Air Force Institute of Technology [AFIT Scholar](https://scholar.afit.edu/)

[Theses and Dissertations](https://scholar.afit.edu/etd) **Student Graduate Works** Student Graduate Works

3-22-2007

Feasibility Study of Variance Reduction in the Logistics Composite Model

George P. Cole III

Follow this and additional works at: [https://scholar.afit.edu/etd](https://scholar.afit.edu/etd?utm_source=scholar.afit.edu%2Fetd%2F3068&utm_medium=PDF&utm_campaign=PDFCoverPages)

Part of the [Other Operations Research, Systems Engineering and Industrial Engineering Commons](http://network.bepress.com/hgg/discipline/310?utm_source=scholar.afit.edu%2Fetd%2F3068&utm_medium=PDF&utm_campaign=PDFCoverPages)

Recommended Citation

Cole, George P. III, "Feasibility Study of Variance Reduction in the Logistics Composite Model" (2007). Theses and Dissertations. 3068. [https://scholar.afit.edu/etd/3068](https://scholar.afit.edu/etd/3068?utm_source=scholar.afit.edu%2Fetd%2F3068&utm_medium=PDF&utm_campaign=PDFCoverPages)

This Thesis is brought to you for free and open access by the Student Graduate Works at AFIT Scholar. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of AFIT Scholar. For more information, please contact richard.mansfield@afit.edu.

FEASIBILITY STUDY OF VARIANCE REDUCTION

IN THE LOGISTICS COMPOSITE MODEL

THESIS

George P. Cole, III, Captain, USAF

AFIT/GLM/ENS/07-04

DEPARTMENT OF THE AIR FORCE AIR UNIVERSITY *AIR FORCE INSTITUTE OF TECHNOLOGY*

Wright-Patterson Air Force Base, Ohio

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government.

[FEASIBILITY STUDY OF VARIANCE REDUCTION](#page-1-0)

[IN THE](#page-1-0) COMPOSITE MODEL

THESIS

Presented to the Faculty

Department of Operational Sciences

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Logistics Management

George P. Cole, III, MS

Captain, USAF

March 2007

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

[FEASIBILITY STUDY OF VARIANCE REDUCTION](#page-1-0)

[IN THE](#page-1-0) COMPOSITE MODEL

George P. Cole, III, MS Captain, USAF

Approved:

Alan W. Johnson (Advisor) date

 \mathcal{L}_max

 \mathcal{L}_max John O. Miller (Reader) date

Abstract

 The Logistics Composite Model (LCOM) is a stochastic, discrete-event simulation that relies on probabilities and random number generators to model scenarios in a maintenance unit and estimate optimal manpower levels through an iterative process. Models such as LCOM involving pseudo-random numbers inevitably have a variance associated with the output of the model for each run, and the output is actually a range of estimates. The reduction of the variance in the results of the model can be costly in the form of time for multiple replications. The alternative is a range of estimates that is too wide to realistically apply to real-world maintenance units.

This research explores the application of three different methods for reducing the variance of the model's output. The methods include Common Random Numbers (CRN), Control Variates, and Antithetic Variates. The differences in the 95% confidence intervals were compared between the variance reduction techniques and the original model to determine the degree of variance reduction. The result is a successful variance reduction in the primary output statistics of interest using the application of the Control Variates technique, as well as a methodology for the implementation of Control Variates in LCOM.

iv

For My Wife and Daughter

Acknowledgments

I owe a huge debt of gratitude and thanks for support of my wife and daughter. My wife has sustained me throughout this effort with unwavering love and encouraged me every step of the way. She is truly my best friend in every sense. Our brand new daughter has been the light of my life these past few months, making all the late nights and "speed bumps" worth it with a loving smile.

I also owe many thanks to the rest of my family, as they have all been there for me with perfectly-timed words of encouragement and well-needed humor. I thank God for blessing me with such a wonderful family, and I love you all.

I owe a big thank you to the members of the ASC LCOM office for their time and assistance, especially Frank Erdman. Frank did not hesitate to give up a great deal of time to answer questions and teach me about the inner workings of LCOM. All of the members of the LCOM office were extremely helpful; this thesis effort would not have been possible without their support.

I must thank my thesis advisor, Dr Alan Johnson, and my reader, Dr J.O. Miller, for their guidance and patience along this journey. Their expertise, knowledge, and experience were instrumental to the completion of this thesis.

Finally, I must acknowledge my classmates here at AFIT. I could not have asked for a finer group of men and women with whom to share this experience, and I look forward to continuing these friendships for years and years to come.

George P. Cole, III

vi

Table of Contents

List of Figures

List of Tables

FEASIBILITY STUDY OF VARIANCE REDUCTION IN THE LOGISTICS COMPOSITE MODEL

1. Introduction

1.1. Background

The Logistics Composite Model (LCOM) is one of the Air Force's primary tools for determining optimal logistics and maintenance manpower levels. Additionally, it can be used to model other logistics resources such as equipment and facilities. The LCOM is a stochastic, discrete-event simulation that relies on probabilities and random number generators to model scenarios in a maintenance unit by manipulating certain variables. Manpower levels are attained through an iterative process in which the variables consisting of supply, facilities, and equipment are set based on command standards. Manpower levels are adjusted after each run until a desired Sortie Generation Rate (SGR) is attained (Boyle 1990).

 In a model such as LCOM, many real-life characteristics exhibit random behavior. As Law and Kelton (2004) state, "A simulation of any system or process in which there are inherently random components requires a method of generating or obtaining numbers that are *random,* in some sense" (Law and Kelton 2004). The random number generators aid the customer in simulating the randomness of the system by producing a stream of continuous, uniformly distributed numbers between 0 and 1. The intent of the random number generator is to produce these numbers independently. However, the computer is actually using a recursive algorithm that produces numbers that

seem independent, but instead follow a pattern that can be repeated over and over, called a stream (Kelton et al. 2004). These types of random number generators are called pseudo-random generators.

 Models involving pseudo-random numbers inevitably have a variance associated with the output of the model for each run, and the output is actually a range of estimates. The reduction of the variance in the results of the model can be costly in the form of time for multiple replications or producing a range of estimates that is too wide to realistically analyze. Since simply increasing the number of replications is not always realistic for reducing variance, this paper proposes the application of other methods for reducing the variance of the model's output. The methods include Common Random Numbers (CRN), Control Variates (CV), and Antithetic Variates (AV). When applied to a model such as LCOM, these variance reduction techniques may significantly reduce the variance without increasing the number of replications (Law and Kelton 2000).

1.2. Problem Statement

Due to the complexity of the LCOM's main model in simulating a real-life maintenance unit, the model contains many instances involving pseudo-random number generation. The large amount of randomness in the model causes results that display a significant amount of variance. As with any other model, large variances in the output are unfavorable. The LCOM office in the Aeronautical Systems Center's (ASC) Systems Supportability Analysis Branch, Wright-Patterson AFB, Ohio, is interested in finding ways to reduce the variance in the model's output.

