Constant Variance of Residuals-Homoskedasticity

The measure of a models constant variance in its residuals, or determining if the model has heteroskedasticity, can be mitigated by using “robust” estimation such as the “White” heteroskedastic invariant variance-covariance matrix (White,1980). This option is what econometricians refer to as “white-washing” the residuals; essentially this removes the presence of heteroskedasticity in a model. Failure to remove heteroskedasticity does not in of itself bias the model coefficients, but it can signify an omitted independent variable. It is more often associated with lower efficiencies in the standard errors. However, heteroskedasticity in conjunction with other regression violations has a profound impact on the usability of a regression model. All models developed in this research were subjected to the robust standard errors option (Kennedy, 2003: 145-148).

Independence of Residuals

The non-independence of the residuals is caused by autocorrelation of the residuals. That is, each residual is affected by the previous one. Failure to meet this post-estimation assumption can cause several grave problems with a model. First, if autocorrelation is present, the F-tests and t-tests are invalid along with the prediction intervals. This is due to the standard errors of the coefficients being smaller than really are. This also leads to the second problem, "spurious regression." Spurious regression is the appearance of significant independent variables, when in fact these variables are not significant. A commonly used test for determining the independence of the residuals (heteroskedasticity) is the Durbin-Watson statistic test (Kennedy, 2003:149). The Durbin-Watson statistic test is a hypothesis test where: H_o = there is no lag one autocorrelation and H_a = lag one autocorrelation is present. The range of values for the Durbin-Watson statistic is between 0 and 4 with a mean of 2. If autocorrelation is not present then the Durbin-Watson distribution is symmetrically centered on its mean of 2 (Makridakis, 2003:303). In interpreting the results of this statistic, the further away from the mean of 2, the more uncertain it is that autocorrelation is not present.

Panel Model Validation

The accuracy measures described in the following paragraphs tested each of the models developed on their adequacy to accurately predict the DLR and CONS rates.

Regression Testing

The first test of each of the model's accuracy was to regress the predicted values against the actual values. If a model is robust in its ability to estimate, the regression model should have high R^2 and Adj R^2 values. If the opposite is observed, the models accuracy is questionable if not poor.

Common Accuracy Measures

Two accuracy measures, Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE), were used to gauge the ability of the models to forecast the FY2005 historical values:. MAE measures the average of the absolute errors between each pair of actual vs. predicted value. This statistic provides a measure of accuracy that is stated in the same terms as the values. For example, if the values being measured are in dollars, the MAE is stated in dollars. On the other hand, the MAPE gives the user a percentage of the error between the two values. It is commonly used because of this attribute, and easy to interpret no matter the scale of the values being assessed. It is especially useful in this context since the scale remains largely constant over the observed time period (2001-2005).

F-15CD Fleet Model—DLR and CONS

The following paragraphs explain the models developed using panel data for the F-15CD fleet DLRs and CONS. They will describe the models themselves, the interpretation of the coefficients, results of the post-estimation analysis, and finally the validation tests are presented.

DLR Model Interpretation

The first model is the F-15CD DLR model. The results of the panel model with robust standard errors for this data are presented in Table 12.

Table 12. F-15CD DLR Model Regression Results

F-15CD DLR Panel Model					
Fixed-effects (within) Regression				Number of obs =	336
Group variable (i): base_index				Number of groups =	7
R-sq:	within	**0.486**		Obs per group: min =	48
	between =	**0.001**		avg =	48
	overall =	**0.063**		max =	48
				F(19,394) =	**12.250**
				Prob > F =	**0.000**

DLR Rate	**Coef.**	**Robust Std. Err.**	**t-stat**	**P>|t|**
CONS Rate	**4.35*****	0.831	5.23	0.000
PPI Aerospace	**-386.00**	368.372	-1.05	0.296
Total Avg Sortie Duration	**-963.80 ****	358.619	-2.69	0.008
ZBT Program Change	**-1390.84***	782.868	-1.78	0.077
War Dummy	**1077.49**	871.935	1.24	0.217
Commercial Jet Fuel for Resale	**39.042*****	12.167	3.21	0.001
Avg Aircraft Age	**198.50†**	136.485	1.45	0.147
Mean Temp Difference	**-11.00**	84.155	-0.13	0.896
November Dummy Variable (DV)	**877.49**	774.861	1.13	0.258
December DV	**1696.57****	742.987	2.28	0.023
January DV	**989.15**	790.894	1.25	0.212
February DV	**1556.28****	773.202	2.01	0.045
March DV	**1996.90****	976.449	2.05	0.042
April DV	**-165.66**	797.638	-0.21	0.836
May DV	**609.76**	880.916	0.69	0.489
June DV	**1448.46***	868.478	1.67	0.096
July DV	**1676.38***	882.617	1.9	0.058
August DV	**2433.61****	1013.836	2.4	0.017
September DV	**4926.92*****	1237.021	3.98	0.000
Constant	**15946.31**	29197.190	0.55	0.585

*** significant to 0.01 level, ** 0.05 level, * 0.10 level, † 0.15 level

The initial examination of the data indicated the overall F-test to be significant to at least three digits ($p < 0.001$). The R^2 values show that the model explains 48.6% f the variation in the depot level Reparable rates (DLR) within each base. Interestingly though, this model explains a very insignificant amount (less than 1$) of the variation that was between the bases; there was very little variation in the DLR rates due to interaction between the bases. Even though the R^2 value for within the bases is not very high, the model does have some interesting findings. First, it is apparent that the Consumables Rate was highly statistically significant ($p < 0.001$) within the model; however, the coefficient is economically insignificant (4.35). Consequently, there were two reasons this variable was not incorporated into the final model. One, this variable was only known at the same time the actual DLR rate was known; therefore, it was useless in forecasting DLR rates. It did indicate a correlation between the DLR rate and the Consumables rate, which was intuitive since consumables are used up during the replacement of most DLRs. Two, the coefficient was insignificant with respect to the overall DLR rate (4.35). This is true with the Jet Fuel variable, also (39.04). Next, nine of the eleven seasonal dummies were highly significant with significant coefficients. This illustrated a definite seasonal component to the model. Lastly, even though the war variable was not significant, it was not highly insignificant (p-value = 0.21) and the coefficient was significant. This illustrated a possible link to the increase in the DLR rate in times of war. This was exogenous to the types of sorties flown during this time. This variable potentially captured the holistic affect of war described in Chapter 3. Thus, the finalized equation for the model is:

$$
\begin{aligned}
DLR_{it} = {} & -963.80(TotAvgDur_{it}) - 1390.84(ProgChng_{it}) + 39.04(JetFuel_{it}) \\
& + 198.50(AvgAge_{it}) + 1696.54(DecDmy_{it}) + 1556.28(FebDmy_{it}) \\
& + 1996.90(MarDmy_{it}) + 1448.46(JunDmy_{it}) + 1676.38(JulDmy_{it}) \\
& + 2433.61(AugDmy_{it}) + 4926.92(SepDmy_{it}) + \varepsilon_{it} \qquad (6)
\end{aligned}
$$

Independence of Residuals

Table 13 displays the Durbin-Watson statistic for this model. The statistic is below 2, but not far enough away that would cause major concern. Based on this measure, there is no significant concern with the possibility of a spurious regression.

Table 13. F-15CD DLR Model Durbin-Watson test—First Order Autocorrelation

F-15CD Fleet			
modified Bhargava et al. Durbin-Watson			
	DLR	=	**1.8715875**

CONS Model Interpretation

The next model is the F-15CD CONS model. The results of the panel model with robust standard errors for this data are presented in Table 14.

Table 14. F-15CD CONS Model Regression Results

F-15CD CONS Panel Model				
Fixed-effects (within) Regression			Number of obs =	336
Group variable (i): base_index			Number of groups =	7
R-sq: within	**0.412**		Obs per group: min =	48
between =	**0.235**		avg =	48
overall =	**0.250**		max =	48
			F(19,394) =	**10.190**
			Prob > F =	**0.000**
CONS Rate	**Coef.**	**Robust Std. Err.**	**t-stat**	**P>|t|**
PPI Aerospace	**-11.520**	24.574	-0.470	0.640
Total Avg Sortie Duration	**-234.02*****	29.315	-7.980	0.000
ZBT Program Change	**-70.078**	56.417	-1.240	0.215
War Dummy	**49.247**	68.011	0.720	0.470
Commercial Jet Fuel for Resale	**0.024**	1.118	0.020	0.983
Avg Aircraft Age	**6.073**	9.769	0.620	0.535
Mean Temp Difference	**-8.38**†	5.899	-1.420	0.150
November Dummy Variable (DV)	**81.783**	47.768	1.710	0.088
December DV	**210.10*****	52.050	4.040	0.000
January DV	**203.18*****	53.129	3.820	0.000
February DV	**221.23*****	53.985	4.100	0.000
March DV	**158.04*****	43.319	3.650	0.000
April DV	**175.79*****	47.643	3.690	0.000
May DV	**163.22*****	54.731	2.980	0.003
June DV	**128.17****	49.801	2.570	0.011
July DV	**156.78*****	48.543	3.230	0.001
August DV	**267.36*****	54.726	4.890	0.000
September DV	**562.17*****	70.676	7.950	0.000
Constant	**1424.087**	1813.844	0.790	0.433

