

Pilot development: an empirical mixed-method analysis

Pilot upgrade training

Jonathan Slottje, Jason Anderson, John M. Dickens and
Adam D. Reiman

*Department of Operational Sciences, Air Force Institute of Technology,
Wright-Patterson AFB, Ohio, USA*

21

Received 1 October 2021
Revised 28 January 2022
Accepted 10 February 2022

Abstract

Purpose – Pilot upgrade training is critical to aircraft and passenger safety. This study aims to identify variances in the US Air Force C-130J pilot upgrade training based on geographic location and provide a model to enhance policy that will impact future pilot training efforts that lower cost and increase operator quality and proficiency.

Design/methodology/approach – This research employed a mixed-method approach. First, the authors collected data and analyzed 90 C-130J pilots' aviation records and then contextualized this analysis with interviews of experts. Finally, the authors present a modified version of Six Sigma's define–measure–analyze–improve–control (DMAIC) that identifies and reduces the variances in C-130J pilot training, translating into higher quality outcomes.

Findings – The results indicate significant statistical variances across geographically separated C-130J pilot training organizations. This leads some organizations to have higher proficiency levels in specific tasks and others with comparative deficiencies. Additionally, the data analysis in this study enabled a recommended number of flight hours in several distinct categories that should be obtained before upgrading a pilot to aircraft commander to enhance standards.

Research limitations/implications – This research was limited to C-130J pilot upgrades, but these results can be implemented within any field that utilizes hours as a measure of experience. Implications from this research can be employed to scope policy that will influence pilot training requirements across all airframes in civilian and military aviation.

Originality/value – This research proposes a process improvement methodology that could be immediately implemented within the C-130J community and, more importantly, in any upgrade training where humans advance into higher echelons of a profession.

Keywords C-130J, Upgrade, Variance Reduction, Six Sigma DMAIC

Paper type Research paper

Introduction

Today, many companies have hubs or franchises spread worldwide (e.g. McDonald's, Toyota and Foxconn). Due to enhanced competition, customers expect low variability in product quality, whether that is a fast-food meal, a critical component for an automobile or a smartphone (Vardhan, 2021). To be and remain competitive, companies must ensure employees consistently produce quality, regardless of geographic location. Aviator production is similar in that variability in quality should be low irrespective of geographic location. Air Force regulations ensure critical training events are accomplished by each co-pilot, no matter where they are stationed, but a lack of standardization in training exists, thereby increasing the variability in proficiency, training and ultimately capability. This variability puts both aircrew and passengers at greater risks (Davis, 2017). To address this problem, this research analyzes C-130 co-pilot upgrades with six units across four distinct locations. We highlight training



variances and provide a method to reduce those shortcomings and enhance quality outcomes that impact aircraft, crew and passenger safety.

A US Air Force C-130J co-pilot must obtain 700 flight hours before reaching upgrade eligibility to aircraft commander, i.e. a fully upgraded pilot status. Co-pilots are continuously monitored through developmental phases, and continual instructor pilot feedback is provided to leadership to ensure standardization and proficiency. Once the co-pilot has accumulated 700 h in the C-130J mobility aircraft and other prerequisites have been met, the final approval to upgrade the individual resides with unit leadership. Besides the number of flight hours, other qualitative inputs, such as human perception of individual performance, are considered by leadership. When paired with the fact that the quality of the hours is not effectively measured, co-pilots could be prematurely upgraded, which may lead to aircraft and passenger safety risks. For instance, a co-pilot can receive upgrade hours while not at the controls of the aircraft. Those “observer” hours degrade the quality of the training. Moreover, it is possible to have pilots receive more instrument time, nighttime or night vision goggle (NVG) time than other pilots by a significant amount. Yet, all pilots will be equally qualified.

While information is provided to unit leadership on co-pilot determinations, there is a need to provide a formal approach to this process. By implementing the Define–Measure–Analyze–Improve–Control (DMAIC) methodology introduced in this paper, we offer the quantitative decision tools to improve the C-130J upgrade process that could also be employed across the mobility enterprise and the multi-national civilian aviation sector.

Historically, practitioners have employed minimum quality standards across multiple industries to include healthcare (Rooney and Van Ostenberg, 1999), manufacturing (Degirmenci *et al.*, 2013) and commercial trucking (Shortliffe, 2009). In retrospect, this iterative approach has not been fully realized in the US Air Force co-pilot upgrade training program. This research reveals that this unduly constrained training program results partially from a single standardized flying hour metric, which could unintentionally mask other deficiencies that may lead to the safety of flight concerns. This oversimplified approach requires a holistic redesign to focus on both quality and standardization across a variety of skills. However, the primary motivation behind this research stems from an upgrade training concern identified in the civilian aviation sector. In 2010, legislators passed a new law that increased minimum flight time requirements for civilian co-pilots hired by US air carriers. Consequently, this law sparked a debate on the controversial refocus towards the number of hours accumulated versus the quality of hours while in co-pilot upgrade training (Werfelman, 2010). Opponents of this law believe that the quality of flight hours should be the measurement of experience, not a specific quantity. The results of these debates led to the Federal Aviation Administration (FAA) establishing minimum hours required in specific categories of flight hours. While this practice has been employed in civilian aviation, it has not yet been adopted by its military counterparts. Thus, this research sought to examine the relationship between flight hour quantity and quality in a military aviation context. This research intends to illustrate the existence of statistical differences in co-pilot upgrade training among six individual flying units, which could create the conditions for increased safety risks. Additionally, this research aims to pinpoint the significant variances between individual pilots within the respective flying units, thus giving leadership a roadmap to correct these deficiencies across many upgrade training categories (e.g. instrument flying, night flying, NVG flying, etc). More compelling, these variances often impact the ratio of quality flight hours individual co-pilots are obtaining. Ultimately, the C-130J flying community lacks a standardized method to measure the quality of flight hours each co-pilot obtains, which leads to an unintended risk-assumption condition where aviators may achieve higher training status without the necessary prerequisites. By analyzing 90 upgrade training records, this research

established an average percentage of total hours for 11 categories of flight hours within the C-130J community. These averages can be employed as minimum training thresholds that decision-makers can use to gauge co-pilot training progress where experience gained can be more quantifiably measured and assessed.

Variance reduction and quality products are terms used extensively in the manufacturing discipline (Flynn *et al.*, 1997). However, these are nonstandard terms in the C-130J pilot production process. Leading manufacturing companies have turned to Six Sigma to enhance productivity and products (Chakraborty *et al.*, 2013). Researchers have applied Six Sigma in many businesses because improvement leads to greater profits (Tonini *et al.*, 2006). When C-130J pilot production is viewed as a product, opportunities to apply Six Sigma's DMAIC methodology emerge where they may otherwise go unnoticed. Using this methodology in pilot training is a literature gap this research seeks to address. Moreover, the primary purpose of this research is to provide the C-130J community with an approach to monitor co-pilot development effectively and reduce variances that exist between flying units and their pilots. By doing so, we seek to address the following research questions: (1) Do C-130J co-pilot flight hours significantly differ between flying units? (2) What is the correct distribution of flight hours a C-130J co-pilot should obtain to reduce variance and ensure quality outcomes? (3) What insights, if any, can the Six Sigma DMAIC methodology provide in C-130J co-pilot development?

The remainder of this paper is organized as follows. First, we will cover the pertinent literature. Next, we will introduce the research methodology. Then, we will analyze the results and present our key findings. Finally, we will provide a brief discussion, conclusions and future research opportunities.

Literature background

The literature that pertains to this research is parsed into four distinct streams. First, the current C-130J co-pilot upgrade requirements are presented to provide context. The second section will focus on variances in production, including the causes and the impacts of these variances. The third stream will examine the relevant FAA literature that relates to establishing minimum flight hour requirements in various categories and discuss quality flight hours. The final stream will explain the Six Sigma's DMAIC method and how the C-130J community could employ it.

C-130J upgrade requirements

Air Force directives state flight hour prerequisites are "based on a crewmember having gained the knowledge and judgment required to safely and effectively perform assigned duties in support of the unit's mission" (AMC/A3TA, 2020, p. 46). Air Mobility Command's Chief of Aircrew Force Management Branch explained how 700 h would ensure co-pilots receive the right mix of experience and knowledge based on decades of input from subject matter experts (Personal Correspondence, December 2019). While contentious, this research does not intend to scrutinize this quality indicator but provides additional controls within 700 h to measure pilot proficiency and reduce variances. The following section highlights why the C-130J pilot community should seek ways to reduce the variances in experience gained by co-pilots.

Variance reduction in production

A vast amount of research has been conducted on reducing variation in production and supply chains (Elmuti, 2002; Flynn *et al.*, 1997; Flynn and Flynn, 2005). Due to safety of flight concerns, identifying variances in pilot production is also essential. Nolan and Provost (1990)

explained that managers must interpret variation within their organizations. These scholars describe how managers should determine whether observed variances are based on a trend or random variation (Nolan and Provost, 1990). Mackay and Steiner make a similar argument as they believe a more consistent output can improve a product's performance (Mackay and Steiner, 1997). Matson and Prusak (2003) explained that the impact of variations is not all equal for each category. They argue that managers should decide which areas need the most attention through key metrics (Matson and Prusak, 2003). Consequently, the results of this research have identified several key metrics where variation occurs in C-130J co-pilot development.

Civil aviation

Congress signed a law in 2010 increasing the minimum flight time requirement of pilots seeking to be hired by US air carriers (Werfelman, 2010). This requirement sparked debate in the civilian aviation community of whether the number of flight hours should be the correct measurement of pilot experience. Opponents of this requirement state quality of flight hours are more important than quantity (Depperschmidt, 2013). This section will cover the studies that found the quality of flight hours to be a stronger indicator of pilot experience. Additionally, the FAA regulation Title 14, Part 61, Subpart G – Airline Transport Pilots (ATP) guidance will be explicated as this will be the basis for the final recommendations of this research.