1.3. Research Objective

The objective of this research is to replicate prior efforts by Bednar (2005) by studying the impact of various variance reduction techniques on LCOM. The research will apply these techniques to the model and analyze the results, determine whether or not each technique is effective and identify the most effective method.

1.4. Research Focus

This research first examines the capabilities and effectiveness of the model's random number generators. We will then use this information to investigate the application of several variance reduction techniques. In particular, three classic techniques – common random numbers, antithetic variates, and control variates – are applied to the LCOM model with the intention of reducing the variance of the model's output.

1.5. Methodology

Before we can apply these techniques, we must examine the random number generator used in the LCOM and determine if the generator is suitably robust for the application of the variance reduction methods. If not, the random number generator must be replaced with a more robust generator capable of facilitating synchronization and numerous replications.

The common random numbers approach involves multiple scenarios in the same model. Using this approach, the individual sources of randomness, or random variates, are synchronized using the same random number stream across the two scenarios. Then, configurations in two different scenarios will use the same random numbers so that the

different scenarios can experience similar experimental conditions (Law and Kelton 2000).

 The second technique, control variates, involves identifying potential control variates within the model that can be used to reduce the variance in the output. This method requires the identification of a particular random variable or variables with known expected values that are thought to correlate to an output variable, either positively or negatively. Then, using these potential control variates and the estimated correlation, the expected value of the output variable is adjusted up or down based on the differences between the observed values of the control variates and their known expected values (Kelton et al. 2004).

 The last technique, antithetic variates, "attempts to induce negative correlation between the results of one replication and another, and uses this correlation to reduce variance" (Kelton et al. 2004). This involves a second replication that replaces each random number U_i with the random number $1-U_i$. For example, where U_i is used for a particular purpose, $1-U_i$ is used in the second replication for the same purpose. The pairs are averaged, possibly replicating this for several pairs (Law and Kelton 2000).

 The approach for the comparison of these three methods involves acquiring the source code for the LCOM model and manipulating the calls, or random number generators, to perform each particular method and comparing the output variances of each method as well as the output variance of the original model.

1.6. Assumptions/Limitations

This research assumes the LCOM's source code can be manipulated and recompiled to allow the application of the three variance reduction techniques – common random numbers, antithetic variates, and control variates.

1.7. Implications

 This research has several implications for positive impact. First, the variance reduction techniques could potentially benefit the LCOM users by requiring fewer replications and shorter process time to achieve a given confidence in a model's output. This could allow the users to save processing time and test additional scenarios instead of performing additional replications. Second, the ASC LCOM office has developed an optimizer, called an auto-constraining wrapper, which automatically performs the constraining process of defining resource levels (Boughton 2006). These variance reduction techniques could potentially improve the performance of the optimizer and enhance the overall effectiveness of the LCOM. In the long term, this could lead to more precise manpower levels to meet mission requirements and more accurate sortie generation rate forecasts.

1.8. Overview

 There are five chapters in this research. Chapter 1, the Introduction, contains background information, problem statement, research objective, a synopsis of the methodology, and the assumptions of the research. Chapter 2, the Literature Review, contains descriptions of LCOM, other LCOM research, random number generators, common random numbers, antithetic variates, and control variates. Chapter 3, the Methodology, discusses how the variance reduction techniques are applied to LCOM.

Chapter 4, Results and Analysis, presents the output of the results from the application of the three techniques. These results are analyzed to determine the effectiveness of each technique. Chapter 5, Conclusions, provides a discussion of the analysis and results in the previous chapter as well as recommendations for future research in the application of variance reduction techniques and LCOM.

2. Literature Review

2.1. LCOM

 The U.S. Air Force's Logistics Composite Model has existed since the late 1960s, created through a combined effort by the Rand Corporation and the Air Force Logistics Command to "relate base-level logistics resources with each other and with sortie generating capability" (Boyle 1990). While the model is capable of studying the interactions between several variables, it has evolved to be known as one of the Air Force's primary tools for establishing manpower levels in operational maintenance units and exists as part of the Air Force's Standard Analysis Toolkit (AFSAT) (Dierker 2006).

 Two separate versions of the LCOM exist in the Air Force today, one at the Air Force Manpower Agency (AFMA), Randolph AFB, Texas, and the other at the Aeronautical Systems Center's (ASC) Systems Supportability Analysis Branch, Wright-Patterson AFB, Ohio (Pettingill 2003). These two separate models essentially perform the same function, with some minor differences in the user interface (Dierker 2006). The AFMA version has four primary users: Air Combat Command, Air Mobility Command, Air Force Special Operations Command, and Air Education and Training Command. The ASC is the only primary user for the ASC version (Dawson 2006). The AFMA version is used by the MAJCOMs to derive 65-70% of their maintenance manpower requirements (Dawson 2006). The rest comes from Air Force Instructions and other guidance. The ASC version is used "to analyze manpower requirements for acquired weapon systems (as well as evaluate manpower requirement changes resulting from modifications to current weapon systems)" (Dawson 2006).

 The LCOM model consists of multiple submodels, including an input model, a main model, and several post processors. The input model analyzes input data from the user and makes assumptions and corrections when necessary so that the data can be used by the main model and post processors. This data typically includes maintenance data from the Air Force's Maintenance Data Collection systems, essential tasks needed to be performed to service each aircraft, mission requirements and flying times. The main model is the heart of the simulation, and the primary source of data for our research. It uses maintenance data and sortie data together with the process logic shown in Figure 2.1.

Figure 2.1: LCOM Simulation Logic (ASC/ENM 2004)

Various post processors show simulation results as a function of time, such as manpower demands, resource and facility usage, parts availability, and depot workload (ASC/ENM 2004).

 The LCOM is subject to four forms of variance: stochastic model variance, interviewee variance, analyst variance, and MAJCOM procedure variance (Dawson 2006). Stochastic model variance occurs because the random number generators generate failures and repairs randomly, resulting in a random output. The random output is the type of variance this research examines. The other forms of variance are associated with data collection, evaluation and interpretation of results by analysts, and differing processes and philosophies across MAJCOMs (Dawson 2006).

2.2. Previous LCOM Studies

 Because of its longevity, complexity, and broad application, the LCOM has been the subject of many studies. Most studies have been focused on either the applicability of LCOM to various levels and resources of the Air Force, or the comparing LCOM to other similar models. Below is a brief summary of the most notable research efforts involving the inner workings of LCOM.

 In 2006, the Air Force Logistics Management Agency performed a study that examined LCOM process reengineering. The study examined the LCOM development process and the steps required for developing and conducting an LCOM study for a particular MAJCOM. The report recommended seven potential changes aimed at reducing the overall development time (Dawson 2006).

Captain Kirk B. Pettingill performed a study in 2003 on the LCOM's ability to determine a maintenance unit's current capacity, rather than just a front-end tool for setting initial manpower levels. He collected actual data from three flying units and compared the data to LCOM outputs. He concluded that the LCOM is an effective tool for determining a maintenance unit's current capacity to produce sorties, but with limitations (Pettingill 2003).