*** significant to 0.01 level, ** 0.05 level, * 0.10 level, † 0.15 level

For this model, the overall F-test is significant to at a minimum three places (p-value < 0.001) also. The R^2 value for the within portion is 0.412. However, for this model, the between R^2 is 0.24 which means there is significantly more interaction between the bases in regards to the consumables rate than with depot level Reparable (DLR) rates. This could be due to the commonality of the consumption of consumables for the F-15CD fleet. Looking at the coefficients of the significant variables, some commonality is apparent between the DLR and CONS models. First, TotAvgDur is highly significant (p-value < 0.001) with a significant coefficient (-234.02) as it was in the DLR model. Lastly, the monthly variables are significant (p-value < 0.05) with again, significant coefficients, but unlike the DLR model, all the months are significant here. There still is correlation, as with DLRs, between the higher magnitude coefficients and the Air Force fiscal year quarters. For example, the highest cumulative values occur in the last quarter of the fiscal year and then again in the later two months of the first quarter. Again, this signifies a strong seasonal/business cycle component in the model as was the case with the DLR model. Finally, the Mean_Diff variable is somewhat significant (p-value = 0.15) and the economic magnitude of the variable does not appear to be highly significant (81.78). However, the magnitude of the variable is based on a one degree difference in the average monthly high and low. Therefore, with the average change in temperature for the entire time-series across all bases being 9.52 degrees Celsius, it is not uncommon for the monthly impact to be ten-times the coefficient in the equation. With this information, Mean_Diff is has a significant economic magnitude and is subsequently a significant variable. The finalized equation for the model is displayed on the following page:

$$
\begin{aligned}
CONS_{it} = & -234.02(TotAvgDur_{it}) + 81.78(NovDmy_{it}) + 210.10(DecDmy_{it}) \\
& + 203.18(JanDmy_{it}) + 221.23(FebDmy_{it}) + 158.04(MarDmy_{it}) \\
& + 175.79(AprDmy_{it}) + 163.22(MayDmy_{it}) + 128.17(JunDmy_{it}) \\
& + 156.78(JulDmy_{it}) + 267.36(AugDmy_{it}) + 562.17(SepDmy_{it}) + \varepsilon_{it} \qquad (7)
\end{aligned}
$$

Independence of Residuals

Table 15 presents the Durbin-Watson statistic for this model. The statistic is just slightly below 2, well with in the range to ascertain there is no lag one autocorrelation present.

Table 15. F-15CD CONS Model Durbin-Watson—First Order Autocorrelation

F-15CD Fleet		
modified Bhargava et al. Durbin-Watson		
	CONS =	**1.918559**

Validation Testing for F-15CD Models

Table 16 displays the results of the validation tests; regression, mean absolute error (MAE), and mean absolute percent error (MAPE) for the two F-15CD models. The first indication of accuracy, the regression of the predicted vs. actual values, indicates neither of the two models were very robust in predicting the actual values. In addition, the DLR model had excessively high MAE and MAPE measures—the average DLR rate for this time frame was $6966.37. The MAE was almost equal to the average; this indicates a very large error which is also evident in the 131.1 MAPE score. This indicates that the predicted amount, on average, was 131 percent greater than the actual value. However, the CONS model performs much better; it has an average error of $223.04 on an average CONS rate of $749.76. This better performance is also seen in the MAPE—on average a 30 percent error rate. These measures for the CONS model are still not very good. However, these measures are for the monthly errors. In Table 22, Comparison against Currently Available Models, it will be shown at the quarterly and yearly levels, these models perform as well as or better than the current models available.

Table 16. F-15CD Fleet Summation of Accuracy Measures

Accuracy of Panel Data Model				
		Measures		
Model	R^2	Adj R^2	MAE	MAPE
F-15CD DLR	**0.0602**	**0.0487**	**6,607.22710**	**131.08955**
F-15CD CONS	**0.1102**	**0.0993**	**223.03829**	**30.34290**

F-15E Fleet Model—DLR and CONS

The following paragraphs will explain the models developed using panel data for the F-15E fleet depot level Reparable (DLR) and consumables (CONS). They will describe the models themselves, the interpretation of the coefficients, results of the post-estimation analysis, and finally the validation tests will be presented.

DLR Model Interpretation

The first model for the F-15E fleet is the DLR model. The results of the panel model with robust standard errors for this data are presented in Table 17.

Table 17. F-15E DLR Model Regression Results

F-15E DLR Panel Model				
Fixed-effects (within) Regression			Number of obs =	240
Group variable (i): base_index			Number of groups =	5
R-sq: within		**0.211**	Obs per group: min =	48
between =		**0.480**	avg =	48
overall =		**0.105**	max =	48
			F(19,394) =	**5.080**
			Prob > F =	**0.000**
DLR Rate	**Coef.**	**Robust Std. Err.**	**t-stat**	**P>|t|**
CONS Rate	**2.46****	1.430	1.72	0.086
PPI Aerospace	**251.33**†	156.584	1.61	0.110
Total Avg Sortie Duration	**-1942.75*****	447.635	-4.34	0.000
ZBT Program Change	**-2526.44****	1237.890	-2.04	0.042
War Dummy	**660.81**	1016.356	0.65	0.516
Commercial Jet Fuel for Resale	**-1.31**	22.401	-0.06	0.953
Avg Aircraft Age	**9.98**	30.760	0.32	0.746
Mean Temp Difference	**-336.74*****	120.173	-2.8	0.006
November Dummy Variable (DV)	**441.85**	1244.963	0.35	0.723
December DV	**589.55**	1222.907	0.48	0.630
January DV	**603.70**	1605.282	0.38	0.707
February DV	**1376.80**	1308.626	1.05	0.294
March DV	**392.59**	1206.148	0.33	0.745
April DV	**390.10**	1063.760	0.37	0.714
May DV	**1345.79**	1199.358	1.12	0.263
June DV	**2880.80****	1260.562	2.29	0.023
July DV	**1214.91**	1180.005	1.03	0.304
August DV	**1270.27**	1370.577	0.93	0.355
September DV	**1960.58**	1778.568	1.1	0.272
Constant	**-29446.67**	21257.980	-1.39	0.167

*** significant to 0.01 level, ** 0.05 level, * 0.10 level, † 0.15 level

This model did not perform like expected based on the assumptions, previous literature, and the F-15CD DLR model. The first and of most concern atypical performance is with the monthly variables. The previous two models and the F-15E CONS model below all had significant evidence of a seasonality/business cycle component; however, this model had only one significant month: Jun. Additionally, the

R^2 measures seem to be reversed from the other models. This model is able to measure the between variation of the DLR rate better than the within variation. Further investigation of the data does confirm a large amount of variation within the years in the DLR rate. This model does have similarities with the other models: TotAvgDur is significant (p-value > 0.001) as in the other models, and Prog_Chng is significant (p-value = 0.042) as it is in the F-15CD model. Having Prog_Chng significant in both DLR models is very interesting since the majority of the items involved in the zero based transfer (ZBT) move were consumables. Lastly, this model had Mean_Diff as highly significant (p-value = 0.006) and the coefficient was also significant (-336.74). However, counter-intuitively, the direction of impact was negative. This means as the difference in the monthly average temperature increase by one degree Celsius, the DLR rate decreases by $337. Below is the finalized equation for the model:

$$\begin{aligned} DLR_{it} = {} & 251.33(PPI\ Aero_{it}) - 1942.76(TotAvgDur_{it}) - 2526.44(ProgChng_{it}) \\ & - 336.74(MeanDiff_{it}) + 2880.80(JunDmy_{it}) -+ \varepsilon_{it} \end{aligned} \tag{8}$$

Independence of Residuals

Table 18 displays the Durbin-Watson statistic for this model. The statistic is below 2, but only slightly and well within an acceptable amount. This model does not have an issue with lag one autocorrelation.

Table 18. F-15E DLR Model Durbin-Watson test—First Order Autocorrelation

F-15E Fleet			
modified Bhargava et al. Durbin-Watson			
	DLR	**=**	**1.9066482**

CONS Model Interpretation

The final model is the F-15E CONS model. The results of the panel model with robust standard errors for this data are presented in Table 19.