Over 20 years ago, the Senate Committee on Commerce, Science and Transportation endorsed a study concerning the adequacy of Federal standards and programs (US Congress, 1988). Although this was a comprehensive study of the civilian aviation community, a sizable portion was dedicated to pilot selection and training. The study analyzed numerous factors concerning pilots, such as age, health, experience, training programs and total time. The conclusion of this study explained that logged hours or years in a crew position do not provide fidelity or enough insight into the skill level of the operator and that other alternative measures of skill and experience should be employed (US Congress, 1988, p. 122).

Smith *et al.* (2013) conducted a study on quality flight hours, which focused on analyzing pilot performance based on previous aviation experience. Their study found that total flight hours produced inconclusive results (Smith *et al.*, 2013). The study also highlighted the quality of the experience, not the total number of hours, to be a better predictor of pilot performance. Finally, the authors stated, “that using a quantity measure of total flight hours as the predictor of success is not suitable for the aviation industry that constantly strives to improve safety and training performance” (Smith *et al.*, 2013, p. 22).

In 2012, the FAA proposed several new rules following the 2009 Colgan Air accident near Buffalo, NY (New York). This accident brought attention to air carriers' training processes and co-pilot development (Department of Transportation, 2012). In response, the FAA adopted many of the newly proposed rules, of which this research focuses specifically on Part 61, Subpart G – ATP. The DoT (Department of Transportation) research resulted in recommending a breakdown in the types of flight hours required before being eligible to apply for an ATP. The types of flight hours include total time, cross-country time, nighttime, multiengine time, instrument time and a maximum amount of simulator time (Electronic Code of Federal Regulations, 2020). Collectively, these types of performance criteria are espoused to be a better indicator of pilot quality than simply flight hours or years in a crew position.

Although the previous studies in this section recommend different measures to determine experience versus the number of hours, the FAA continues to employ the quantity of flight hours as a significant factor towards co-pilot upgrade eligibility. What the FAA did change, also adopted in this research effort, was a breakdown in the type of flight hours obtained towards the total hour requirement.

Six Sigma's DMAIC

Variance reduction is a heavily researched topic, and often a quick search for “variance reduction” will include results with Six Sigma. Six Sigma was introduced in the 1980s with the primary goal of improving products through variance reduction and has since been applied with significant improvements by some well-known companies such as: Motorola, General Electric, Black and Decker, and Bombardier (Klefsjo, Wiklund and Edgeman, 2001; Caulcutt, 2001). Tang *et al.* (2007) explained that “Six Sigma makes use of sound statistical methods and quality management principles to improve processes and products via the DMAIC quality improvement framework.” De Mast and Lokkerbol (2012) conducted an analysis of the Six Sigma DMAIC method from the perspective of problem-solving. These scholars highlight the type of problems the DMAIC method works well in improving and provide critical analysis of problems where the method proved to be ineffective (De Mast and Lokkerbol, 2012). Chakraborty *et al.* (2013) focused on implementing the Six Sigma approach to reduce the variability in food processing. DMAIC demonstrated its utility in a services environment (Chakrabarty and Tan, 2007). Our research employs the DMAIC methodology to determine if it can garner any insights into the problem of co-pilot production.

Method

This research employed a mixed methodology through three distinct phases. First was the data collection phase, which involved extracting 90 flight records from six active-duty military units spread across four locations. Next, we ran a Levene's test and the Kruskal–Wallis H test amongst the six units to test for statistical equivalency. To complete this phase, we collected subject matter experts' qualitative inputs to explain why these statistical differences occurred. Finally, we incorporated the DMAIC methodology to establish a recommendation to improve the C-130J upgrade process.

Phase I

Initially, we selected five co-pilots from each of the six units for a sample of 30 observations of recently upgraded co-pilots to aircraft commander status. This ensured we pulled observations from those individuals with 700 flight hours in the C-130J, which both standardized and maximized the sample's number of flight hours. Next, we randomly selected ten pilots from each location, with the only standardization being they were listed as a Flight Qualified Pilot. Utilizing this group of pilots in the research ensures the maximum number of flight hours, as these individuals are in the last phase of development before starting the aircraft commander upgrade.

The data analysis began by determining which flight hours would fall into each category by employing Air Force guidance. To limit the categories, the terms training and operational were utilized. For this research, training was any mission symbol starting with N2 (tactical training), N1 (training and standardization), T1 (student training), T2 (formal MWS training) and T3 (operational training). In this research, all other mission symbols were given the term “Operational.” In addition, the terms “Operational” and “Tasked” are equivalent and are used interchangeably throughout. Examples of “Operational” include positioning, repositioning, air evacuation, cargo, passenger or patients, contingency, TWCF (Transportation Working Capital Fund), SAAM (Special Airlift Assignment Mission) and channels where aircrew training is unlikely to occur. Joint airborne or air transportability training and simulator times were assigned their own categories, and the mission symbols start with M8 and Q1-3, respectively.

Once all 90 records were sorted by mission symbols, total hours for each of the 11 categories were calculated. Next, these totals were combined for each of the six units to

determine a squadron average. Finally, the total for each squadron was combined to determine a C-130J average. These averages were then utilized to set recommended minimum hours for each category and additionally move on to Phase II, statistically analyzing each squadron.

Phase II

Due to failures in normality tests, we used a nonparametric test to explore variances within the data. Consequently, we employed the Levene's test for each category. The Levene's test is utilized to verify that variances are equal for all samples when the data comes from a nonnormal distribution (Glen, 2014), and it tests variances amongst two or more groups (NIST/SEMATECH, 2020). After completing the tests for variances, a similar test was conducted to test the mean from the units within each category to determine statistical equivalence. The last step was the qualitative input of subject matter experts on the results to investigate potential causal factors. These data were then combined to define why each variation existed.

Phase III

The final phase of this research combined the statistical results and subject matter expert input to establish a DMAIC method for the C-130J co-pilot upgrade process. Figure 1 displays the model that will be employed in this research, which was developed by Pan *et al.* (2007). We chose this model as it guides the process for variance reduction. We made minor modifications to this model for specific employment of co-pilot upgrade process, and the results are discussed in subsequent sections of this paper.

Data analysis

Categories of flight hours

This section presents the first three of the eleven categories of flight hours and the results of the statistical tests. Appendix discusses all remaining categories (i.e. 4–11). The Levene's test was performed under the hypothesis that all six population (units) variances are equal at a 0.05 alpha for all categories. Next, the Kruskal–Wallis H test was performed under the hypothesis that each sample (squadron) came from the same distribution based on the means at a 0.05 alpha. If either test resulted in a rejection of the null hypothesis, post hoc analysis was employed to determine which population(s) variance or mean caused the null rejection. Finally, subject matter expert' inputs aided in determining the practical reasons for the differences.

In the following sections, each flight hour category (for categories 1–3; see Appendix from categories 4–11) will be represented with a scatter plot for all 90 co-pilots. Each figure is then broken up into six sections referencing each of the six C-130J units and is labeled accordingly. The green line on each figure is the C-130J average based on the 90 co-pilots' flight hours for each respective category (as a percentage of total C-130J flight hours). Within each unit's block, a box with a red line across the middle is displayed, indicating the unit average, while the spread of the box illustrates the standard deviation.

Category 1: primary flight time – time at the aircraft controls while actively controlling the aircraft unless logging instructor or evaluator hours

Table 1 highlights the basic statistical data derived from the 90 co-pilots, while Figure 2 visually portrays these data. The Kruskal–Wallis H test showed a p -value of 0.000, indicating a statistical difference in the mean primary flight time. The Levene test resulted in a p -value of 0.339, which fails to reject the null hypothesis, indicating no statistical difference in the

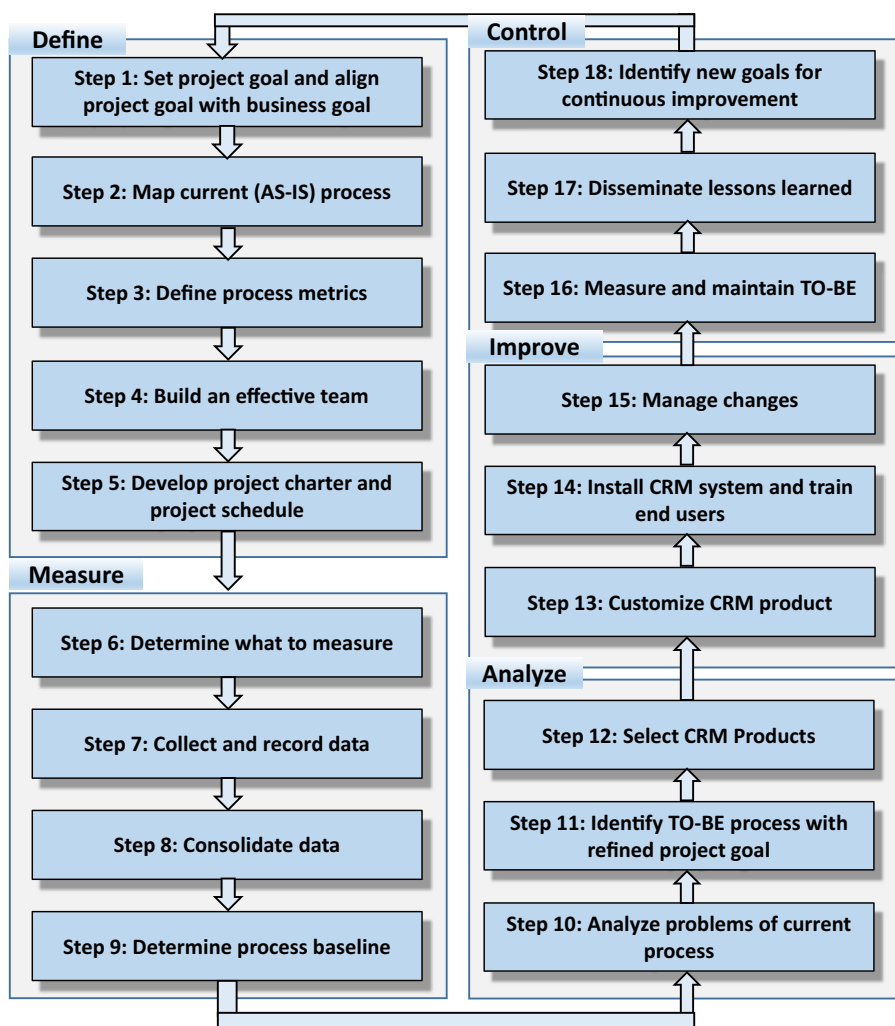


Figure 1. DMAIC method

variance of primary flight time. In practical terms, this means that Ramstein 37AS co-pilots are logging more primary flight time than co-pilots in the other five units. Through post hoc analysis, the higher-than-average mean results from lower-than-average other times that will be discussed later in this section.