 In 1996, Captain Todd Carrico and Patrick K. Clark of the Human Resources Directorate, Logistics Research Division performed research on the automated

conversion of LCOM into Integrated Model Development Environment (IMDE) objects. IMDE is a simulation development system that embeds an object-oriented modeling approach within an interface to improve the user-friendliness of the LCOM. This research helped update the LCOM model to a more graphical interface orientation for the user (Carrico and Clark 1996).

 In 1981, a research effort performed by Robert Garcia and Joseph P. Racher, Jr. focused on the variance in workcenter performance based on skill level mixtures. They measured skill level effects and determined that skill mixture has an impact on workcenter performance. Garcia and Racher developed a methodology for capturing this effect and incorporating it into the LCOM model (Garcia and Racher 1981).

2.3. SIMSCRIPT II.5

The LCOM model is programmed using the California Analysis Center, Inc's (CACI) simulation language called SIMSCRIPT II.5. The software is relatively easy to use, with interactive graphical user interfaces and animated graphics (CACI 2006). The following excerpt is from CACI's manual titled, *Building Simulation Models with SIMSCRIPT II.5*:

> SIMSCRIPT II.5 is an integrated, interactive development environment controlled by SimLab. SimLab includes the complete SIMSCRIPT II.5 programming language, utilities for editing and managing SIMSCRIPT II.5 programs, the SIMGRAPHICS I and II graphical interface and utilities, and comprehensive online-help (Russell 1999).

Typical SIMSCRIPT II.5 applications include telecommunications, network analysis, transportation, manufacturing, health care, and military operations to include wargaming and logistics planning (CACI 2006).

 The current compiling environment for SIMSCRIPT II.5 has evolved from SimLab and is called Simulation Studio, or SimStudio (CACI 2006). It provides support for projects with hierarchical directories, and is available on all supported SIMSCRIPT II.5 platforms including Windows, PC Linux, and UNIX workstations (CACI 2006). SimStudio "…has a more intuitive graphical user interface, a modern look-and-feel, and incorporates SIMSCRIPT II.5 Syntax Color Coded Text Editor and all Graphical Editors for SIMSCRIPT II.5 Graphics" (CACI 2006).

2.4. Variance Reduction Techniques

Since models with random input, such as LCOM, consequently produce random output, there is a variance associated with the output over a given number of replications (Law and Kelton 2000). For the purpose of analysis and interpretation, the amount of computational time and appropriate statistical analysis can often be great. As Law and Kelton (2000) state, "Sometimes the cost of even a modest statistical analysis of the output can be so high that the precision of the results, perhaps measured by confidenceinterval width, will be unacceptably poor," so therefore the analyst should "...try to use any means possible to increase the simulation's efficiency." By the term efficiency they are talking about statistical efficiency, measured by the variances of the output random variables from the simulation (Law and Kelton 2000). Certain variance reduction techniques, when properly applied, may result in greater precision over fewer replications without disturbing its expectation (Law and Kelton 2000). This research focuses on three of these techniques: common random numbers (CRN), antithetic variates (AV), and control variates (CV). These techniques are described in depth below.

2.4.1 Common Random Numbers

 The common random numbers technique is unlike the other two techniques in that it is a multiple model technique, while CV and AV are both single model techniques. This means that CRN involves the comparison of two or more system configurations instead of investigating a single configuration (Law and Kelton 2000). CRN requires that the different configurations not only use the same random numbers, but that the numbers are synchronized to induce similar experimental conditions. This technique, also known as matched streams, or matched pairs, ensures that "…any observed differences in performance are due to differences in the system configurations rather than to fluctuations of the 'experimental conditions' " (Law and Kelton 2000). The following is a summary of the common random numbers theory described in Law and Kelton's *Simulation Modeling and Analysis* (2000):

Consider the case of two alternative configurations, where X_{1j} and X_{2j} are the observations from the first and second configurations on the *j*th independent replication, and we want to estimate equation (2.1):

$$
\zeta = \mu_1 - \mu_2 = E[X_{1j}] - E[X_{2j}]
$$
\n(2.1)

For *n* replications where $j = 1,2,...n$, and $E[Z_j] = \zeta$, and $Z_j = X_{1j} - X_{2j}$, then equation (2.2) is an unbiased estimator of ζ:

$$
\overline{Z} = \frac{1}{n} \sum_{j=1}^{n} Z_j
$$
 (2.2)

Furthermore,

$$
Var(\overline{Z}) = \frac{Var(Z_j)}{n} = \frac{Var(X_{1j}) + Var(X_{2j}) - 2Cov(X_{1j}, X_{2j})}{n}
$$
(2.3)

Obviously, if the simulations of the two different configurations are done independently, then

$$
Cov(X_{1j}, X_{2j}) = 0
$$
\n(2.4)

However, if there was a way to induce a positive correlation, then

$$
Cov(X_{1j}, X_{2j}) > 0
$$
\n(2.5)

and the value for $Var(\overline{Z})$ will be reduced (Law and Kelton 2000).

 The key to the induction of this positive correlation is the synchronization of random variate draws across the different system configurations on the same replication. As Law and Kelton (2000) state, "*Ideally*, a specific random number used for a specific purpose in one configuration is used for *exactly the same* purpose in all other configurations." The process of synchronization involves two steps. First, all points in the model where a random number or variate is drawn must be identified. Second, each point is assigned its own random number stream (Bednar 2005). Care should also be taken to ensure that streams of random numbers do not overlap. For this reason, a robust random number generator with built-in functions that keep track of random number streams should be used (Bednar 2005).

2.4.2 Antithetic Variates

 The antithetic variates technique, like CRN, induces a correlation between separate runs. However, AV is conducted on a single configuration. Also, the desired

correlation is negative rather than positive (Law and Kelton 2000). AV tries to induce this negative correlation by making pairs of runs of the model where complementary random variates drive the two runs in a pair. The complementary random variates are such that if U_i is a particular random number used for a particular purpose in the first run *j*, then $1 - U_j$ is used for the same purpose in the second run. The use of $1 - U_j$ is valid since $U \sim U(0,1)$ then $[1-U]$ is also $\sim U(0,1)$ (Law and Kelton 2000). Like CRN, synchronization is essential. The following theory is paraphrased from Law and Kelton's *Simulation Modeling and Analysis* (2000):

Suppose we make n pairs of replications resulting in observations of pairs $(X_{1i},$ X_{2j} ,…, (X_{1n}, X_{2n}) , where X_{1j} is essentially U_j and X_{2j} is $1 - U_j$. Both X_{1j} and X_{2j} are legitimate observations of the simulation model, therefore

$$
E[X_{1j}] = E[X_{2j}] = \mu \tag{2.6}
$$

If each pair is independent of the other pair, then the total number of replications is 2*n*, and

$$
X_n = \frac{\left(X_{1j} + X_{2j}\right)}{2} \tag{2.7}
$$

The average of the X_i 's, \overline{X} , is the unbiased point estimator of

$$
\mu = E[X_{1j}] = E[X_j] = E[\overline{X}]
$$
\n(2.8)

And

$$
\text{Var}\left[\overline{X}\right] = \frac{\text{Var}\left[X_j\right]}{n} = \frac{\text{Var}\left[X_{1j}\right] + \text{Var}\left[X_{2j}\right] + 2\text{Cov}\left[X_{1j}, X_{2j}\right]}{4n} \tag{2.9}
$$

At this point, like CRN, with no synchronization the covariance portion of equation (2.9) is zero. If we synchronize the pairs and induce a negative correlation between X_{1j} and X_{2i} , then

$$
Cov\left[X_{1j}, X_{2j}\right] < 0\tag{2.10}
$$

This will result in a lower $\text{Var}[\bar{X}]$, the overall goal of AV (Law and Kelton 2000).