Table 19. F-15E CONS Model Regression Results

F-15E CONS Panel Model					
Fixed-effects (within) Regression				Number of obs =	240
Group variable (i): base_index				Number of groups =	5
R-sq:	within	**0.339**		Obs per group: min =	48
	between =	**0.676**		avg =	48
	overall =	**0.334**		max =	48
				F(19,394) =	**6.470**
				Prob > F =	**0.000**

CONS Rate	**Coef.**	**Robust Std. Err.**	**t-stat**	**P>|t|**
PPI Aerospace	**10.25**†	7.099	1.44	0.150
Total Avg Sortie Duration	**-238.79*****	35.949	-6.64	0.000
ZBT Program Change	**106.09**†	71.045	1.49	0.137
War Dummy	**-12.397**	70.430	-0.18	0.860
Commercial Jet Fuel for Resale	**-1.182**	1.025	-1.15	0.250
Avg Aircraft Age	**-4.16*****	1.557	-2.67	0.008
Mean Temp Difference	**-17.64***	9.107	-1.94	0.054
November Dummy Variable (DV)	**61.802**	58.342	1.06	0.291
December DV	**96.29***	57.569	1.67	0.096
January DV	**152.38*****	56.549	2.69	0.008
February DV	**173.81****	71.920	2.42	0.016
March DV	**253.24**†	157.580	1.61	0.109
April DV	**133.15****	61.751	2.16	0.032
May DV	**130.33***	66.861	1.95	0.053
June DV	**208.22*****	78.249	2.66	0.008
July DV	**151.09****	69.098	2.19	0.030
August DV	**318.41*****	88.722	3.59	0.000
September DV	**595.21*****	103.688	5.74	0.000
Constant	**103.482**	992.406	0.1	0.917

*** significant to 0.01 level, ** 0.05 level, * 0.10 level, † 0.15 level

In examining this model, it is apparent that it too is better at accounting for the between variation than the within. The R^2 (0.676) for the between is higher than any other R^2 in any of the other models. As with all the other models, except the F-15E DLR model, this model shows a distinct seasonal/business cycle component to it. Again, the seasonality/business cycle identified has close ties to the Air Force's FY quarters. Additionally, Tot_Avg_Dur is very significant (p-value < 0.001) with the magnitude of the coefficient being noteworthy also (-238.79). Although Avg_Age has a low p-value (0.008), the coefficient's magnitude is small (-4.16) and not significant. Also, the sign on this coefficient does not follow the previous research findings that as an aircraft ages, the maintenance costs also increase (Hawkes, 2005:15). Lastly, the Prog_Chng variable is only slightly significant (p-value = 0.137), but the coefficient's magnitude is considerable when compared to the series mean of $772. As with the F-15CD CONS and F-15E DLR models, Mean_Diff is significant (p-value = 0.054) and negative, again this is counter-intuitive. Below is the finalized equation for the model:

$$
\begin{aligned}
CONS_{it} = {} & -238.79(TotAvgDur_{it}) + 106.09(ProgChng_{it}) - 4.16(AvgAge_{it}) \\
& - 17.65(MeanDiff_{it}) + 96.29(DecDmy_{it}) + 152.38(JanDmy_{it}) \\
& + 173.81(FebDmy_{it}) + 253.24(MarDmy_{it}) + 133.15(AprDmy_{it}) \\
& + 130.33(MayDmy_{it}) + 208.22(JunDmy_{it}) + 151.09(JulDmy_{it}) \\
& + 318.41(AugDmy_{it}) + 595.20(SepDmy_{it}) + \varepsilon_{it} \qquad (9)
\end{aligned}
$$

Independence of Residuals

The Durbin-Watson statistic for this model is displayed in Table 20. The statistic is below 2, but not far enough away that would cause major concern. Based on this measure, the residuals are believed to be independent.

Table 20. F-15E CONS Model Durbin-Watson test—First Order Autocorrelation

F-15E Fleet		
modified Bhargava et al. Durbin-Watson		
	CONS =	**1.8858912**

Validation Testing for F-15E Models

The results of the validation tests; regression, mean absolute error (MAE), and mean absolute percent error (MAPE) for the two F-15E models are presented in Table 21. As with the F-15CD fleet models, the accuracy measures were not very strong. The depot level Reparable (DLR) model performed better than the F-15CD DLR model (R^2 = 0.060), but only by a small margin. The MAE measure was about half the average DLR rate ($7103.91) for this series and the MAPE is just as poor at an 83% error rate. Alternatively, as with the F-15CD consumables (CONS) model, the F-15E CONS model performed better. The MAE of $204.80 was only about one-third the value of the average of $772.28 and the MAPE was only 26%. Even though these measures were not, at first look, very robust, they were for monthly predictions which were at the most micro level of measurement for cost per flying hour data. As stated with the F-15CD models, when these models were compared to the current available models, they performed as well or better in most cases.

Table 21. F-15E Fleet Summation of Accuracy Measures

Accuracy of Panel Data Model				
Measures				
Model	**R^2**	**Adj R^2**	**MAE**	**MAPE**
F-15E DLR	**0.1140**	**0.0987**	**4,308.30692**	**83.97277**
F-15E CONS	**0.3242**	**0.3125**	**204.79888**	**26.02951**

Comparison against Currently Available Models

The most important step in assessing the usability of the developed models is to perform a basic comparison of the accuracy measures with those that are currently available. The current models used for comparison were the model by Hawkes (2005) and the Physics Based Model (2000). Within the Physics Based Model literature, there are data showing how well the proportional model (current model used in 2000) performed. The proportional model, as explained in the Physics Based Model literature, uses flying hours to predict maintenance needs (removals). To forecast future flying hour costs, this model uses a historical CPFH rate and multiplies it by the forecasted hours. Thus, the performance of this research's models will be compared to the proportional model also.

The Physics Based Model was used for several different Mission MDSs; however, it was only used for one fighter aircraft, the F-16C. Hence, the comparison between this model and the panel models will be limited to the F-16C and no others. Each of these models were discussed at length in the review of literature. Hawkes' model was built and measured based on yearly data, so the comparison for the panel model results was yearly. The Physics Based Model and the proportional model are built on 60 months of data; separated into three periods or calibration sets. The length of the calibration sets were 20, 19 and 20 months; subsequently, their accuracy measures for these three calibration sets were considered approximately yearly. Therefore, these models were compared to the F-15 models yearly measures.

Table 22 depicts the panel models, F-15CD and F-15E DLR and CONS, in comparison to the current available models.

Table 22. Comparison Against Currently Available Models[2,3]

Comparison Against Currently Available Models				
		Yearly		
Panel Model		**MAPE**	**Relative Error**	**RMS Relative Error**
F-15CD DLR		**16.0**	**16.1**	**2.3**
F-15CD CONS		**6.7**	**6.1**	**0.3**
F-15E DLR		**11.7**	**10.5**	**0.9**
F-15E CONS		**10.5**	**6.0**	**0.3**
Hawkes (2005)				
F-16CD DLR		**15.4**		
Physics Based				
F-16C Removals	**Set 1**		**-24.0**	**24.7**
	Set 2		**-1.8**	**10.3**
	Set 3		**-1.2**	**9.8**
Proportional Model				
F-16C Removals	**Set 1**		**23.5**	**29.7**
	Set 2		**25.4**	**31.5**
	Set 3		**14.2**	**22.1**

[2] The error estimates in this thesis are derived form underlying monthly estimates, hence are subject to more error than the 12 month estimates by Hawkes [2005]. This overstates the comparison of the MAPE in the model presented in this thesis with that of Hawkes.

[3] Forecast comparisons were made using reported Relative Error, RMS, and MAPE form earlier studies. It is difficult to make a full set of forecast comparisons without the underlying data (which would allow for a broader set of comparables). One notable outcome is that for comparisons which consistently over or under predict costs, the RMS may be greater than the Relative Error. Whereas those that fluctuate between over and under predicting, the RMS can be smaller than the Relative Error.

Importantly, from a managerial aspect, both the direction and the magnitude of the error matter. While it is important to pursue accuracy, errors over-predicting costs are far less onerous than errors under-predicting costs. These cost overruns can have a significant impact on the budgeting process and the operational readiness of the USAF. Therefore, this research's models are well suited to be used in the budgeting process, because they overestimate the actual costs, but not to the severity of the other models. Table 22 clearly identifies this research model's forecast accuracy far exceeds the proportional model, and in most cases the Physics Based model[4]. Perhaps most importantly, the forecasts in this research's models captured both the short run dynamics (as evidenced by the very low RMS values) and the steady-state dynamics (as evidenced by the low relative error values). However, where the errors in relative error were made, they were on the over-predicting costs side. Again, form a managerial aspect; this is an improvement over the existing monthly or quarterly models, if available.

Summary

The purpose of this chapter was to fully describe the processes used to answer the research questions and ultimately the research objective. First, the theoretical models were described, to include the intuitive direction of impact on the dependent variable and the analysis of the correlation matrices. Then, the model pre-estimation assumptions, stationarity and lag structure, were described in detail. After the pre-estimation

[4] There is likely no statistical difference in the comparison with the Hawkes model, but since that research did not examine Consumables or alternative models, and was limited to annual estimates, the conclusion is difficult to make with certainty.

assumptions, the individual panel model results with the corresponding post-estimation test were presented. This included the detailed description of the final models and the implications of each model. Lastly, the models were assessed for accuracy, first using the common measures of MAE and MAPE and then in comparison to previous models. Overarching conclusions will be presented in the next chapter.