Category 2: secondary flight time – time at controls but not actively controlling the aircraft, instructing or evaluating

Table 2 displays the basic statistical data derived from the 90 co-pilots, while Figure 3 visually portrays this data. The Kruskal–Wallis H test provided a p -value of 0.000, indicating a statistical difference in the mean secondary flight time. Post hoc analysis highlighted that the means of the Ramstein (37AS), Yokota (36AS) and Dyess (39AS) are lower than that of Dyess (40AS), Little Rock (41AS) and Little Rock (61AS). The Levene’s test resulted in a

p -value of 0.495 and failed to reject the null hypothesis, thus indicating that the variances are not significantly different between the squadrons' secondary flight time.

Category 3: other flight time – time not at aircraft controls, instructing or evaluating, but on the flight authorization

Similarly, Table 3 displays the basic statistical data derived from the 90 co-pilots, while Figure 4 visually portrays these data. The Kruskal–Wallis H test resulted in a p -value of 0.000, indicating a statistical difference for the mean other flight time. The Levene test resulted in a p -value of 0.000, which rejects the null hypothesis, indicating a statistical

Unit Tested	Sample		Mean of Sample	Median of Sample	Variance of Sample	Standard Deviation
	Size from Unit					
RMS37	15		70.87	73	33.27	5.77
YOK36	15		58.8	57	32.31	5.68
DYS39	15		58.8	58	64.31	8.02
DYS40	15		54.13	56	30.7	5.54
LRF41	15		54.93	57	27.78	5.27
LRF61	15		52.73	53	33.64	5.8

Table 1.
Primary flight time basic statistics

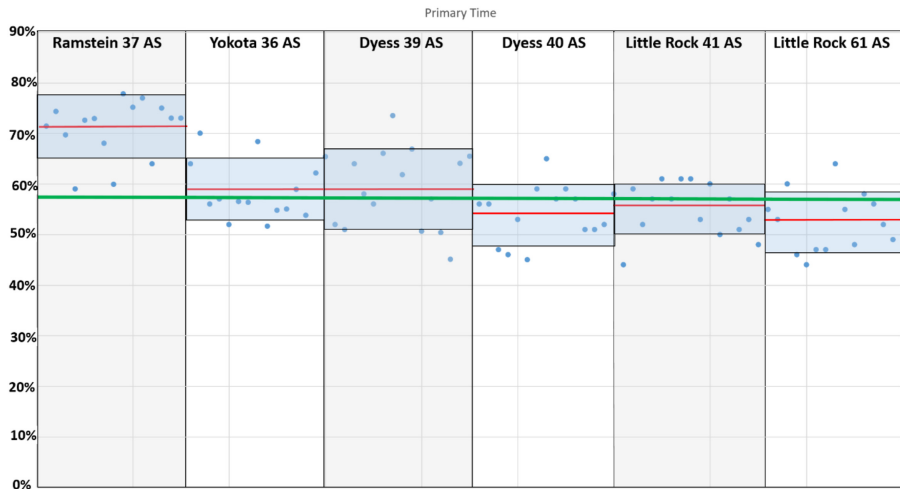


Figure 2.
Primary flight time as a percentage of total C-130J hours

Unit Tested	Sample		Mean of Sample	Median of Sample	Variance of Sample	Standard Deviation
	Size from Unit					
RMS37	15		22.27	22	31.21	5.59
YOK36	15		20.2	21	18.89	4.35
DYS39	15		24	25	39.14	6.26
DYS40	15		30	31	24.29	4.93
LRF41	15		32	32	40.14	6.34
LRF61	15		31.53	31	25.12	5.01

Table 2.
Secondary flight time basic statistics

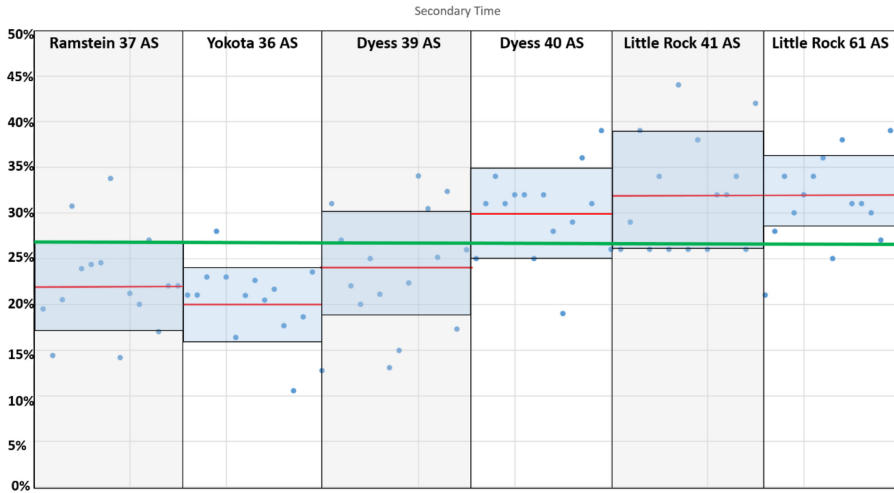


Figure 3.
Secondary flight time as percentage of total C-130J hours

Unit Tested	Sample		Mean of Sample	Median of Sample	Variance of Sample	Standard Deviation
	Size from Unit					
RMS37	15		6.8	7	7.89	2.81
YOK36	15		21.13	23	50.12	7.08
DYS39	15		17.13	17	24.27	4.93
DYS40	15		16.13	17	27.55	5.25
LRF41	15		13.07	13	8.64	2.94
LRF61	15		15.93	17	41.78	6.46

Table 3.
Other flight time basic statistics

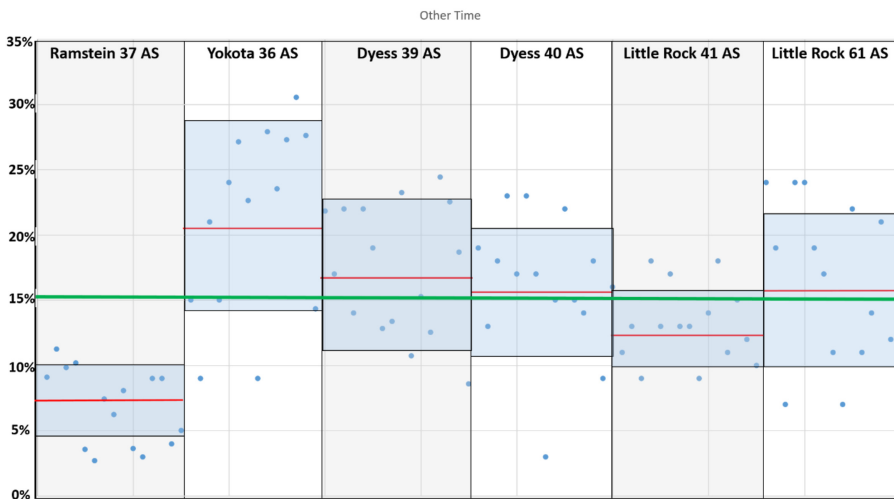


Figure 4.
Other flight time as a percentage of total C-130J hours

difference between the variance of the six squadrons' other flight times. Ramstein 37AS and Yokota 36AS differed from the other four units in their respective means. Regarding variance, Yokota 36AS and Little Rock 41AS are significantly different from the other four units. Subject matter expert insight was collected on Ramstein 37AS' low mean and variance.

First, the low secondary time mean results from a high percentage of the primary time. A pilot can attain one of three types of flight hours (primary, secondary and other). When one increases, the other two will naturally decrease. A subject matter expert provided additional factors that limit the amount of other time logged at Ramstein 37AS. Ramstein 37AS directed the squadron to limit "bleacher flights" (Personal Correspondence, March 2020). Bleacher flights are a common term used in the Mobility Air Force community and are flights where extra pilots are on board the aircraft to accomplish one or two specific training events. For illustration, imagine five pilots onboard, two primary and three extras. The extra pilots would rotate, occupying a primary crew position to accomplish their specific training event(s) and then turn the controls over to the next pilot. On a standard training sortie of 4 h, the pilot would log 0.5 h primary and 3.5 h other time. Hence, with the direction of limiting bleacher flights, Ramstein 37AS results are lower than average other time. The second reason the subject matter expert provided was the number of operational taskings and efficient scheduling focused on accomplishing Air Force requirements, contributing to a limited amount of other time (Personal Correspondence, March 2020).