2.4.3 Control Variates

 Like the previous two methods, "control variates attempts to take advantage of a correlation between certain random variables to obtain a variance reduction" (Law and Kelton 2000). However, this correlation in CV is not induced, but already exists during the course of the simulation. Furthermore, the sign of the correlation does not matter for CV (Law and Kelton 2000). Bednar (2005) gave a complete explanation of the theory behind CV in his thesis, *Feasibility Study of Variance Reduction in the Thunder Campaign-Level Model*, derived from lecture notes given by Bauer (2005) at the Air Force Institute of Technology. His explanation includes a full derivation of the methodology for control variates. This full derivation is quoted from Bednar and found in Appendix A.

 As Bednar (2005) points out, the selection of the controls is of importance to the analyst. The selection of potential controls has two steps. First, the analyst must identify an input of a random number or random variate. Then, the expected value of the random variate must be determined, given the random input parameters (Bednar 2005). Bednar describes the next steps in the following paragraph:

After all the potential candidates for controls are identified, the average of all the realizations of each potential candidate must be calculated and output in addition to the MOE [measure of effectiveness] of interest for each replication. Since the correlation between the control candidates and the MOE is unknown, a stepwise regression must be performed to identify which control candidates are significantly correlated to the selected MOE (Bednar 2005).

Once the controls are identified, the analyst applies the steps in Appendix A to arrive at a reduced confidence interval (Bednar 2005).

2.5. Previous Variance Reduction Studies

 Variance reduction techniques were first introduced in the early days of computers using Monte Carlo simulations (Law and Kelton 2000). The literature on this subject is extremely large, and the analysis of research would take far too long to discuss in the context of this research. However, the application of these variance reduction techniques is really the focus of this research. We have identified two sources of research that relate directly to the subject of our research effort. In 1983, Lieutenant Colonel Mohamed Elhefny performed a thesis exploring the application of different variance reduction techniques, comparing the results of the various techniques. He concluded, "there is no single technique which is the most suitable technique for every simulation problem," and recommended further research to identify techniques effective for specific simulation models (Elhefny 1983). This is significant because it tells us that the success or failure of these techniques can not be foreseen ahead of time. We must apply each method to the LCOM model to analyze the effectiveness of each method.

The Air Force Studies and Analysis Agency sponsored a master's thesis by Bednar (2005) while at the Air Force Institute of Technology. He applied numerous variance reduction techniques to the Air Force's THUNDER combat simulation model. The THUNDER model is also coded in SIMSCRIPT II.5 programming language.

Bednar applied CRN, control variates, antithetic variates, and various combinations of the three methods in order to determine the most effective variance reduction techniques for the THUNDER model. His results indicated that the control variates method performed the best of the variance reduction methods, while CRN and antithetic variates did not produce successful results. When combining the various techniques, he found that the combination of control variates and antithetic variates produced the most favorable results (Bednar 2005). Given that both models are written using SIMSCRIPT II.5, we hope to incorporate the techniques used by Bednar in THUNDER to apply the same variance reduction techniques to the LCOM model.

2.6. Random Number Generator

Since random number generators lie at the root of stochastic simulation, it is important that we define and investigate the types of random number generators that we will be dealing with in our research effort. According to Kelton, Sadowski, and Sturrock, the most common form still built into simulation models is a method called the linear congruential generator (LCG). This method was first introduced in 1951 by Lehmer. Most currently used LCG's have been thought to produce a moduli, or cycle length, of 2^{31} – 1, or a little over 2 billion numbers. Once considered an impressive cycle length, these numbers can be consumed in a matter of minutes by today's ordinary PC. This means the generator can potentially "lap" itself within a few minutes of simulation time, given an extremely high consumption rate for random numbers within the model. Furthermore, based on the research of L'Ecuyer and Simard (2001), the poor structure of the random numbers in these types of generators "…can dramatically bias simulation results for sample sizes much smaller than the period length" (L'Ecuyer et al. 2002).

The particular pseudo-random number generator coded in SIMSCRIPT II.5 is the Lehmer Pseudo-random Number Generator. The recursion is shown below (Lehmer 1969):

$$
X_{n+1} = K X_n \pmod{m} \tag{2.11}
$$

This equation is based on some modulus *m* and multiplier *K*. The starting seed is multiplied by a constant *K* to produce a new seed and a sample (Lehmer 1969). This generator performs like any normal LCG and produces a cycle length of 2^{31} -1 when the modulus *m* is a prime modulus and the multiplier *K* is one of more than 534 million full period multipliers (Law and Kelton 2000).

 This LCG found in SIMSCRIPT II.5 gives the modeler 10 separate random number streams from which to choose (Russell 1999). These 10 streams will be used to link or synchronize various sources of randomness in the model across multiple replications or various scenarios.

 Although several alternatives to LCG's exist in the random number generation world, we will apply the three variance reduction techniques using the current generator found in the SIMSCRIPT II.5 programming language.

2.7. Minitab and Regression

 Minitab 14 is a computer program designed to perform statistical functions. According to the software's website, Minitab is used by thousands of companies in more than 80 countries for implementation of Six Sigma and other data-driven quality improvement programs (Minitab 2007).

 Minitab 14 features several methods for performing regression analysis, such as linear regression, polynomial regression, logistic regression, partial least squares,

stepwise and best subsets, and residual plots (Minitab 2007). The stepwise regression analysis is used to investigate and model the relationship between a response variable and two or more predictors. Stepwise regression removes and adds variables to the regression model for the purpose of identifying a useful subset of the predictors, using what is called a forward selection to add variables and then a backward elimination to remove variables (Minitab 2007). The user simply specifies the response variable, the starting set of predictor variables, and the alpha value for adding or removing a variable to or from the model. The following is paraphrased from the explanation of the stepwise procedure as stated in the Minitab user manual:

The first step in stepwise regression is to calculate an *F*-statistic and *p*-value for each variable in the model. If the model contains *n* replications where $j = 1,2,...n$, then *F* for any variable, x_i , is

$$
F_{(1,n-j-1)} = \frac{(SSE_{(j-X)} + SSE_j)}{MSE_j}
$$
\n(2.12)

where $SSE_{(j - Xj)}$ = Sum of Squared Error for the model that does not contain x_j , and SSE_j = Sum of Squared Error and MSE_i = Mean Square Error for the model that contains x_i . For the backwards elimination, if the *p*-value for any variable is greater than the value specified in α to remove, then Minitab removes the variable with the largest *p*-value. If Minitab cannot remove a variable, the procedure attempts to add a variable. Minitab calculates an *F*-statistic and *p*-value for each variable that is not in the model. If the model contains *j* variables, then F for any variable, x_i , is

$$
F_{(1,n-j-1)} = \frac{\left(SSE_j - SSE_{(j+Xj)} \right)}{MSE_{(j+Xj)}}
$$
(2.13)

where $n =$ number of observations, $SSE_i =$ Sum of Squared Error before x_i is added to the model, and $SSE_{(i + X_i)}$ = Sum of Squared Error and $MSE_{(i + X_i)}$ = Mean Square Error after x_i is added to the model. For the forward selection, if the *p*-value corresponding to the *F*statistic for any variable is smaller than the value specified in *α* to enter, Minitab adds the variable with the smallest *p*-value to the model. When no more variables can be entered into or removed from the model, the stepwise procedure ends (Minitab 2007).