V. Conclusions and Recommendations

Chapter Overview

This final chapter ties the previous chapters together by describing how the analysis and results of Chapter 4 were used to answer the research questions identified in Chapter 1. Next, this chapter discussed the overall conclusions of this research to include how well the research performed in reaching the research objective. This performance will also be summed up in the significance of the research. Lastly, two areas of recommendations will be addressed, action and future research. The last recommendation represents the personal desires of the research team regarding the direction and essence of related future research.

Discussion of Research Questions

Is there a seasonal trend/business cycle for the F-15 fleet CPFH rates?

In three of the four models developed by this research, there was significant evidence of a cyclical/seasonal component within the CPFH data. The only model that did not show evidence was the F-15E DLR model. There is no apparent reason this model did not show this cyclical/seasonal component. In the other three models, it was also evident the cyclical pattern matched that of the USAF's quarterly budget pattern. This was supported by the coefficients, as an aggregate, were greater in the fourth quarter of a fiscal year than in any other quarter—intuitive since a majority of the expenditures occur in the last quarter of the year along with "fall-out" money. The second quarter on aggregate was higher than the third quarter, which also is intuitive because historically the authority to execute the budget (i.e. the bases finally get the money loaded to spend it) occurs late in the first quarter or early second quarter. Therefore, bases increase their spending in the quarters just identified—second and fourth. Figures 4 thru 7 also support the evidence of a seasonal or cyclical component in the cost per flying hour program.