From the definition within the governing regulation, other time is logged when not in a primary crew position. A C-130J co-pilot could log other time when in the augmented seat, still actively engaged in the mission, or it could be logged when in the back of the aircraft, not engaged in the mission. Other time should be limited in a co-pilot's development due to the limited experience gained. Visually depicted in Figure 4, some sampled pilots have over 25% of their total C-130J flight time as other time. This is a significant portion of total flight time and is an indicator of the quality of upgrade training. In the case of other flight time, a low mean and low variance would assist in ensuring the co-pilots are acquiring quality flight hours.

See Appendix for all remaining flight training categories and tables (Tables A1–A8), figures (Figures A1–A8), and the associate narrative.

Discussion

This section discusses the results of our analysis and will flow in accordance with each research question presented in the introduction of our paper.

- (1) Do C-130J co-pilot flight hours significantly differ between flying units?

Most of our results highlighted differences between location and units, based on both the mean and variance. This research aims to highlight where these variances are and provide data-backed tools to assist unit leaders in decision-making when it comes to co-pilot development. The following three tables provide consolidated results of this research. First, Table 4 consolidates the results of the variances produced from running the Levene's test.

	Primary	Secondary	Other	Night	INS	NVG	Training	Tasked	Simulator	JA/ATT	CMB/Spt
RMS37	33.27	31.21	7.89	7.35	59.41	1.89	54.11	60.84	17.5	11.27	93.64
YOK36	32.31	18.89	50.12	2.12	25.21	2.74	104.35	88.84	10.64	2.4	0
DYS39	64.31	39.14	24.27	56.69	32.35	45.07	44.81	153.21	68.21	11.86	366.78
DYS40	30.7	24.29	27.55	30.67	48.41	29.69	70.38	138.55	8.26	11.6	205.4
LRF41	27.78	40.14	8.64	15.7	16.78	20.81	58.38	121.83	77.27	19.38	155.95
LRF61	33.64	25.12	41.78	19.14	39.21	22.03	111.92	453.1	163.6	22.24	449.84

Table 4.
Consolidated Levene's
test results

Table 5 then provides the consolidated results of the Kruskal–Wallis H test. Finally, Table 6 combines all results to provide insight into which units vary the most.

In Table 4, grey cells indicate the variances resulted in failing to reject the null hypothesis (not a significant difference between the units). Green cells indicate a lower variance, and yellow cells indicate a higher variance that was causal in the rejection of the null hypothesis.

Within Table 5, the grey cells indicate the sample mean resulted in failing to reject the null hypothesis (not a significant difference between the units). Orange cells indicate that within that category (column), the squadron had a significantly higher mean, while yellow cells indicate a significantly smaller mean.

Finally, Table 6 consolidates all test results. Within the table, each cell contains a 0, 1 or 2. Zero indicates neither test resulted in a rejected null hypothesis, 1 indicates one of either test hypotheses was rejected and 2 indicates both tests resulted in a rejection. The far-right side Totals column is the sum for each squadron, and the lower Totals row is the sum for each flight hour category. Table 6 provides evidence there are differences between the units and are most significant between the overseas locations (Ramstein and Yokota) versus the stateside locations.

- (2) What is the correct distribution of flight hours a C-130J co-pilot should obtain to reduce variance and ensure quality outcomes?

Through the analysis of 90 flight records, averages were set for each category of flight hour. After an average was established as a percentage of total C-130J flight hours, the percentage was multiplied by 700 h. A total of 700 flight hours were chosen as this is the minimum amount of flight hours a co-pilot requires before upgrading to aircraft commander. After calculating the raw number of hours, rounding was applied based on standard deviations resulting in the recommended hours. Both the raw data hours and rounded hours are provided in Table 7. Consequently, this research offers a roadmap for future co-pilot training and sets new standards that should be followed to guide production and quality control efforts. Although we cannot say that reaching these minimum hours will result in the quality of upgrade per individual, it would create safe and war-ready pilots. We know that the standardization of hours results in consistent experience levels across the different flight hours, leading to organizational safety based on known proficiency. For example, we cannot

	Primary	Secondary	Other	Night	INS	NVG	Training	Tasked	Simulator	JA/ATT	CMB/Spt
RMS37	70.87	22.27	6.8	12.93	29.47	6.2	28.4	56.53	8.07	6.87	7.07
YOK36	58.8	20.2	21.13	12.13	26.73	5.8	55.93	36.53	7.07	0.4	0
DYS39	58.8	24	17.13	31.4	29.73	22.07	17.33	69.73	8.73	4	45.07
DYS40	54.13	30	16.13	27.67	25.87	18.6	21.67	65.13	8.6	4.8	31.6
LRF41	54.93	32	13.07	26.47	28.07	18.33	26.33	48.4	19.53	5.67	29.33
LRF61	52.73	31.53	15.93	24	25.27	14.2	28.27	49.67	14.8	7.33	29.13

Table 5. Consolidated Kruskal–Wallis H test results

	Primary	Secondary	Other	Night	INS	NVG	Training	Tasked	Simulator	JA/ATT	CMB/Spt	TOTALS
RMS37	1	1	2	2	0	2	0	1	0	0	2	11
YOK36	0	1	1	2	0	2	1	1	0	2	2	12
DYS39	0	1	0	1	0	0	1	1	0	0	0	4
DYS40	0	1	0	0	0	0	0	1	1	0	0	3
LRF41	0	1	1	0	0	0	0	1	2	0	0	5
LRF61	0	1	0	0	0	1	0	2	2	0	1	7
TOTALS	1	6	4	5	0	5	2	7	5	2	5	

Table 6. Consolidated research results

say how many nighttime hours a particular pilot needs to reach a level of quality (i.e. safety or mission ready), but we will know that minimizing the differences between pilots enables commanders confidence in known proficiency and standardization. We believe Table 7 will provide the benchmark for future training to ensure the safety of crew and passengers through higher quality outcomes.

- (3) What insights, if any, can the Six Sigma DMAIC methodology provide in C-130J co-pilot development?

This research recommends implementing a quality control program capable of monitoring variance in pilot development. Figure 5 presents the recommended DMAIC

Table 7.
Recommended flight hour distributions

	Primary	Secondary	Other	Night	INS	NVG	Training	Tasked	Simulator	JA/ATT	CMB/Spt
C-130J FP AVERAGE	58%	27%	15%	22%	28%	14%	30%	54%	11%	5%	24%
STANDARD DEVIATION	8%	7%	7%	9%	6%	8%	15%	17%	9%	4%	21%
Raw Hours (% x 700)	408.6533	186.7301	105.0833	157.1431	192.7251	99.25314	207.5889	379.5556	77.93333	33.91111	165.9
Rounded Recommended	400	200	100	150	175	100	250	350	75	25	100

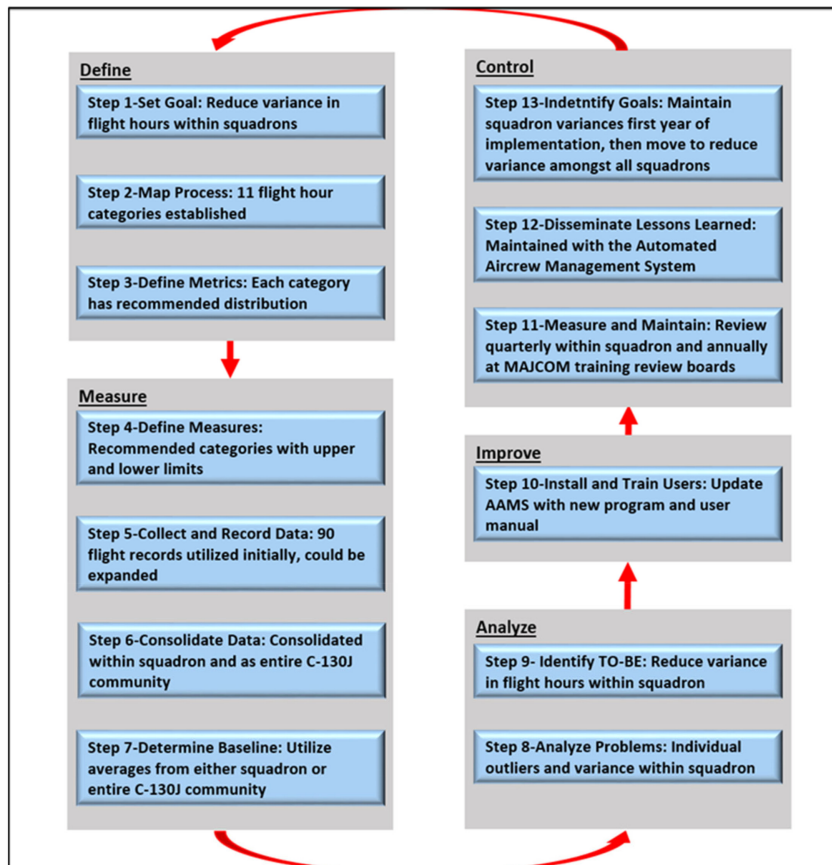


Figure 5.
DMAIC method applied to C-130J pilot production

method as it could be applied to the flight development process. This section discusses each step of DMAIC and how it could be used for C-130J co-pilot development.

Define

The results of this research project highlighted significant variances regarding C-130J co-pilot development. As discussed in the literature review, variance reduction improves processes. The results from this research indicate these variances occur both within the squadron and between the units, and both levels could benefit from variance reduction.

Measure

This research accomplished this step through the data collection and analysis of 90 co-pilots' aviation records. If this recommendation is accepted, further discussion will be required from the C-130J community to determine if the recommended "baselines" should be adjusted. The baselines could be adjusted by individual units based on the findings of this research. For example, overseas locations are unable to meet the same amount of NVG or nighttime training as stateside locations.