2.8. Analysis

Several statistical methods will be used to analyze the LCOM output and determine the effectiveness of each technique when applied to the model. The following sections describe these methods.

2.8.1. Confidence Intervals

For single model tests such as control variates, we will construct confidence intervals for the mean estimate using the following equation:

$$
CI = \overline{X} + \textcolor{red}{\textbf{-}t_{\alpha/2,n-1}} (S/\sqrt{n}) \tag{2.14}
$$

In this equation, μ is the mean, *t* is the quartile of the *t* distribution with *n*-1 degrees of freedom, *S* is the standard deviation, and *n* is the number of replications. This equation assumes the output for the mean follows a normal distribution (Banks et al. 2005). A reduction in CI width will be measured for a successful reduction of variance.

2.8.2. Determining the Number of Replications

In order to determine the number of replications to achieve a desired halfwidth *β*, we will apply the following equation (McClave et al. 2005):

$$
n = \frac{(t_{\alpha/2})^2 \sigma^2}{(SE)^2}
$$
 (2.15)

In equation 2.15, *SE* is known as the sampling error. In this case, the sampling error is equal to the half-width, *β*, of the confidence interval desired (McClave et al. 2005). A significant reduction in the number of replications required will indicate a successful application of a particular variance reduction technique (Bednar 2005).

2.8.3. Paired-t Confidence Interval

Finally, a paired t-test will be used to calculate a confidence interval about the difference between two values. This will tell us if the changes in the model result in a statistically significant values. To form the $100(1 - \alpha)$ confidence interval we use $X_{1j} - X_{2j} = Z_j$ for $j = 1$ to *n*. Therefore, the following equation for the confidence interval applies (Law and Kelton 2000):

$$
CI\left[\overline{Z}\right] = \overline{Z} \pm t_{(1-\alpha/2),n-1} \sqrt{\hat{V}ar\left[\overline{Z}\right]}
$$
 (2.16)

If the confidence interval includes zero we can not conclude that the two values are statistically significant at the specified α level, and the variance reduction technique has been ineffective in reducing the variance of the output.

3. Methodology

This chapter begins with the general research process for each variance reduction method. It then outlines the selection of the particular model used in the analysis, and the selection of output measures. This is followed by a description of the methodology behind the application of each variance reduction technique on the model. The chapter concludes with a discussion of the collection of output data through model runs and the preparation of the data for analysis.

3.1. Research Process

 The objective of this research is to study the impact of various variance reduction techniques on LCOM. The research will apply these techniques to the model and analyze the results, determine whether or not each technique is effective and identify the most effective method. The accomplishment of this objective is achieved through the following general steps:

- 1. Identify a particular measure of effectiveness (MOE) or set of MOE's in the model's output that are commonly used by LCOM users
- 2. Identify the locations in the model where random numbers are generated
- 3. Alter the model's source code in order to apply the desired variance reduction technique
- 4. Analyze the results from each variance reduction technique to determine the effectiveness of each technique

For the purpose of this research, the locations in the model where random numbers are generated and consumed will be referred to as random variate draws or points of

consumption. The two terms will be used synonymously throughout the next three chapters.

3.2. Model Selection and Parameters for LCOM Analysis

Development of an actual, real-world scenario for LCOM can take many months to develop (Erdman 2006). Fortunately, the LCOM program includes two example scenarios for training and education. The first scenario, the Bicycle Model, is a very simplistic model involving a bicycle used for delivering papers early every morning. The second scenario, the Joint Service FX-99 Generic Fighter Model, is loosely based on an F-16 aircraft maintenance unit, with information compiled from data at Hill AFB from July 1979 to June 1980 (ASC/ENM 2004). Since the reliability and maintainability parameters can be quickly changed and updated for generic applications, the name of the model has evolved to be known as the Generic Fighter Model (ASC/ENM 2004).

Unfortunately, the simplicity of the bicycle model is such that there is little or no variance in the original model, so the application of variance reduction techniques would obviously show no significant improvement over the original. On the other hand, the generic fighter model is much more realistic and similar to current LCOM models in use today. Additionally, the variance in the output is sufficiently large enough for a reduction in the variance to be possible and desirable. For these reasons, the Generic Fighter Model was selected for analysis in this study.

The original configuration of the generic fighter model is sufficient for both AV and CV, so the default configuration was used for these techniques. However, since CRN requires the comparison between two different configurations, a modification of the model for this technique was necessary. To modify the model in a way that would have a
significant effect on the outcome of the MOE's, the manpower availability for the FLTL manpower type was adjusted and the model run with a 30-replication production run until a significant difference in the two output statistics of interest were empirically observed. The output statistics of interest will be explained in detail in the next section. Then, a paired difference test of hypothesis was conducted for $\mu_d = (\mu_l - \mu_2)$ in order to verify that the mean for the two scenarios are not equal (for results and calculations of paired difference test see Appendix B). In the original model, parts and labor are both essentially unlimited. For the second scenario, the model simply contains a constrained value for FLTL manpower of 50 workers available for each shift.

The period of interest for this experiment consisted of a 5-day period following a 20-day warm-up period. After consulting with LCOM users it was determined that a 20 day warm-up was likely to be sufficient for the model to exhibit a steady-state behavior, based upon their experiences and recommendations (Erdman 2006). Additionally, a 5 day observation period from day 21 to day 25 would provide a significant amount of variance before the implementation of the variance reduction techniques.

3.3. Output Measure Selection

The LCOM generates dozens of statistics in its output, with most calculations typically experiencing some sort of variance across multiple replications. However, after speaking with expert LCOM users, the MOE that interests the users the most is the statistic labeled C15 on the output tool, known as *Overall Achieved Sorties per Aircraft per Day* (Erdman 2006). This output data is located in the production run merged output report. An example of a typical merged output report for the Bicycle Model is contained in Appendix C.