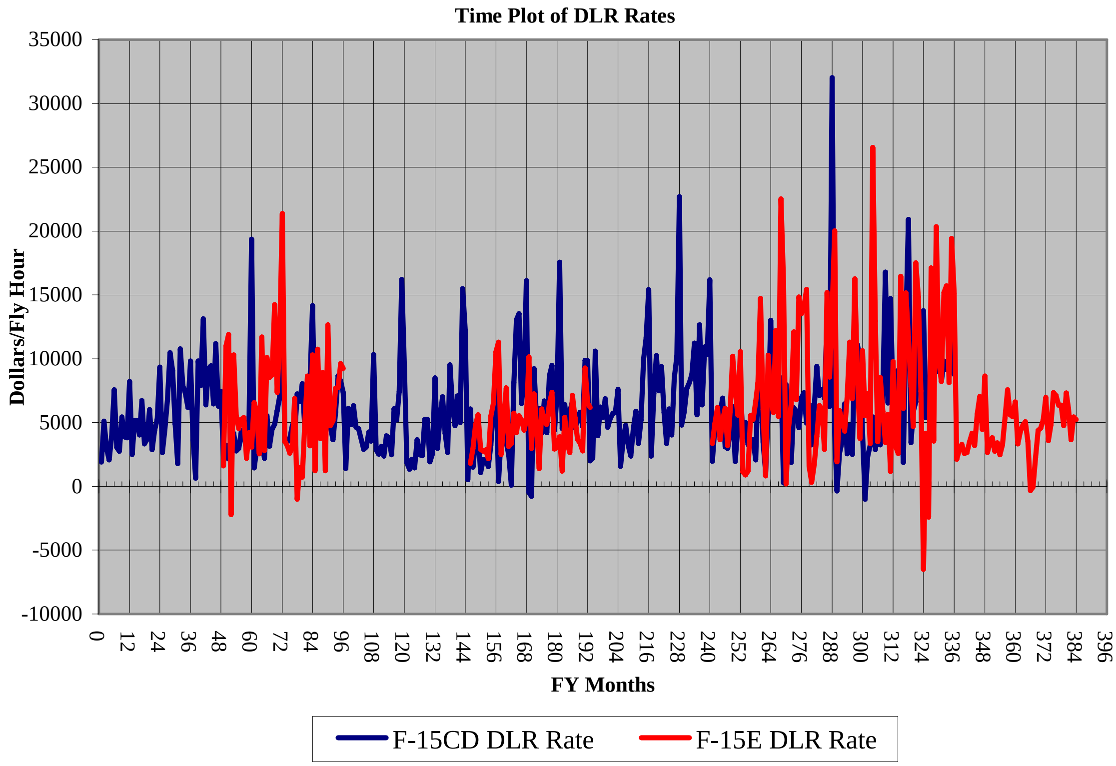


Figure 4. Time Plot of DLR Cost Per Flying Hour Data

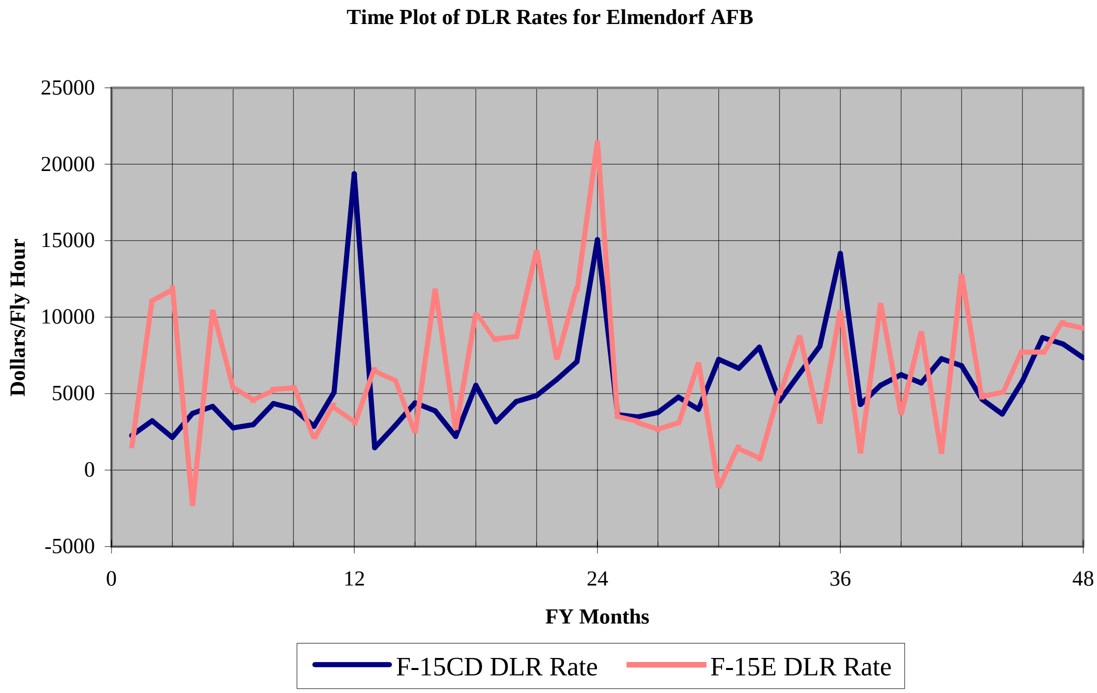


Figure 5. Time Plot of DLR Cost Per Flying Hour Data-Elmendorf AFB

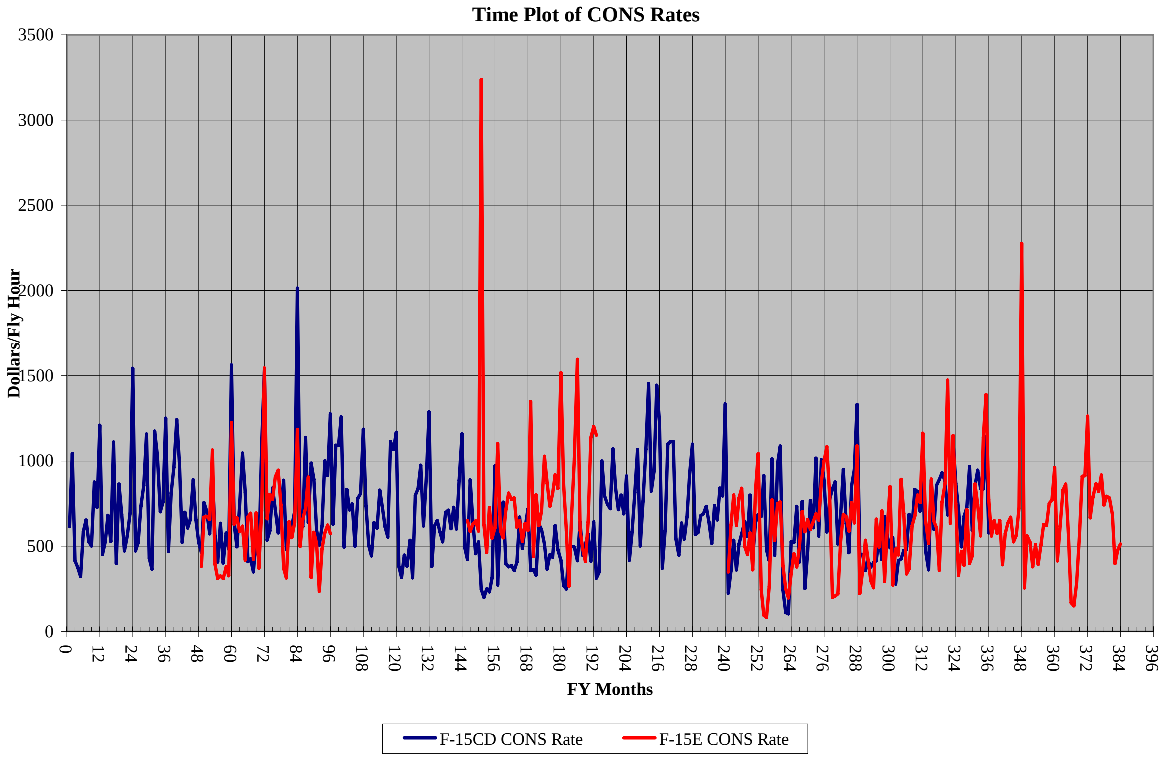


Figure 6. Time Plot of CONS Cost Per Flying Hour Data

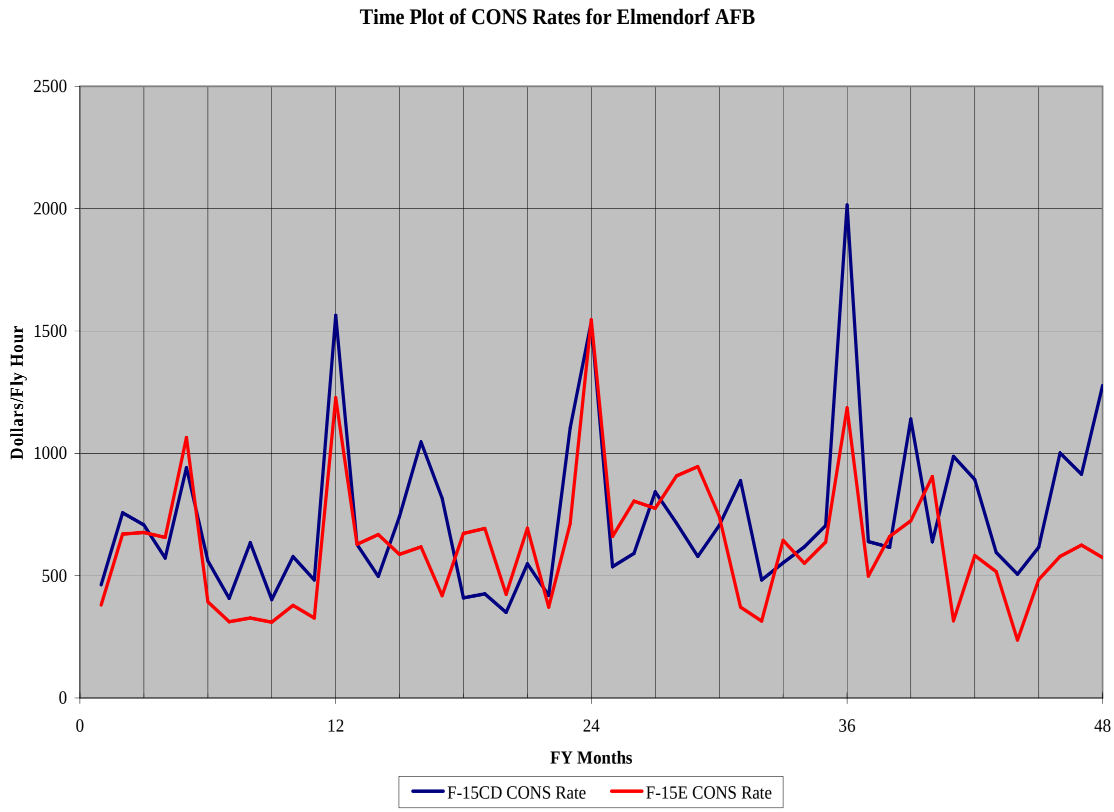


Figure 7. Time Plot of CONS Cost Per Flying Hour Data-Elmendorf AFB

Does the monthly average temperature and salinity at a location influence the F-15 fleet CPFH rates?

This research was unable to unequivocally answer this question because of the inability to find a more robust measure of salinity. Since the proposed measure, a binary dummy variable, was used to proxy for the approximaty of the base to an ocean it was a time invariant (does not change with time) variable and was unable to be used in the panel model. If there could be a measure of salinity that changes over time, as in percentage salinity by month, then this variable could be measured for its significance.

Additionally, if this research was able to obtain the deployment data from each of the bases, the location of the deployment would more than likely have changed the binary variable throughout the data set (making salinity time variant). However, the average monthly difference in temperature variable was significant in three of the four models with the magnitude of the variable being significant. Counter-intuitively, though, the sign of the coefficient was negative. One would expect just the opposite would occur. However, deployment cycles could have influenced this variable significantly during this time period.

Does the average age of the aircraft have an effect on the F-15 fleet CPFH rates?

Based on the results of this research, the average age of an aircraft was not found to be statistically significant in the F-15CD CONS and F-15E DLR models while significant in the F-15CD DLR and F-15E CONS models. Yet, in these last two models, that found average age to be statistically significant, the economic magnitudes of the coefficients was only significant in the F-15CD DLR model. For that reason, this research finds inconclusive evidence that the average age of the aircraft impacts the F-15CD and E fleets' DLR and CONS CPFH rates.

Can an aggregate model be developed for the entire F-15 fleet by MDS?

In the previous chapter, the research models were compared to the currently available models. Based on this comparison of the accuracy measures, a generalized model (panel data) can be used to accurately forecast the DLR and CONS CPFH rates for the Air Force's F-15CD and E fleets. These models either are as accurate as or better than the compared models.

With the answers to the research questions as support, the next question is does this research answer the overall problem statement from Chapter 2: Can an aggregate model be developed for the entire F-15 fleet by Mission Design Series.

Conclusions of Research

Based on the answers to the research questions above, the results from Chapter 4, and the comparison of these models to the currently available models, it can be concluded that the development of an aggregate "marginal CPFH" model can be constructed. Such that, if a Command flies in excess of its PB (programmed baseline) direct hours, the additional funding to pay for contingencies etc. is commensurate with the additional cost for the extra hours flown, not the full value of a flying hour for that weapon system. These models significantly outperformed the current models in almost all cases. In the cases they did not perform as well, they were relatively close to the existing models performance. The remarkable performance of the model presented in this research could be the result of outliers in the comparable periods for the other models, or a similar anomaly. This research effort was able to successfully answer the overall objective, but what is the significance of this research?

Significance of Research

The significance of this research can be found in several different aspects. First, this research proved there is a significant cyclical/seasonal component to the CPFH rate, something that was not previously investigated. Another significant finding was the identification of the mean average difference in temperature has a significant affect on the DLR and CONS CPFH rates, but in a negative way. Most importantly, this research demonstrated there is the capability to forecast the CPFH DLR and CONS rates at an aggregate level using panel data. This allows the analyst to study smaller time-series data sets and still provide robust analysis. This will be significant if a specific time frame needs to be isolated, but only occurs over a short time period. Additionally, this method allows the analyst to use aggregated data, quarterly and yearly, to perform analysis on without having to have a large number of observations and losing degrees of freedom.

From a purely managerial aspect, this research provides the decision maker a tool to better manage their cost per flying hour program. Also, these models lend themselves to be used successfully in war simulation exercises in accurately predicting the cost of the additional flying hours. Even though this research had several significant findings and is the best performing forecast model for the F-15 cost per flying hour program, there is always room for improvement and expansion of the research focus.

Recommendations for Future Research/Actions

Six recommendations for future research/actions are offered as a result of this study. First, expanding the panel model to analyze more Air Force MDSs would be worthwhile. If this model can be used to accurately predict CPFH rates for the F-15 CD and E fleets, can it be applied to other airframes? This researcher believes it can be applied to all the Air Force's airframes.

Second, including the Aviation Fuel portion of the CPFH rate would be of great benefit, especially with the drastic changes in the world oil markets. This would also provide a model that is all encompassing; includes the DLR, CONS, and AVFUEL portions.

Next, a deeper investigation into the effects of climatology on the CPFH rate at each base would be significant. With this, a need for a better measure of the salinity associated with each base. A percent salinity would be the optimum measure to determine if in fact salinity and temperature do impact the CPFH rates at a base. Or, if the deployment data for each base could be obtained, then this data would surely make the salinity variable change over time and therefore could be used in the analysis.

Fourth, an investigation into whether the variable "war" has an impact on the overall CPFH rates is warranted. This would require the acquisition of data that does not include times of conflict.