Analyze

Through subject matter expert inputs and post hoc analysis, reasons for the variances are provided. Understanding why the variance is occurring is the first step in deciding if a reduction in variance is possible or necessary. In some cases, the restrictions imposed by the host nation will not change. For this reason, individual units should have the capability to adjust their control limits for each category. Setting control limits will improve variance reduction and is the basis for the next step.

Improvement

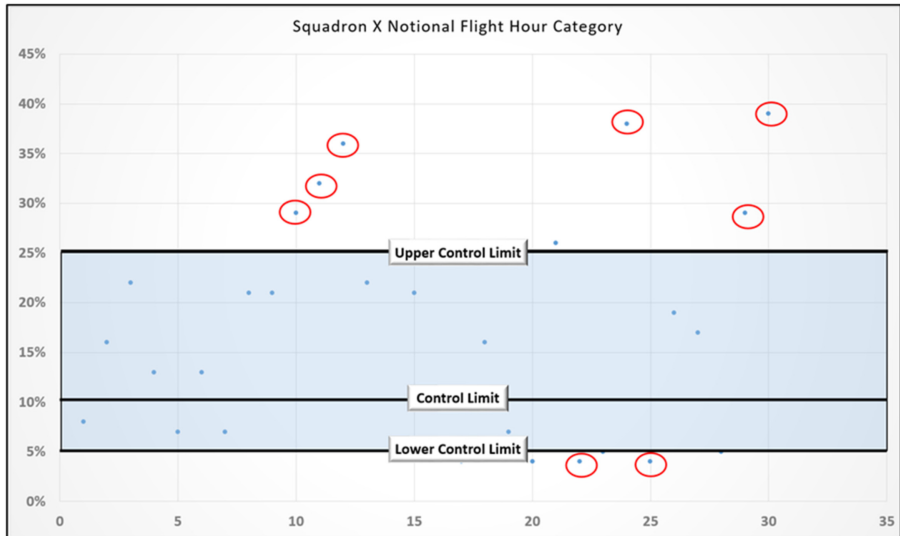
To reduce variance in co-pilot hours, it is recommended to set upper and lower control limits within each flight hour category by the individual units, and all units should know the differences. Then by applying these filters to their respective co-pilot population, individual outliers could be identified and then scheduled more effectively. The envisioned generated product would be similar to [Figure 6](#). In this example, the unit determined the control limit and then set upper and lower limits. All pilots that fall outside these limits are identified and scheduled more efficiently to reduce the variance within the unit.

Control

As the ultimate step of the DMAIC method, documentation would be stored within a database. The data would enable continued analysis from management. At the unit level, the main goal will be to reduce variance. At higher echelons of management, the variance between the units could be utilized to more effectively schedule exercises and seek assistance from headquarters management regarding the Flying Hour Program execution.

Through the DMAIC lens, we have garnered several important insights. First, we have accurately measured the current state of co-pilot production and found some areas for significant improvement across several categories of upgrade training. Next, DMAIC has empowered this research to recommend a future state and provide the appropriate metrics to evaluate its performance. To this end, co-pilot production as a service could benefit from employing DMAIC as it transitions to different states and procedures for developing high-performing pilots.

Figure 6.
Envisioned AAMS
Report on notional
scenario



Conclusions

This is the first known research to both identify and provide an analysis on the variance and mean flight hours in the C-130J pilot upgrade process. Both the mean and the variance have critical implications to the efficient and effective upgrade training as well as safety concerns for co-pilot upgrade efforts in the C-130J (Bahari, 2011; Gholizadeh and Esmaeili, 2020). This research clearly shows statistical differences in the quality of flight hours for pilots based on their respective geographic locations. For example, Ramstein provided more primary flight time, while Yokota co-pilots were receiving significantly more hours in other flight time. This suggests that a Ramstein co-pilot will be much more proficient and qualified than the Yokota pilot, even though they meet the requirement based on the total flight hours acquired.

Future research should analyze flight incidents or mishaps based upon the quality of hours the co-pilot received, as common sense would predict more mishaps occur based on relative substandard quality flight hours. Furthermore, identifying the variances leads to more efficient pilot scheduling, resulting in cost savings through more efficient training scheduling, which decreases consumable consumption, such as fuel. Moreover, we anticipate increased pilot production outcomes through the lens of variance control. As the variances are reduced, all pilots will receive similar critical flight skills, thereby increasing flight safety (Xu and Xu, 2021). For location-specific issues that impact the availability of training hours in a given category, identifying alternative modalities to overcome those issues, such as simulation, is an area for future research.

We anticipate higher quality outcomes translating into reduced aircraft accidents or near-miss incidents by providing a roadmap for future co-pilot training. Moreover, this research could be applied to any Air Force co-pilot, but it could also be used as a benchmark to any aviation branch, military or civilian where cumulative flight hours are used to mask deficiencies in upgrade training tasks. More importantly, it could be used as an example for any company or profession that employs hours as a means of upgrading employees. Our research should be a template to drive high-quality task-oriented behaviors where skillsets are valued. Consequently, our study illustrates that a focus on thresholds for quality in specific core tasks versus a threshold directed at the upgrade program is a higher quality standard for the community and its consumers.

References

- AMC/A3TA (2020), "Air Force manual 11-2C-130J, Volume 1", *Flying Operations. C-130J Aircrew Training*.
- AP/2 (2020), "DoD flight information publication (Enroute)", in *Supplement Europe-Africa-Middle East*, National Geospatial-Intelligence Agency, United States Government.
- AP/3 (2019), "DoD flight information publication", *Pacific-Australasia-Antarctica*, National Geospatial-Intelligence Agency, United States Government, Washington DC.
- Bahari, S.F. (2011), *An Investigation of Safety Training, Safety Climate and Safety Outcomes: A Longitudinal Study in a Malaysian Manufacturing Plant*, The University of Manchester (United Kingdom).
- Caulcutt, R. (2001), "Why is Six Sigma so successful?", *Journal of Applied Statistics*, Vol. 28 Nos 3-4, pp. 301-306.
- Chakrabarty, A. and Tan, K.C. (2007), "The current state of six sigma application in services", *Managing Service Quality: An International Journal*, Vol. 17 No. 2, pp. 194-208.
- Chakraborty, R.K., Biswas, T.K. and Ahmed, I. (2013), "Reducing process variability BY using DMAIC model: a case study IN Bangladesh", *International Journal for Quality Research*, Vol. 7 No. 1, pp. 127-140.
- Davis, J.C. (2017), *Mobility Air Force Aircrew Flight Training Requirements Validation through the Use of Line Oriented Safety Audit Data*, Air Force Institute of Technology, Wright Patterson AFB.
- Degirmenci, T. *et al.* (2013), "Potential of standardization and certification for successful lean implementations", *Journal of Enterprise Transformation*, Vol. 3 No. 3, pp. 211-232.
- De Mast, J. and Lokkerbol, J. (2012), "An analysis of the Six Sigma DMAIC method from the perspective of problem solving", *International Journal of Production Economics*, Vol. 139 No. 2, pp. 604-614.
- Department of Transportation (2012), "Pilot certification and qualification requirements for air carrier operations", *Federal Aviation Administration*, Federal Register, Vol. 77 Nos 40/Wednesday, February 29.
- Depperschmidt, C.L. (2013), "Public Law 111-216: effects of new legislation on collegiate aviation flight training programs", *The Collegiate Aviation Review International*, Vol. 31 No. 1.
- Electronic Code of Federal Regulations (2020), "Part 61, Subpart G – Airline Transport Pilots", *US Government Publishing Office*, available at: https://www.ecfr.gov/cgi-bin/text-idx?SID=762dc9e1dc3ebb7f4ea478978a57ae7c&mc=true&node=se14.2.61_1159&rgn=div8 (accessed 17 January 2020).
- Elmuti, D. (2002), "The perceived impact of supply chain management on organizational effectiveness", *Journal of Supply Chain Management*, Vol. 38 No. 2, pp. 49-57.
- Flynn, B.B. and Flynn, E.J. (2005), "Synergies between supply chain management and quality management: emerging implications", *International Journal of Production Research*, Vol. 43 No. 16, pp. 3421-3436.
- Flynn, B.B., Schroeder, R.G., Flynn, E.J., Sakakibara, S. and Bates, K.A. (1997), "World-class manufacturing project: overview and selected results", *International Journal of Operations and Production Management*.
- Gholizadeh, P. and Esmaeili, B. (2020), "Cost of occupational incidents for electrical contractors: comparison using robust-factorial analysis of variance", *Journal of Construction Engineering and Management*, Vol. 146 No. 7, p. 04020073.
- Glen, S. (2014), "Levene's test for Equality of variances. Statistics how to", March 4, available at: <https://www.statisticshowto.datasciencecentral.com/levene-test/> (accessed 11 March 2020).
- Klefsjö, B., Wiklund, H. and Edgeman, R.L. (2001), "Six Sigma seen as a methodology for total quality management", *Measuring Business Excellence*, Vol. 5 No. 1, pp. 31-35.