LCOM automatically collects and calculates the information for these statistics. Although LCOM generates several dozens of various statistics, the decision was made to analyze the variance of two output statistics in particular, C15 – *Overall Achieved Sorties per Aircraft per Day*, and C24 – *Mission Capable Rate*. This decision was reached for several reasons. Since LCOM users focus almost exclusively on the C15 statistic, the reduction in the variance of any other statistic separate from C15 would have little or no impact on the way the model was run and analyzed. Additionally, *Mission Capable Rate* has a direct impact on the sortie generation rate, so C24 was also included. Second, the Generic Fighter Model produces a fair amount of variance in both output statistics C15 and C24, making variance reduction a feasible and desirable goal for our research and analysis.

3.4. Common Random Numbers

As noted previously, the random number generator in SIMSCRIPT II.5 allows the user to identify up to 10 unique random number streams. In the LCOM's source code for the main model, 33 separate random variate draws or points of consumption were identified. In the original source code, LCOM uses only the first 9 of the 10 available random number streams. The model's random number generators are organized by function, dedicating a stream to different tasks in the manner shown in Table 3.1 (ASC/ENM 2004):

Stream	Function
	Attribute initial values, % sortie to complete (AABORT & ATTRIT) Ram time
2	Task Durations
3	Failure Clock Operations
4	Time Accumulating Attributes Random Setting
5	Probability of Air Abort, Attrition, or Ram Repair
6	Task Selection on A, E, and G Selection Modes
	Random Multiplier for Initial Failure Clock Settings
8	Sortie Length (Task Time Option)
9	Not used
10	Unknown

Table 3.1: LCOM Random Number Streams and Associated Functions

LCOM allows the user to specify or create starting seeds for each of the first 8 streams specified above, using the LCOM interface screen shown in Appendix D. However, several different random variate draws within the model sampled from the same stream. This means that as the model's inputs change, the one-for-one relationship between random variates across different scenarios is lost. Even with the ability to set seeds for each stream and synchronize the random number streams, the lack of a one-for-one relationship between random variates consumed across different scenarios makes the application of CRN unlikely to be successful or consistent in the pursuit of variance reduction.

Each of the 33 points of consumption in the model was investigated one-by-one to determine the relative use of each point in the model where random numbers were generated and consumed. After investigating the 33 random variate draws in the LCOM main model source code, it was determined that 8 random variate draws were in use for FX99 model. In other words, only 8 points of consumption in the default settings of

FX99 actively generated and consumed random numbers. The remaining 25 points of consumption were completely inactive. The random variate draws in use usually shared the same random number stream. Figure 3.1 illustrates the total number of random variate draws assigned to each random number stream in the main model code.

Figure 3.1: Configuration of Random Variate Draws by Stream

The code was modified so that each active random variate draw was synchronized by identifying a unique random number stream to each of the points in the model. The modified configuration of random variate draws is shown in Figure 3.2, with the modified random variate draws shown in red.

Figure 3.2: Modified Configuration of Random Variate Draws by Stream

Since less than 8 of the 33 random variate draws are active in the the FX99 model, the current configuration of LCOM allowed the synchronization of the random variate draws in use by simply identifying the same seed set across scenarios in the ISEEDS tab of the production run interface shown in Appendix D. Then, as an additional precaution, the remaining 25 inactive points of consumption were split to share the remaining $9th$ and $10th$ random number streams.

 Once the source code was modified to reflect these changes, the code was compiled using Simscript's SimStudio compiling environment and an LCOM production run of 30 replications was conducted with the new, modified source code using the default scenario with unconstrained manpower. Then, the scenario was changed to the second scenario with constrained FLTL manpower resources and the 30-replication production run was completed again. The paired difference from each replication was used to create the confidence interval and halfwidth.

 Similarly, the unmodified model was run and a halfwidth calculated for both the unconstrained manpower scenario and the constrained manpower scenario in order to provide a base from which to compare the new results. This paired difference across the two different scenarios was used to calculate the confidence interval and halfwidth in the exact same manner as the CRN model described above. The reduction in the confidence interval halfwidth was calculated to determine the degree of variance reduction.

3.5. Antithetic Variates

Whereas CRN requires a comparison between two different scenarios, antithetic variates relies on an induced correlation among replications within the same scenario. Since no additional scenarios are required for this technique, the default settings for the 48-aircraft Generic Fighter Model were used.

Recall that if U_k is a particular random number used for a particular purpose in the first run, then $1 - U_k$ is used for the same purpose in the second run. Fortunately, SIMSCRIPT II.5 allows the incorporation of antithetic variates with some very simple modifications. According to the manual, *Building Simulation Models with SIMSCRIPT II.5*, "To use an antithetic variate in any random deviate generator in SIMSCRIPT II.5, it is merely necessary to negate the random number stream parameter to the function **random.f"** (Russell 1999). Random.f is the SIMSCRIPT II.5 function for a random number drawn between 0 and 1. For example, a portion of the LCOM source code is shown in Figure 3.1:

> LET VALUE.ATB (OWNER, SEQ, TAIL, 3) = REAL. F (INT. F (RANDOM. F (4)))

Figure 3.3: LCOM Random Variate Example

The random variate in this example, $\text{RANDOM. F}(4)$, is a pseudo-random number uniformly distributed between 0 and 1, drawn from SIMSCRIPT II.5's random number stream 4. In order to implement antithetic variates, the portion of the code RANDOM. F(4) simply becomes RANDOM. F(-4) in order to turn U_k into $1 - U_k$ (Russell 1999) . Every point in the model where the function random.f is used, the random number stream parameter must be negated. This includes all 33 random variate sources, or points of consumption in the model.

The problem of random variates must also be considered. Some sources of randomness in LCOM draw from commonly used distributions, such as normal, lognormal, exponential, and poisson. These points in the model typically appear with the distribution name, distribution parameters, and stream number from which to sample. For example, a random number drawn from a normal distribution may appear as normal.f(MU, SIGMA, 3), with stream 3 specified for this draw. In this case, U_k is still a random number between 0 and 1, generated using random.f before SIMSCRIPT II.5 converts the number according to the distribution and parameters. Consequently, like the random variates created using random.f, the stream is simply negated and SIMSCRIPT II.5 conducts the process of turning $1 - U_k$ into the appropriate random deviate, given the distribution and specified parameters (Russell 1999).

 Like CRN, the new code is modified to reflect the change at each of the 33 points of consumption in the model, compiled in SimStudio, and run using both the unmodified code as well as the modified code with antithetic variates. Like CRN, a one-for-one relationship between U_k and $1 - U_k$ is essential to synchronize the model and induce the correlation required to achieve a reduction in variance. For this reason, careful attention

must be paid to making sure the starting seeds for each stream remained the same for both the original model and the model with the antithetic variates. In this case, unlike CRN, the success of the model depends on the ability to ensure a one-for-one relationship across models without changing the scenario parameters. Consequently, the starting seeds are simply created and identified for each replication using the ISEEDS tab of the production run portion of the LCOM graphical user interface (see Appendix D). A production run of 15 replications was made with the unmodified code and 15 with the antithetic variates code. The output for both productions runs were combined and averaged in order to obtain the new confidence interval calculation and halfwidth. The results were then compared to a 30-replication production run with the unmodified model.