Fifth, although this research investigated numerous explanatory variables, finding a few of them to be significant, this is by no means an all exhaustive analysis. There is need to investigate even further the events/factors that impact the CPFH rates. The investigation should start at the lowest level, base/wing, and then move its way up to a more aggregated level. One possible route to research these factors is to survey those analysts in the field that have been working the CPFH program. These individuals have first hand knowledge on the most significant factors impacting their CPFH rates.

Finally, the ultimate output of this research would be a graphically interfaced model that can be fed down to the base level for analysts to use. Providing this capability to the lowest level of analysis would provide them the capability to accurately forecast the marginal CPFH. This gives them the ability to provide invaluable budgetary analysis to their Commander and to the MAJCOM.

Summary

This research investigated the capability of a panel model to accurately predict the cost per flying hour (CPFH) rates of the Air Force's F-15 CD and E fleets using readily available data from FY01 to FY05. In doing so, it constructed the most accurate forecast model currently available. This research effort expanded the current knowledge of CPFH explanatory variables by concluding there was a significant cycle/seasonality component to the depot level Reparable (DLR) and consumables (CONS) CPFH rate. In addition, it was found that the ZBT program change had a significant impact on all the models with the exception of the F-15CD CONS model (it was close to being significant with a p-value of 0.021). Furthermore, this research ascertained that average of the aircraft was not, overall, a significant determinant of CPFH rates. Lastly, this research solidified the notion that average sortie duration, as a whole, significantly impacts the CPFH rates for DLRs and CONS. Overall, this thesis provides analysts and decision makers a robust and defendable tool to analyze and predict the CPFH rates for the F-15CD and E fleets.

Appendix A: List of Acronyms

ACC	Air Combat Command
AETC	Air Education and Training Command
AFCAA	Air Force Cost Analysis Agency
AFCAIG	Air Force Cost Analysis Improvement Group
AFCCC	Air Force Combat Climatology Center
AFKS	Air Force Knowledge Services
AFRC	Air Force Reserve Command
AFSOC	Air Force Special Operations Command
AFTOC	Air Force Total Ownership Costs
AIC	Akaike Information Criteria
AMC	Air Mobility Command
ANG	Air National Guard
ASD	Average Sortie Duration
AVFUEL	Aviation Fuel
CONS	Consumables
CPFH	Cost Per Flying Hour
DLR	Depot Level Reparable
DV	Dummy Variable
FYDP	Future Years Defense Plan
GAO	Government Accounting Office
GPC	Government Purchase Card

GSD	General Support Division
MA	Moving Average
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MD	Mission Design
MDS	Mission Design Series
ME	Mean Error
MSD	Mission Support Division
O&M	Operations and Maintenance
PACAF	Pacific Air Forces
PB	Programmed Baseline
POM	Program Objectives Memorandum
PPBE	Planning, Programming, Budgeting, and Execution
REMIS	Reliability and Maintenance Information System
RMS	Root Mean Square
SES	Single Exponential Smoothing
SRRB	Spares Requirements Review Board
VIF	Variance Inflation Factor
ZBT	Zero Based Transfer

Appendix B: Hausman Specification Test Results

F-15CD Fleet Model—DLR and CONS

DLR Model Specification

Table 23 displays the results of the Hausman Specification test. As indicated, the H_o failed to be rejected; thus, the model is properly specified with the fixed effects panel model.

Table 23. F-15CD DLR Hausman Specification Results

F-15CD DLR Hausman Specification Test				
	Coefficients			
	(b)	(B)	(b-B)	sqrt(diag(V_b-V_B))
	Fixed	Random	Difference	S.E.
consum_rate	4.347868	4.193739	-0.1541291	.
ppi_aero	-385.9973	106.9811	492.9784	.
tot_avg_dur	-963.7975	-1089.902	-126.1041	214.5085
prog_chng	-1390.841	-1158.542	232.2994	.
war	1077.494	2201.278	1123.784	.
jet_fuel	39.04278	26.84676	-12.19602	3.641124
avg_age	198.5002	-4.270244	-202.7704	.
mean_diff	-11.00012	57.73199	68.73211	.
nov_dmy	877.4932	1007.575	130.0815	109.8466
dec_dmy	1696.574	1794.179	97.60559	322.3419
jan_dmy	989.1467	861.4407	-127.706	139.273
feb_dmy	1556.277	1560.633	4.355903	222.7151
mar_dmy	1996.898	1778.224	-218.6742	.
apr_dmy	-165.6649	-273.4754	-107.8105	190.2453
may_dmy	609.7575	727.6596	117.9021	.
jun_dmy	1448.455	1429.323	-19.13215	.
jul_dmy	1676.383	1517.663	-158.7202	.
aug_dmy	2433.614	2217.484	-216.1298	.
sep_dmy	4926.917	4867.811	-59.10631	.

b = consistent under Ho and Ha; obtained from xtreg
B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

chi2(19) = (b-B)'[(V_b-V_B)^(-1)](b-B)
= 1.84
Prob>chi2 = 1.0000
(V_b-V_B is not positive definite)

CONS Model Specification

Table 24 displays the results of the Hausman Specification test. As indicated, the H_o failed to be rejected; thus, the model is properly specified with the fixed effects panel model.

Table 24. F-15CD CONS Hausman Specification Results

F-15CD CONS Hausman Specification Test				
	Coefficients			
	(b)	(B)	(b-B)	sqrt(diag(V_b-V_B))
	Fixed	Random	Difference	S.E.
ppi_aero	-11.51981	0.0864405	-11.60625	23.83681
tot_avg_dur	-234.0228	-227.6946	-6.328224	9.438531
prog_chng	-70.07827	-64.06877	-6.009496	19.51137
war	49.24749	73.78812	-24.54063	51.34768
jet_fuel	0.0244294	-0.2546498	0.2790792	0.6458614
avg_age	6.072581	1.366648	4.705933	9.726796
mean_diff	-8.384272	-7.63164	-0.7526315	2.266458
nov_dmy	81.78348	83.53422	-1.750742	.
dec_dmy	210.101	210.0419	0.059083	.
jan_dmy	203.1761	198.5752	4.60096	.
feb_dmy	221.2262	219.0511	2.175112	.
mar_dmy	158.0413	150.4355	7.605823	.
apr_dmy	175.7908	173.3996	2.391221	.
may_dmy	163.2168	166.5771	-3.360262	.
jun_dmy	128.1669	127.4843	0.6825922	.
jul_dmy	156.7772	153.252	3.525181	.
aug_dmy	267.3618	261.5944	5.767419	.
sep_dmy	562.171	557.6658	4.50519	36.28008

b = consistent under Ho and Ha; obtained from xtreg
B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

chi2(19) = (b-B)'[(V_b-V_B)^(-1)](b-B)
= 0.35
Prob>chi2 = 1.0000
(V_b-V_B is not positive definite)

F-15E Fleet Model—DLR and CONS

DLR Model Specification

Table 25 displays the results of the Hausman Specification test. As indicated, the H_o failed to be rejected; the model is properly specified with the fixed effects panel model. However, the p-value is not as robust as the other models. This is probably due to the higher value for the between R^2 than the within R^2.

Table 25. F-15E DLR Hausman Specification Results

F-15E DLR Hausman Specification Test				
	Coefficients			
	(b)	(B)	(b-B)	sqrt(diag(V_b-V_B))
	Fixed	Random	Difference	S.E.
consum_rate	2.462423	1.783091	0.6793321	0.9607419
ppi_aero	251.3278	246.2832	5.044619	68.38952
tot_avg_dur	-1942.746	-2362.521	419.7749	.
prog_chng	-2526.438	-2082.812	-443.6258	.
war	660.8119	734.7067	-73.89474	.
jet_fuel	-1.311595	-3.058407	1.746812	.
avg_age	9.976045	0.4817864	9.494258	28.06572
mean_diff	-336.7388	161.1553	-497.894	60.07657
nov_dmy	441.8548	1172.591	-730.7363	.
dec_dmy	589.5491	1627.465	-1037.915	.
jan_dmy	603.6995	1174.88	-571.1806	761.5178
feb_dmy	1376.802	1683.702	-306.8999	.
mar_dmy	392.5887	-44.54753	437.1363	.
apr_dmy	390.0959	-414.1768	804.2726	.
may_dmy	1345.785	377.2367	968.5479	.
jun_dmy	2880.804	2010.631	870.1727	.
jul_dmy	1214.905	647.6361	567.2689	.
aug_dmy	1270.272	758.3354	511.9366	.
sep_dmy	1960.579	1956.073	4.506012	716.7527

b = consistent under Ho and Ha; obtained from xtreg
B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

chi2(19) = (b-B)'[(V_b-V_B)^(-1)](b-B)
= 21.58
Prob>chi2 = 0.3055
(V_b-V_B is not positive definite)

CONS Model Specification

Table 25 displays the results of the Hausman Specification test. As indicated, the H_o failed to be rejected; thus, the model is properly specified with the fixed effects panel model.