- Mackay, J.R. and Steiner, S.H. (1997), *Strategies for Variability Reduction*, Department of Statistics and Actuarial Sciences. University of Waterloo.
- Matson, E. and Prusak, L. (2003), "The performance variability dilemma", *MIT Sloan Management Review*, Vol. 45 No. 1, p. 39.
- NIST/SEMATECH (2020), "E-handbook of statistical methods", available at: <http://www.itl.nist.gov/div898/handbook> (accessed 10 March 2020).
- Nolan, T.W. and Provost, L.P. (1990), "Understanding variation", *Quality Progress*, Vol. 23 No. 5, pp. 70-78.
- Pan, Z., Ryu, H. and Baik, J. (2007), "A case study: CRM adoption success factor analysis and Six Sigma DMAIC application", *5th ACIS International Conference on Software Engineering Research, Management and Applications (SERA 2007)*, IEEE, pp. 828-838.
- Rooney, A.L. and Van Ostenberg, P.R. (1999), *Licensure, Accreditation, and Certification: Approaches to Health Services Quality*, Center for Human Services, Quality Assurance Project.
- Shortliffe, L.M.D. (2009), "Certification, recertification, and maintenance: continuing to learn", *Urologic Clinics of North America*, Vol. 36 No. 1, pp. 79-83.
- Smith, G.M., Herchko, D., Bjerke, E., Niemczyk, M., Nullmeyer, R., Paasch, J. and NewMyer, D.A. (2013), "The 2012 pilot source study (phase III): response to the pilot certification and qualification requirements for air carrier operations", *Journal of Aviation Technology and Engineering*, Vol. 2, p. 2.
- Tang, L.C., Goh, T.N., Lam, S.W. and Zhang, C.W. (2007), "Fortification of six sigma: expanding the DMAIC toolset", *Quality and Reliability Engineering International*, Vol. 23 No. 1, pp. 3-18.
- Tonini, A.C. and de Mesquita Spinola, M. and Fernando Jose, B.L. (2006), "Six Sigma and software development process: DMAIC improvements", *2006 Technology Management for the Global Future-PICMET 2006 Conference*, Vol. 6, IEEE, pp. 2815-2823.
- US Congress Office of Technology Assessment (1988), *Safe Skies for Tomorrow: Aviation Safety in a Competitive Environment. OTA-SET-381*, US Government Printing Office, Washington, DC.
- Vardhan, H.G. (2021), "Customers' service quality Expectations from quick service restaurants", *The Retail and Marketing Review*, Vol. 17 No. 1, pp. 79-88.
- Werfelman, L. (2010), "Counting the hours", *Flight safety foundation*, available at: <https://flightsafety.org/asw-article/counting-the-hours/> (accessed 10 January 2020).
- Xu, Q. and Xu, K. (2021), "Analysis of the characteristics of fatal accidents in the construction industry in China based on statistical data", *International Journal of Environmental Research and Public Health*, Vol. 18 No. 4, p. 2162.

Further reading

- A3TF, AF/ (2019), *Air Force Instruction 11-412, Flying Operations, Aircrew Management*.
- AMC/A3VX (2019), "Air Force manual 11-2C-130J, Volume 3", *Flying Operations. C-130J Operations Procedures*.
- Automated Aircrew Management System (AAMS) (2019), *User's Manual*, Version 3.17.
- Conn, R. (2002), "Developing software engineers at the C-130J software factory", *IEEE Software*, Vol. 19 No. 5, pp. 25-29.
- Gastwirth, J.L., Gel, Y.R. and Miao, W. (2009), "The impact of Levene's test of equality of variances on statistical theory and practice", *Statistical Science*, Vol. 24 No. 3, pp. 343-360.
- Hopp, W.J. and Spearman, M.L. (2007), *Factory Physics: Foundations of Manufacturing Management*, 3rd ed., McGraw-Hill, Chicago.
- HQ AMC/A3D (2018), "Air mobility Command instruction 10-2101", *Operations, Joint Airborne/Air Transportability Training*.

- HQ USAF/A30-T (2010), "Air Force Instruction 11-401", *Aviation Management* (Incorporates Air Force Guidance Memorandum AFGM2019-01. 31 January 2019).
- Lynch, D.P., Bertolino, S. and Cloutier, E. (2003), "How to scope DMAIC projects", *Quality Progress*, Vol. 36 No. 1, pp. 37-41.
- Morrissey, T. (2004), "Analysis of pilot performance using precision visual flight rules", Master's Thesis, University of Tennessee.
- Morton, P., Adams, B., Byrnes, K., Morrison, G., Hibler, T., Greubel, D., Haugaard, L., Buyer, J., Dee, M., Winter, J. and Panhans, J. (2016), "Panel 2: ATP/CTP experience Report and new ideas in flight education", Vol. 11, *National Training Aircraft Symposium (NTAS)*.
- Panagopoulos, I., Atkin, C.J. and Ivan, S. (2016), "Lean Six-Sigma in Aviation Safety: an implementation guide for measuring aviation system's safety performance", *Journal of Safety Studies*, Vol. 2 No. 2, pp. 1-12, ISSN 2377-3219 2016.
- Poiger, M. (2010), *Improving Performance of Supply Chain Process by Reducing Variability*, Diss. WU Vienna University of Economics and Business.
- Sarkar, D. and Zangwill, W.I. (1991), "Variance effects in cyclic production systems", *Management Science*, Vol. 37 No. 4, pp. 444-453.
- Thompson, N. (2018), *Avoiding a Pilot Retention Death Spiral: The Pilot Shortage and DOD's Challenge to Maintain an Effective Fighting Force*, Joint Forces Staff College/NDU NORFOLK.

Appendix

Categories 4–11 with associate tables, figures and narrative

Category 4: Nighttime – flying that occurs between sunset and sunrise

Similarly, [Table A5](#) displays the basic statistical data derived from the 90 co-pilots, while [Figure A3](#) visually portrays these data. The Kruskal–Wallis H test resulted in a p -value of 0.001, indicating that there is statistical difference with respect to the mean nighttime. The Levene test resulted in a p -value of 0.000, which rejects the null hypothesis, indicating there is a statistical difference between the variance of the six squadrons' nighttime. Ramstein 37AS and Yokota 36AS differed from the other four units in their respective low mean. Furthermore, Dyess 39AS differed from the other five units due to their high mean. Regarding variance, Ramstein 37AS and Yokota 36AS are significantly different from the other four units. Subject matter expert insight was collected on Ramstein 37AS and Yokota 36AS low mean and variance.

The subject matter expert pointed out that these low means are caused by the host nation restrictions related to noise abatement procedures directed in their respective aviation regulations (Personal Correspondence, March 2020). In both Japan and Germany, the units are restricted in the amount of nighttime they can legally fly. In these locations, training flights are only authorized between the hours of 0600–2200 local time, directed in each of their respective DoD Flight Information Publication ([AP/3, 2019](#), pp. 3–77) ([AP/2, Feb 2020:B-422](#)). These host nation restrictions impact the amount of

Sample						
Unit Tested	Size from Unit	Mean of Sample	Median of Sample	Variance of Sample	Standard Deviation	
RMS37	15	12.93	13	7.35	2.71	
YOK36	15	12.13	12	2.12	1.46	
DYS39	15	31.4	31	56.69	7.53	
DYS40	15	27.67	28	30.67	5.54	
LRF41	15	26.47	27	15.7	3.96	
LRF61	15	24	25	19.14	4.38	

Table A1.
Night flight time basic statistics

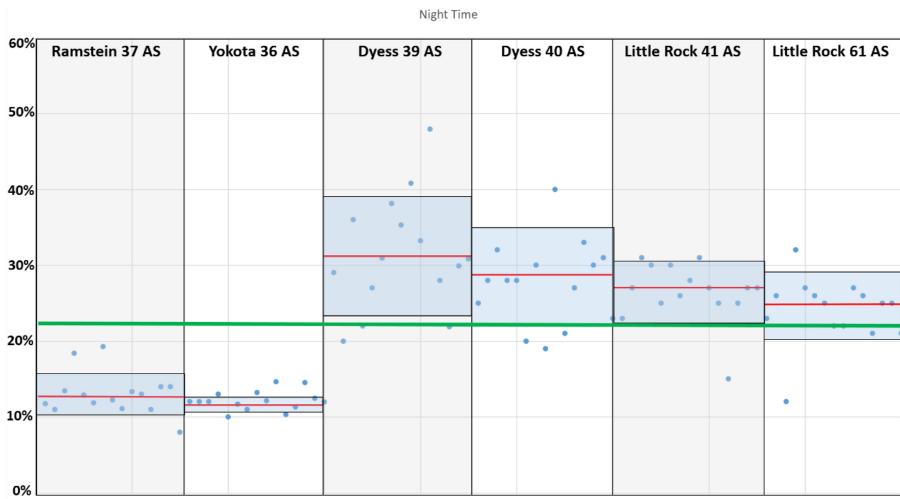


Figure A1.
Night flight time as a percentage of total C-130J hours

“night” and “NVG” flight time that can be accomplished. At stateside locations, there is no limit to the time at which the training must end. Hence, there is more opportunity to log night and NVG flight time. Consequently, co-pilots trained in Ramstein and Yokota are currently considered equals for co-pilots trained at Dyess even though they clearly have less training at night by comparison.

Category 5: Instrument time – flying that occurs with the use of primary vision (instrument only)

As before, [Table A6](#) displays the basic statistical data derived from the 90 co-pilots, while [Figure A4](#) visually portrays these data. The statistical testing confirmed there is not a significant difference in variance or mean between any of the units. The Kruskal–Wallis H test resulted in a *p*-value of 0.249, indicating that there is no statistical difference with respect to the mean instrument time. The Levene test resulted in a *p*-value of 0.146, which fails to reject the null hypothesis, indicating there is not statistical difference between the variance of the six squadrons’ instrument time. Lastly, instrument flight time is the only category that did not indicate a significant statistical difference between any of the units.