3.6. Control Variates

The Control Variate method relies on the relationship between a random variate input and the output variable of interest. This correlation along with the deviation from the known expected value of the input variable are used to adjust the output variable up or down, closer to the true but unknown mean. In LCOM, the number of potential controls can reach the hundreds. However, each potential control must be considered by capturing each random variate value drawn in the model, along with the expected value, distribution type and parameters over the course of the 5-day period.

The first step in this process involved modifying the LCOM main model source code in order to capture each random variate drawn and consumed in the model. Capturing the random variates required the use of the "print" command to dump the random variates in an output file called a PSR report for each replication in the LCOM

production run. Two different examples of modified random variate draws are shown in

Figure 3.2, the first from a logical expression and the second through a subroutine:

```
IF VALUE EO O
         LET VALUE = RANDI.F(1,100000,6)
print 1 line with value thus
* * * * * * *
```

```
LET DURATION = DRAW(DIST, FIRST, SECOND, 2, TYPE. TASK(TDIM1, TDIM2))
print 1 line with duration, dist, first, second thus
always
```
Figure 3.4: Two Modified Random Variate Draws for Control Variates

In the first case, RANDI.F is a command that generates a discrete random integer between 1 and 100000 using stream number 6. In the second case, DRAW references a routine that generates a random number based on the variables listed in parentheses. For the purposes of this exercise, the first four variables are the only variables of significance. The first variable is the distribution type, the second is the first parameter of the specific distribution, the third is the second parameter, and the fourth is the random number stream. All random variates used in LCOM were modified in the same manner as the two shown in Figure 3.2.

Once in the PSR report, the random variates were cut and pasted to an Excel spreadsheet where the numbers were sorted and separated by distribution type and parameter. Once sorted, each unique distribution formed the basis for a potential control to be analyzed. This means that different distribution types and parameters could come from a single random variate draw. The total number of potential controls was 23. After eliminating those potential controls with no random variates drawn in any single replication, the number of potential controls decreased to 20. In actuality, 13 of the 20

potential controls came from a single random variate draw in the task duration function. Since regression theory requires at least as many replications as potential controls, 21 replications were performed (McClave et al. 2005).

Calculations for each of the 21 replications were performed in Excel, recording the control number, the first parameter (Parameter1), the second parameter (Parameter2), the observed mean (Value), and the calculated difference between the expected and observed means. Finally, the output statistics C15 and C24 for each replication were recorded. An example is shown in Table 3.4:

Control	Parameter1	Parameter ₂	difference Value
1	0.004167	0.00125	0.004296 -0.00013
$\overline{2}$	0.007083	0.002083	0.006909 0.000174
3	0.008333	0.0025	0.008226 0.000107
$\overline{4}$	0.010417	0.002917	0.010755 -0.00034
5			0.011666670.0029166670.012967 -0.0013
6	0.0125	0.00375	0.011788 0.000712
$\overline{7}$	0.012917	0.00375	0.013393 -0.00048
8	0.020833	0.005833	0.020451 0.000382
9	0.02125	0.005833	$0.022362 - 0.00111$
10	0.025	0.007083	0.025813 -0.00081
11	0.033333	0.009583	-0.0004 0.033729
12	0.034167	0.009583	0.03227 0.001897
13	0.041667	0.024167	0.037604 0.004063
14	50000.5	3000	52359.79 -2359.29
15	50000.5	6000	48885.22 1115.282
16	50000.5	10000	47602.14 2398.362
17	50000.5	15000	48855.5 1144.995
18	50000.5	50000	47690.16 2310.335
19	50000.5	75000	60801.71 -10801.2
20	50000.5	1-100000	49346.3 654.2028
C15	1.72		
C ₂₄	67.91		

Table 3.2: Potential Control Observations, Single Replication

A copy of the observations for each replication can be seen in Appendix F. Controls 1-13 are all lognormally distributed, with the expected mean in Parameter1 and the expected standard deviation in Parameter 2. Controls 14-20 are all uniform, discrete random variables between 1 and 100,000. In the case of controls 14-20, Parameter1 refers to the expected mean and Parameter2 refers to a cutoff point for decision tree task selection.

 The differences between the observed and expected means for each potential control were arranged in a 20 X 21 matrix and exported to Minitab for the stepwise regression calculation (See Appendix G). The predictors identified in the stepwise regression automatically became the control variates used to perform the calculations listed in Appendix A, ultimately concluding with a new, smaller confidence interval and halfwidth.

3.7. Data Collection

When conducting a production run in LCOM, the user specifies the number of replications desired for the run using the screen shown in Appendix E. After the run, the output from each replication is stored in a PSR report. One PSR report is created for each replication. Then, a merged post processor merges statistics from the individual PSR statistics files output over all replications of the main model specified in the production run (ASC/ENM 2004). The output of this file, shown in Appendix B, includes the mean value, the standard deviation, minimum and maximum values across replications, and automatically computes a 95% confidence interval for each statistic (ASC/ENM 2004). For the purpose of this research, the confidence interval, mean, and standard deviation are used primarily for the analysis of each technique. Additionally, the output statistic for each replication must be obtained for use in the paired difference when analyzing the

common random numbers technique. This information is obtained from the PSR file for each individual replication. Likewise, all random variate draws were coded in the main model source code to dump to the PSR file after each individual replication.

3.8. Analysis

Chapter 2 discusses the statistical calculations involved in the analysis of this experiment. These calculations include confidence intervals about the mean, confidence intervals about a difference between two values, and replications for a desired halfwidth. The calculations are used to determine the variance reduction achieved by each technique.

First, for the base, unmodified model, a confidence interval about the mean was identified and a confidence interval halfwidth was calculated using the results of a production run of 30 replications. For techniques involving a single scenario, such as antithetic variates and control variates, a confidence interval about the mean was identified for each of the two techniques in the merged output report, along with the calculation of the new confidence interval halfwidths. The percent improvement over the base model in the confidence interval halfwidth achieved by the each variance reduction technique determines the degree of variance reduction.

Additionally, if any level of variance reduction is observed in the output, the improvement can be approximated by determining the number of additional replications required by the unmodified model in order to achieve the halfwidth observed in the model with the variance reduction technique applied. This demonstrates the additional effort required to achieve the precision of the new model.

For the common random numbers technique, the base model must include the same two scenarios included in the CRN models. Recall that the modification included unconstrained and constrained manpower scenarios. The CRN code and the unmodified code were compiled and run for the two different scenarios. Then, a paired t-test was used to calculate the confidence interval of the difference between the two values across the different scenarios. This calculation included the differences in the two scenarios for both the new, CRN code and the unmodified source code. Now, using the same method as the other two techniques, the confidence interval halfwidths were calculated and the percent improvement determined the degree of variance reduction.

4. Results and Analysis

This section contains results and analysis performed to determine the effectiveness of each variance reduction technique applied to LCOM. The chapter describes the results achieved from common random numbers, antithetic variates, and control variates.