Table 26. F-15E CONS Hausman Specification Results

F-15E CONS Hausman Specification Test					
	Coefficients				
	(b)	(B)	(b-B)		sqrt(diag(V_b-V_B))
	Fixed	Random	Difference		S.E.
ppi_aero	10.25074	5.494324	4.756421	.	
tot_avg_dur	-238.7857	-227.4698	-11.31593		3.503243
prog_chng	106.0934	101.2987	4.794679	.	
war	-12.39749	-21.31983	8.922345		11.98222
jet_fuel	-1.181954	-0.9614112	-0.2205429	.	
avg_age	-4.161089	-2.149272	-2.011816		1.014735
mean_diff	-17.64494	-19.91884	2.273907		4.473704
nov_dmy	61.80243	57.92338	3.879052	.	
dec_dmy	96.29407	91.74216	4.551914	.	
jan_dmy	152.3781	152.5945	-0.2163515	.	
feb_dmy	173.8139	173.2031	0.6108433	.	
mar_dmy	253.2398	255.538	-2.298242		129.5871
apr_dmy	133.1459	136.1239	-2.978015	.	
may_dmy	130.3299	131.2031	-0.873175	.	
jun_dmy	208.2206	210.8668	-2.646252	.	
jul_dmy	151.0934	154.6518	-3.558391	.	
aug_dmy	318.4064	322.1175	-3.711161	.	
sep_dmy	595.2018	596.2114	-1.009601		42.80952

b = consistent under Ho and Ha; obtained from xtreg
B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

chi2(19) = (b-B)'[(V_b-V_B)^(-1)](b-B)
= 0.22

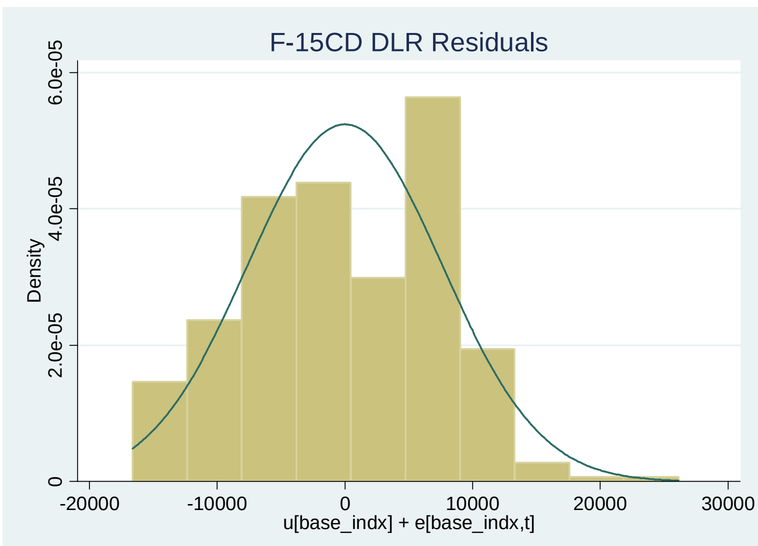
Prob>chi2 = 1.0000
(V_b-V_B is not positive definite)

Appendix C: Shapiro-Wilk W Test and Histogram of Residuals

F-15CD Fleet Model—DLR and CONS

DLR Normality of Residuals

Figure 8 displays the histogram plot and Shapiro-Wilk test for this model. This model does not meet the assumption of normality of the residuals based on the Shapiro-Wilk test statistic; however, the histogram does not look too far deviated from the normal distribution. Since this model is not being used for hypothesis testing, the deviation from this assumption is not a major concern.

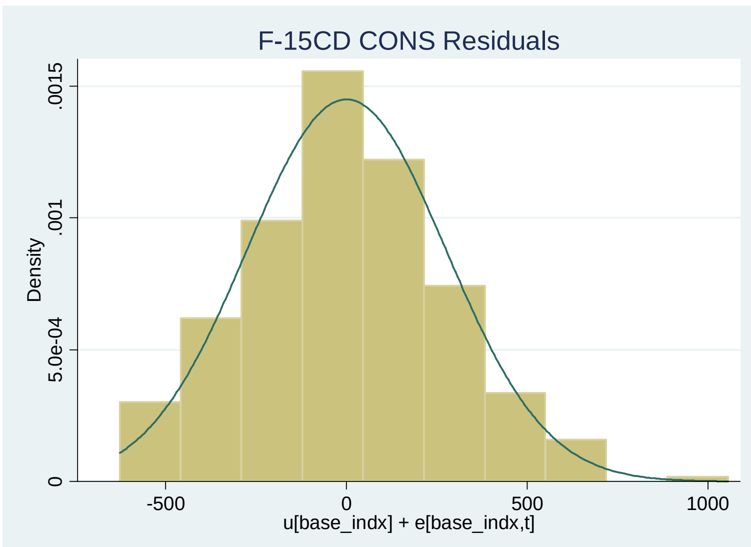


F-15CD Fleet					
Shapiro-Wilk W test for normal data					
Variable	Obs	W	V	z	Prob>z
dlr_resids	336	0.973	6.329	4.355	0.000

Figure 8. Histogram Plot of Residuals for F-15CD DLR Model

CONS Normality of Residuals

Figure 9 displays the histogram plot and Shapiro-Wilk test for this model. The visual inspection of the residuals leads to the conclusion the residuals are normally distributed. However, the Shapiro-Wilk test rejects the null hypothesis, at a 90% confidence level, but it is very close. Again, since this model is not being used for hypothesis testing, the slight deviation from normality is not a major concern.



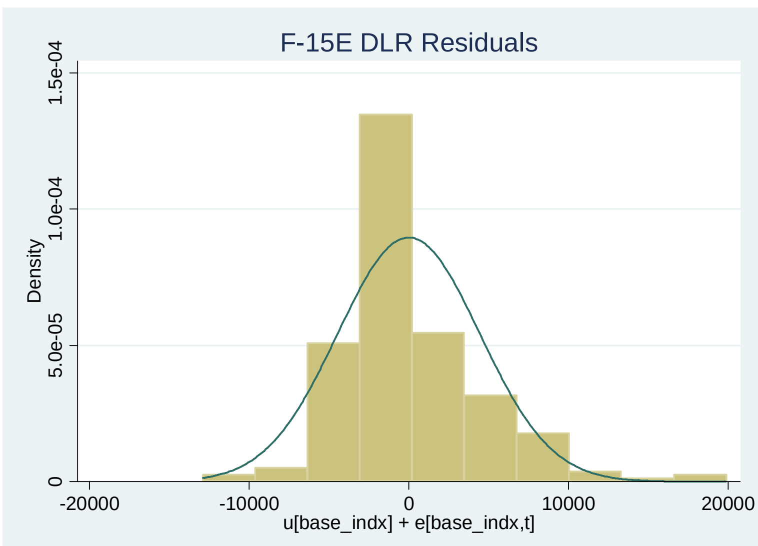
F-15CD Fleet					
Shapiro-Wilk W test for normal data					
Variable	**Obs**	**W**	**V**	**z**	**Prob>z**
consum_resids	336	0.993	1.73	1.294	0.098

Figure 9. Histogram Plot of Residuals for F-15CD CONS Model

F-15E Fleet Model—DLR and CONS

DLR Normality of Residuals

Figure 10 displays the histogram plot with a normal curve and Shapiro-Wilk test for this model. This model does not meet the assumption of normality of the residuals; it is slightly skewed to the left. As with all the other models, it is not being used for hypothesis testing so the deviation from this assumption is not a major concern.

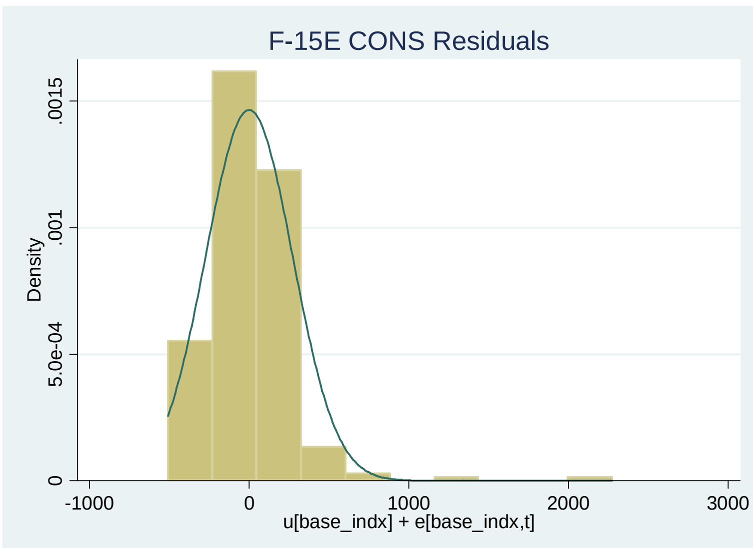


F-15E Fleet					
Shapiro-Wilk W test for normal data					
Variable	**Obs**	**W**	**V**	**z**	**Prob>z**
dlr_resids	240	0.918	14.387	6.191	0.000

Figure 10. Histogram Plot of Residuals for F-15E DLR Model

CONS Normality of Residuals

Figure 11 displays the histogram plot with a normal curve and Shapiro-Wilk test for this model. The visual inspection of the residuals shows the distribution skewed to the right due to a couple of large positive errors. This is supported by the Shapiro-Wilk test. Since this model is not being used for hypothesis testing, the possible deviation from this assumption is not a major concern.