Unit Tested	Sample		Mean of Sample	Median of Sample	Variance of Sample	Standard Deviation
	Size	Unit from				
RMS37	15		29.47	29	59.41	7.71
YOK36	15		26.73	26	25.21	5.02
DYS39	15		29.73	30	32.35	5.69
DYS40	15		25.87	25	48.41	6.96
LRF41	15		28.07	29	16.78	4.1
LRF61	15		25.27	24	39.21	6.26

Table A2.
Instrument flight time basic statistics

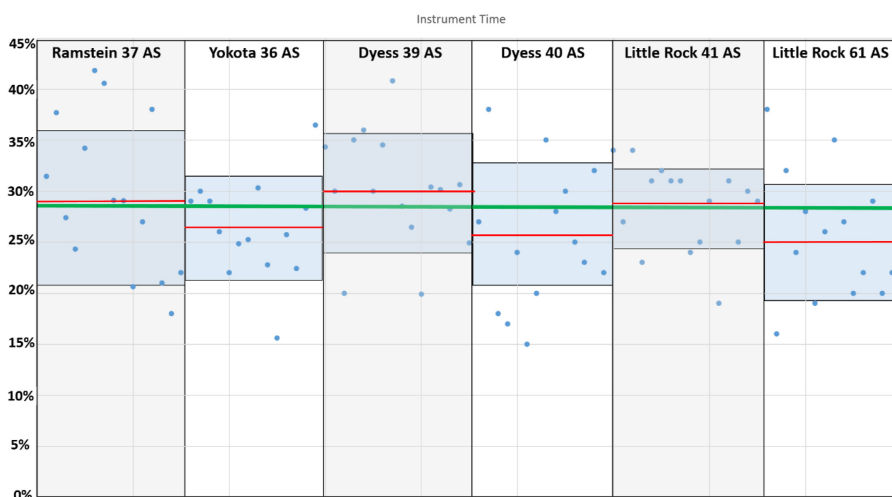


Figure A2.
Instrument flight time as a percentage of total C-130J hours

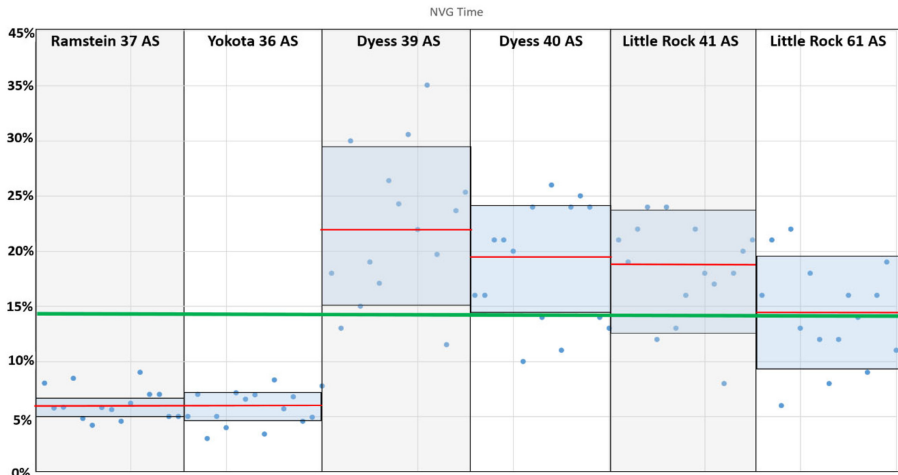
Category 6: Night vision goggle (NVG) time – flying that occurs with the primary employment of a night vision enhancement device

Similarly, [Table A7](#) displays the basic statistical data derived from the 90 co-pilots, while [Figure A5](#) visually portrays these data. The Kruskal–Wallis H test resulted in a *p*-value of 0.000, indicating that there is statistical difference with respect to the mean NVG time. The Levene test resulted in a *p*-value of 0.000, which rejects the null hypothesis indicating there is a statistical difference between the variance of the six squadrons’ NVG time. More specifically, the statistical test results show that Ramstein 37AS and Yokota 36AS are statistically different in the equality of variances and their respective mean, compared to the other four units. Like the nighttime flight hours, the subject matter expert suggested NVG time is limited in Japan and Germany due to the host nation restrictions and noise abatement procedures (Personal Correspondence, March 2020).

Unit Tested	Sample		Mean of Sample	Median of Sample	Variance of Sample	Standard Deviation
	Size from Unit					
RMS37	15		6.2	6	1.89	1.37
YOK36	15		5.8	6	2.74	1.66
DYS39	15		22.07	22	45.07	6.71
DYS40	15		18.6	20	29.69	5.45
LRF41	15		18.33	19	20.81	4.56
LRF61	15		14.2	14	22.03	4.69

Table A3.
NVG flight time basic statistics

Figure A3.
NVG flight time as a percentage of total C-130J hours



Category 7: Training flight time – flying time that is categorized as nonoperational

As before, [Table A8](#) displays the basic statistical data derived from the 90 co-pilots, while [Figure A6](#) visually portrays these data. The Kruskal–Wallis H test resulted in a p -value of 0.000 indicating that there is statistical difference with respect to the mean training flight time. The Levene’s test resulted in a p -value of 0.372 and failed to reject the null hypothesis, thus indicating that the variances are not significantly different between the squadrons in training flight time. More specifically, post hoc analysis revealed that Yokota 36AS and Dyess 39AS are different from the other four units based on their respective means. The data presented indicate that Yokota 36AS is different due to a high mean, and Dyess 39AS differs due to a low mean. Both cases can be explained due to their respective tasked time ratios. A subject matter expert from Yokota suggested during the timeframe of data collection; the squadron was receiving fewer than average tasked missions from its tasking authority. However, no specific reason could be provided for this lower tasking rate.

Table A4.
Training flight time basic statistics

Unit Tested	Sample		Median of Sample	Variance of Sample	Standard Deviation
	Size from Unit	Mean of Sample			
RMS37	15	28.4	30	54.11	7.36
YOK36	15	55.93	57	104.35	10.22
DYS39	15	17.33	17	44.81	6.69
DYS40	15	21.67	21	70.38	8.39
LRF41	15	26.33	25	58.38	7.64
LRF61	15	28.27	26	111.92	10.58

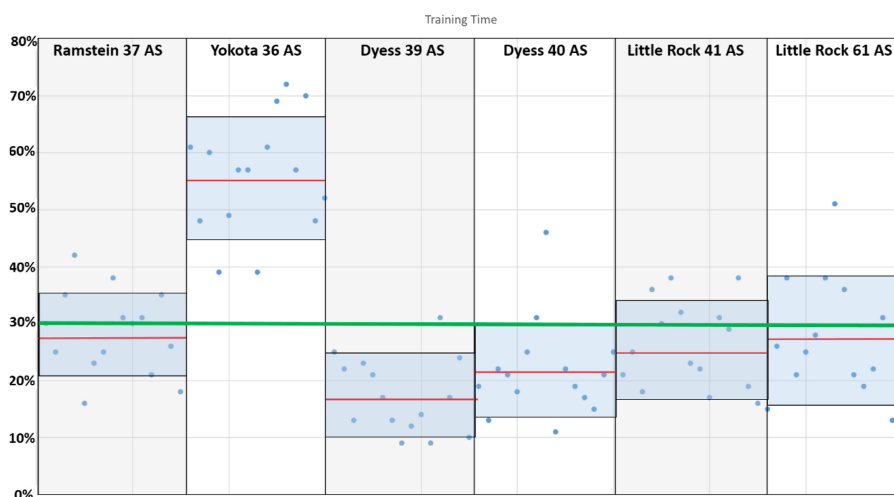


Figure A4. Training flight time as a percentage of total C-130J hours

Category 8: Tasked flight time – flying time that is categorized as operationally tasked from higher headquarters

Table A9 displays the basic statistical data derived from the 90 co-pilots, while Figure A7 visually portrays this data. The Kruskal–Wallis H test resulted in a *p*-value of 0.000 indicating that there is statistical difference with respect to the mean tasked flight time. The Levene’s test resulted in a *p*-value of 0.003 and rejected the null hypothesis, thus indicating that the variances are significantly different between the squadrons in tasked flight time. More specifically, Little Rock 61AS is significantly different based on variance than the other five units. Based on post hoc analysis of the means, the units are divided in half with three units having a significantly higher mean of tasked time compared to the other three units. Tables A1–A8, with those units having a high mean boxed with red and those with a low mean boxed in blue. Subject matter experts were not available to comment on this variance.

Unit Tested	Sample Size from Unit	Mean of Sample	Median of Sample	Variance of Sample	Standard Deviation
RMS37	15	56.53	56	60.84	7.8
YOK36	15	36.53	37	88.84	9.43
DYS39	15	69.73	70	153.21	12.38
DYS40	15	65.13	67	138.55	11.77
LRF41	15	48.4	51	121.83	11.04
LRF61	15	49.67	57	453.1	21.29

Table A5. Tasked flight time basic statistics

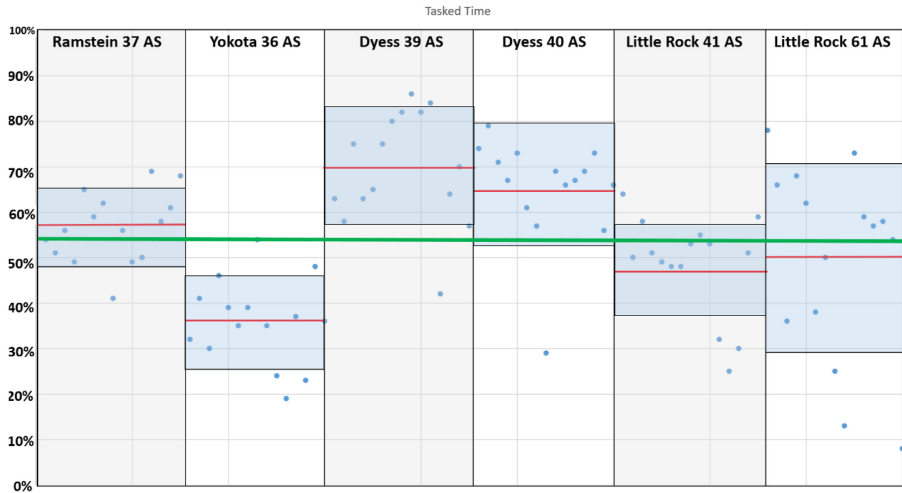


Figure A5.
Tasked flight time as a percentage of total C-130J hours

Category 9: Simulator time – flying time that is logged in a virtual environment

As with above, [Table A10](#) displays the basic statistical data derived from the 90 co-pilots, while [Figure A8](#) visually portrays these data. The Kruskal–Wallis H test resulted in a p -value of 0.001, indicating that there is statistical difference with respect to the mean simulator time. The Levene’s test resulted in a p -value of 0.000 and rejected the null hypothesis, thus indicating that the variances are significantly different between the squadrons in simulator time. More specifically, the results indicate both units at Little Rock, 61AS and 41AS, are significantly different in both equality of variance and means from the other locations. Dyess 40AS differs significantly from the other five units based on variance. A subject matter expert at Little Rock explained that at Little Rock, there are more simulators than other locations. Little Rock AFB is host to the C-130J schoolhouse, where all pilots and loadmasters begin their C-130J training.

For this reason, Little Rock has four simulators compared to one simulator at the three other locations (Personal Correspondence, March 2020). With more simulators come more opportunities for the units to complete simulator training explaining the high variance and mean for the two Little Rock units. A subject matter expert from Dyess confirmed this analysis by explaining Dyess has two units using one simulator. Additionally, this individual explained that occasionally their single simulator is utilized by crews from Germany or Yokota due to availability at their respective locations. These explanations confirm the statistical results; Little Rock units are obtaining more simulator hours, which in turn increases the opportunity for variance (Personal Correspondence, February 2020).

Unit Tested	Sample Size from Unit	Mean of Sample	Median of Sample	Variance of Sample	Standard Deviation
RMS37	15	8.07	7	17.5	4.18
YOK36	15	7.07	7	10.64	3.26
DYS39	15	8.73	7	68.21	8.26
DYS40	15	8.6	8	8.26	2.87
LRF41	15	19.53	21	79.27	8.9
LRF61	15	14.8	7	163.6	12.79

Table A6.
Simulator time basic statistics

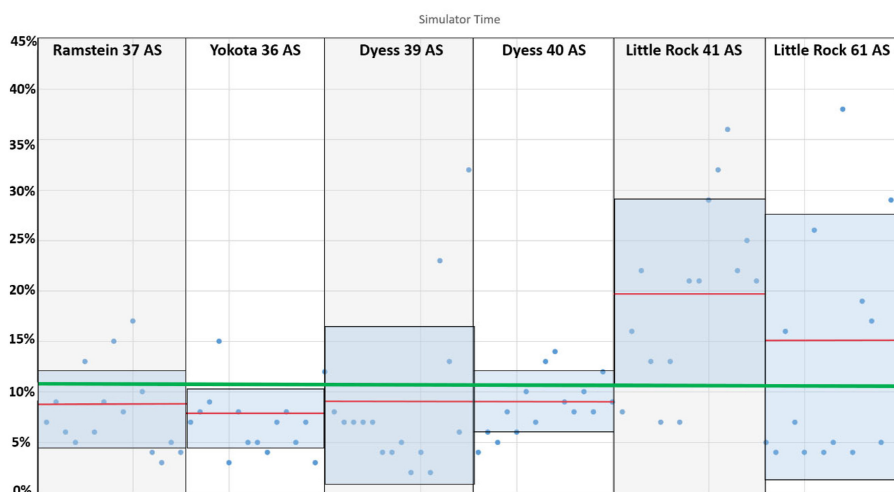


Figure A6. Simulator time as a percentage of total C-130J hours

Category 10: Joint airborne or air transportability training time – flying time that involves jettisoning cargo or passengers via a parachute

Like above, [Table A11](#) displays the basic statistical data derived from the 90 co-pilots, while [Figure A9](#) visually portrays these data. The Kruskal–Wallis H test resulted in a *p*-value of 0.000, indicating that there is statistical difference with respect to the mean joint airborne or air transportability training time. The Levene’s test resulted in a *p*-value of 0.003 and rejected the null hypothesis, thus indicating that the variances are significantly different between the squadrons in joint airborne or air transportability training time. More specifically, Yokota 36AS is significantly different based on variance and mean from the other five units. Of the 15 samples from Yokota 36AS, only one had any joint airborne or air transportability training time, which results in a significantly low mean and variance.

Unit Tested	Sample		Mean of Sample	Median of Sample	Variance of Sample	Standard Deviation
	Size from Unit					
RMS37	15		6.87	6	11.27	3.36
YOK36	15		0.4	0	2.4	1.55
DYS39	15		4	4	11.86	3.44
DYS40	15		4.8	4	11.6	3.41
LRF41	15		5.67	5	19.38	4.4
LRF61	15		7.33	6	22.24	4.72

Table A7. Joint airborne or air transportability flight time basic statistics

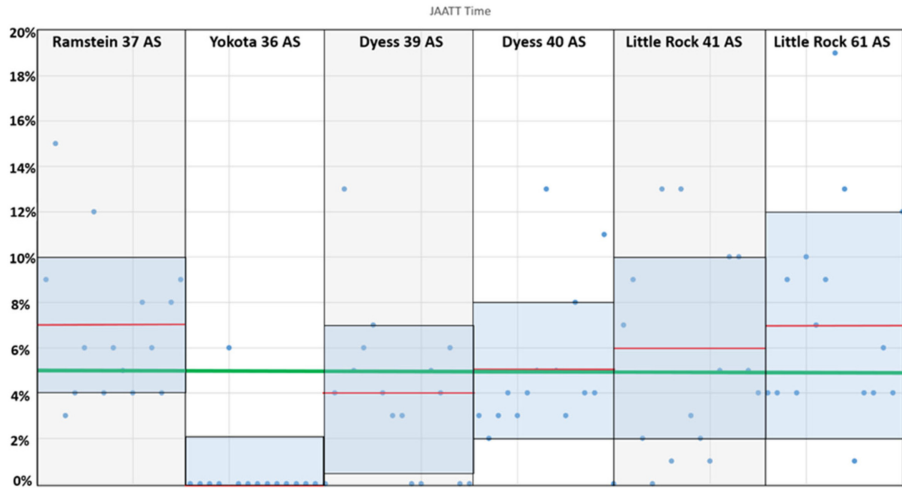


Figure A7.
Joint airborne or air transportability flight time as a percentage of total C-130J hours

Category 11: Combat and combat support flight time – flying time that encapsulates a combat related scenario (i.e. in a combat zone)

In accordance with the above, [Table A12](#) displays the basic statistical data derived from the 90 co-pilots, while [Figure A10](#) visually portrays these data. The Kruskal–Wallis H test resulted in a *p*-value of 0.000, indicating that there is statistical difference with respect to the mean combat and combat support flight time. The Levene’s test resulted in a *p*-value of 0.000 and rejected the null hypothesis, thus indicating that the variances are significantly different between the squadrons in combat and combat support flight time. More specifically, Ramstein 37AS and Yokota 36AS are significantly different from the other four units based on variance and their respective mean. Little Rock 61AS is also significantly different than

Unit Tested	Sample Size from Unit	Mean of Sample	Median of Sample	Variance of Sample	Standard Deviation
RMS37	15	7.07	4	93.64	9.68
YOK36	15	0	0	0	0
DYS39	15	45.07	50	366.78	19.15
DYS40	15	31.6	38	205.4	14.33
LRF41	15	29.33	32	155.95	12.49
LRF61	15	29.13	38	449.84	21.21

Table A8.
Combat/comb support flight time basic statistics

the five other units based on a high variance. Yokota 36AS had zero combat or combat support flight time within the sample of 15, while Ramstein 37AS recorded the next lowest mean of 7.07. A subject matter expert explained Ramstein 37AS and Yokota 36AS are not in a theater of operations that are currently supporting operations that would result in significant combat or combat support flight time. Little Rock 61AS's high variance can be explained by examining the samples collected. Of the 15 samples, four had zero combat time while the remaining 11 did and some with considerable amounts, which is the cause for the significant level of variance.

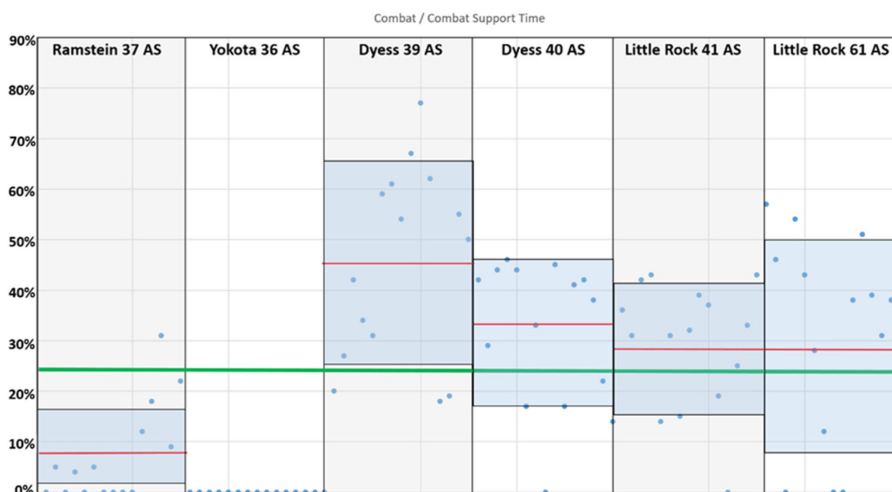


Figure A8. Combat/ combat support flight time as a percentage of total C-130J hours

Corresponding author

Jason Anderson can be contacted at: jason.anderson@afit.edu

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgrouppublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com