F-15E Fleet					
Shapiro-Wilk W test for normal data					
Variable	Obs	W	V	z	Prob>z
consum_resids	240	0.825	30.644	7.947	0.000

Figure 11. Histogram Plot of Residuals for F-15E CONS Model

Bibliography

Baltagi, Badi H. and others. "Testing Panel Data Regression Models with Spatial Error Correlation," 10th International Conference on Panel Data, Berlin, July 5-6, 2002 B6-4, International Conferences on Panel Data.

Department of Defense. *The Planning, Programming, and Budgeting System (PPBS)*. DoD Directive 7045.14. Assistant Secretary of Defense, May 1984.

Edwards, Michael V. *Flight Hour Costing at the Type Commander and Navy Staff Levels: An Analytical Assessment*. MS Thesis, Naval Post Graduate School, Monterey CA, December 1992 (ADA261977).

Energy Information Administration. "U.S. Aviation Gasoline Wholesale/Resale Price by All R&G (Cents per Gallon)." Historical database. http://tonto.eia.doe.gov/dnav/pet/hist/a403700002m.htm. 7 January 2006

Faykes, Frank. "FY07 Air Force Budget." PowerPoint presentation. https://www.saffm.hq.af.mil/FMB/pb/FY07%20AF%20Budget%20Rollout.pdf. 10 February 2006.

Guo, M. and others. "Effect of Environmental Factors on the Corrosion of 2024T3 Aluminum Alloy," *Materials Forum*, Volume 28:433-438 (2004).

Hawkes, Eric M. *Predicting the Cost per Flying Hour for the F-16 using Programmatic and Operational Variables*. MS Thesis, AFIT/GOR/ENC/05-01. School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, June 2005 (ADA436138).

Hicks, Michael J. Class lecture, ECON 520, Managerial Economics, School of Systems and Logistics, Air Force Institute of Technology, Wright Patterson AFB OH, January 2005

Hinkle, D.E., Wiersma, W., & Jurs, S.G. *Basic Behavioral Statistics*. Boston: Houghton-Muffin, 1982

Kammerer, Joseph T. "Notes from the: Deputy Assistant Secretary, Cost & Economics." *The Air Force Comptroller*, 36-2:19 (Apr 2002).

Kennedy, Peter. *A Guide to Econometrics* (Fifth Edition). Cambridge: The MIT Press, 2003.

King, David R. and Massey, Donald S. “History of the F-15 Program: A Silver Anniversary First Flight Remembrance,” *Air Force Journal of Logistics*, XXI, No 1, 10-16 (Winter 1997).

Laubacher, Matthew E. *Analysis and Forecasting of Air Force Operating and Support Cost for Rotary Aircraft*. MS Thesis, AFIT/GCA/ENV/04M-05. School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, March 2004 (ADA423137).

Makridakis, Spyros and others. *Forecasting: Methods and Applications* (Third Edition). New York: John Wiley & Sons Inc., 2003.

McCourry, Larry G. “Corrosion: A Formidable Air Force Enemy,” *Air Force Journal of Logistics*, VII, No 1, 5-9 (Winter 1983).

Nordhaus, William D. *The Economic Consequences of a War with Iraq: Cowles Foundation Discussion Paper No 1387*. Cowles Foundation for Research in Economics, Yale University, New Haven CT (Dec 2002)

Rose, Pat A. Jr. “Cost Per Flying Hour Factors: A Background and Perspective of How They Are Developed and What They Do,” *The Air Force Comptroller*, 31-1:4-9 (Jul 1997).

SAF/FMC. “Cost Per Flying Hour (CPFH) Program: Statement of Intent.” Electronic Message. May 2003.

-----. “FY07 POM Action Officer Cost Per Flying Hour (CPFH) Air Force Cost Analysis Improvement Group (AFCAIG) Tutorial.” Electronic Message. 20 Jan 2005.

Stata Corp.. *Longitudinal/Panel Data*. *Stata Statistical Software: Release 9*. College Station TX: Stata Corp LP, 2005.

Tetmeyer, Donald C. *Forecasting Wartime Resource Requirements*. USAF Logistics Capability Assessment Symposium, 1982, IV122-IV136.

United States General Accounting Office. *Defense Acquisitions: Air Force Operating and Support Cost Reductions Need Higher Priority*. Washington DC: Government Printing Office. GAO/NSIAD-00-165. August 2000

-----. *Defense Budget: Observations on the Air Force Flying Hour Program.* Washington DC: Government Printing Office. GAO/NSIAD-99-165. July 1999

Wallace, John M. and others. *A Physics Based Alternative to Cost-Per-Flying-Hour Models of Aircraft Consumption Costs*, Logistics Management Institute, August 2000 (ADA387273).

White, Halbert. "A Heteroskedastic-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity." *Econometrica*, 48: 817-838 (May1980)

Wu, Ming-Cheng, *Estimating Operating and Support Models for U.S. Air Force Aircraft*. MS Thesis, Naval Post Graduate School, Monterey CA, March 2000 (ADA376488).

Vita

Captain Patrick Armstrong graduated from Green Mountain High School in Lakewood, Colorado. He then entered the United States Marine Corps Reserve for five years. In 1990, he enlisted in the active duty Air Force and was assigned Detachment 21, Belle Fourche, South Dakota as a ground radar technician. His next assignment was as an Airman Leadership School Instructor at Ellsworth AFB, South Dakota. It was here where he finished his undergraduate degree from Black Hills State University. He graduated Summa Cum Laude with a Bachelor of Science in Business Administration in December 1999. He was accepted and attended Air Force Officer Training School in 2000 and received his commission in January 2001, graduating as a Distinguished Graduate.

His first assignment as an officer was at Spangdahlem AB, Germany, as the Deputy Financial Services Officer and later that tour served as the Deputy Budget Officer. While he was in Germany, he deployed to Iraq in support of OPERATION IRAQI FREEDOM. He served as the 332nd Air Expeditionary Wing's Financial Officer while deployed. In August 2004, he entered the Graduate School of Engineering and Management, Air Force Institute of Technology, to obtain his Masters in Cost Analysis. Upon graduation, he will be assigned to the Air Force Cost Analysis Agency in Crystal City, Virginia.

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 074-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From – To)*
23-03-2006	Master's Thesis	October 2004 – March 2006

4. TITLE AND SUBTITLE	
Developing an Aggregate Marginal Cost Per Flying Hour Model for the U.S. Air Force's F-15 Fighter Aircraft	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	**5d. PROJECT NUMBER**
Armstrong, Patrick D., Captain, USAF	**5e. TASK NUMBER**
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)	**8. PERFORMING ORGANIZATION REPORT NUMBER**
Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way, Building 640 WPAFB OH 45433-7765	AFIT/GCA/ENV/06M-01
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	**10. SPONSOR/MONITOR'S ACRONYM(S)**
Mr. Tom Lies Air Force Cost Analysis Agency 201 12th Street Arlington, VA 22202 COMM (703) 692-6014 DSN 222-6014	**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

12. DISTRIBUTION/AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This thesis applies econometric techniques to build a "marginal" cost per flying hour model for the U.S. Air Force's F-15CD and E fleets. It used monthly economic, programmatic, operational, and climatology data from FY01-FY04 to construct Depot Level Reparable (DLR) and Consumable (CONS) models on the aggregate level. It incorporated the use of panel data analysis to explore the effect each of the independent variables had on the CPFH rate by time and by base. This allowed it to capture not only the temporal (time) interactions, but also the spatial (cross-sectional) interactions, providing a more robust analysis of the dynamics between the independent variables, bases, time and the CPFH rates. It discovered the DLR and CONS CPFH rates have a significant business cycle/seasonal trend component. Also, the following variables were found to be statistically and economically significant: average sortie duration, mean monthly temperature difference, and the Zero Base Transfer CPFH program change. These models when compared to the currently available models significantly out performed these models. On average, the relative error rate for this research's models was half that of the current models. Therefore, an aggregate CPFH model can be developed to accurately forecast the CPFH rates.

15. SUBJECT TERMS

Cost analysis, Cost per Flying hour, F-15, econometric modeling, panel model, time series analysis, O&M Costs, seasonal component, business cycle trend, DLR, CONS, marginal costs, cross-sectional, longitudinal

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					Dr. Michael J. Hicks (ENV)
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(Include area code)*
U	U	U	UU	103	(937) 255-3636, ext 4605 (michael.hicks@afit.edu)

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18