4.1. Common Random Numbers

The common random numbers model was compared to the base model using the difference between the two scenarios with constrained and unconstrained manpower as specified in Chapter 3. The full results for each replication can be seen in Appendix H. A summary of the results for the C15 statistic, *Overall Achieved Sorties per Aircraft per Day*, with the confidence interval (CI) calculated about the mean difference from replication to replication between the two scenarios, is shown in Figure 4.1:

Base -- C15		
95% CI:	0.2153527750.26398056	
CI halfwidth: 0.024313892		
CRN -- C15		
95% CI:	0.1669436580.23305634	
CI halfwidth: 0.033056342		
provement:	-3596%	

Figure 4.1: Common Random Numbers Results Summary, C15 Statistic

Likewise, the same calculations were conducted for the models using the C24 statistic, *Mission Capable Rate*, as the variable of interest. A summary of the results for the C24 statistic is shown in Figure 4.2:

Figure 4.2: Common Random Numbers Results Summary, C24 Statistic

 A negative improvement indicates an increase in the confidence interval size, the opposite of the desired effect in this experiment. As Figure 4.1 shows, the confidence interval for the C15 statistic did not improve from the base model to the CRN model. In fact, the confidence interval halfwidth was significantly larger in the CRN model. On the other hand, the confidence interval halfwidth for the C24 statistic improved slightly by 7.28 percent.

To put this improvement in the C24 variance into perspective, equation 2.15 was used to determine the number of replications required by the original model in order to achieve the same confidence interval halfwidth observed in the CRN model. In order to achieve the same confidence interval halfwidth as the CRN model, the user would need to run approximately 35 replications in order to achieve the same confidence in the C24 output statistic, versus 30 replications with the CRN model.

4.2. Antithetic Variates

 The results for antithetic variates were even less favorable than the common random numbers results. In this experiment, the base, unmodified FX99 model was run a total of 30 replications. Then, the main model code was modified to incorporate the antithetic variates at each point in the model where a random variate draw occurs. The code was compiled and the FX99 model was run again for a total of 30 replications, 15

with *Uj* and 15 with 1-*Uj*. The full results can be seen in Appendix I. A summary of the results for the C15 output statistic is shown in Figure 4.3:

```
Base -- C15 
95% CI: 2.174569922.230763413
CI halfwidth: 0.02809675 
Antithetic Variates -- C15 
95% CI: 2.145646262.216353742
CI halfwidth: 0.03305634 
Improvement: -25.83%
```
Figure 4.3: Antithetic Variates Results Summary, C15 Statistic

Like the common random numbers experiment, the confidence interval for the C15 output statistic does not improve, but actually worsens by more than 25 percent. Similar results were observed in the C24 output statistic in this experiment. A summary of the results is shown in Figure 4.4:

```
 Base -- C24 
95% CI: 66.593192167.21947454
CI halfwidth: 0.31314121 
Antithetic Variates -- C24 
95% CI: 66.7662962 67.59970375
CI halfwidth: 0.41670375 
Improvement: -33.07%
```
Figure 4.4: Antithetic Variates Results Summary, C24 Statistic

Contrary to CRN, the halfwidth for the C24 output statistic got much worse after the implementation of the antithetic variates.

 It is apparent that the synchronization techniques, common random numbers and antithetic variates, do not improve the variance of the output statistics. In fact, the improvement observed in the variance of C24 in the CRN model more than likely occurs

due to the randomness of the model, and not due to any true synchronization. When examining the random variate draws in the model and the purpose of each specific draw, it is easy to understand why synchronization is not possible the way the model is currently constructed. For example, task durations all originally sampled from random number stream 2. It was determined that the task durations were generated using three different random variate draws in different places in the model. However, upon further examination, the vast majority of task durations are all generated and consumed from a single random variate draw. For a maintenance unit, this includes all scheduled and unscheduled task times generated for every flightline, backshop, and depot task, regardless of the priority. In the current configuration of LCOM, essentially all tasks sample from a single point in the model. As discussed in Chapter 3, even with the ability to set seeds for each stream and synchronize the random number streams, the lack of a one-for-one relationship between random variates consumed across different scenarios renders the current configuration useless for the application of both common random numbers and antithetic variates. If the individual tasks are not separated, the one-for-one relationship between random variates consumed across different scenarios is not achievable.

4.3. Control Variates

 Unlike the previous two methods, the control variates method does not attempt to induce a correlation in the model. Instead, the control variates method attempts to take advantage of a known correlation between a random variate and the output variable of interest.

As described in Chapter 3, and similar to the other two experiments, this method was applied to two different output variables of interest, C15 – *Overall Achieved Sorties per Aircraft per Day*, and C24 – *Mission Capable Rate*. Since 20 different potential controls were initially identified, the model was run with 21 replications in order to satisfy the regression theory requirement for at least as many replications as potential predictors (McClave et al. 2005). Additionally, each potential control was traced to a particular function within the FX99 model. The controls and their associated functions where known are shown in Table 4.1:

	Control Parameter1	Parameter ₂	Function
1	0.004167	0.00125	End of runway check
2	0.007083	0.002083	γ
3	0.008333	0.0025	Start engines
$\overline{4}$	0.010417	0.002917	Load MBRK
5	0.0116666670.002916667UnMSRK		
6	0.0125	0.00375	Load Bomb
7	0.012917	0.00375	Load missiles
8	0.020833	0.005833	Jhalon service
9	0.02125	0.005833	Jlox service
10	0.025	0.007083	Load chaff dispenser
11	0.033333	0.009583	Do preflight
12	0.034167	0.009583	Jtanks
13	0.041667	0.024167	refill lox cart
14	50000.5		multiple networks, multiple tasks
15	50000.5		
16	50000.5		
17	50000.5		
18	50000.5		
19	50000.5		
20	50000.5		

Table 4.1: List of Potential Controls and Associated Functions

Table 4.1 lists the potential control in the first column. The second column, Parameter1, is the expected mean. The column labeled Parameter2 is the expected standard deviation. Finally, the function of each control, if known, is listed in the fourth column. Controls 1- 13 appear to be flightline pre- and post-sortie tasks and are all lognormally distributed. Controls 14-20 are random variates for node selection in task networks with multiple options and a specific probability associated with the task selection. Controls 14-20 are all random variates drawn using the RANDI.F function generating a uniform, discrete random integer between 1 and 100000. The specific function for Control 2 could not be identified.

After calculating the difference between the known and expected value for each potential control following each replication (Appendix F), the values were placed in the 20 X 21 matrix shown in Appendix G. The 20 X 21 matrix provided the data for the stepwise regression, and the data was imported in Minitab 14 for the stepwise regression calculations, first using C15 as the response variable. An initial alpha level of 0.05 was set, and the stepwise regression returned just one control, the UnMSRK task, which involves removing the missile rack from the aircraft.

After performing the control variate calculations in Appendix A, the new, reduced confidence interval was calculated. The improvement was drastic, and can be seen in Figure 4.5:

Figure 4.5: Control Variates Results Summary, C15, Single Control

Using the calculations in Appendix A to take advantage of the correlation between the UnMSRK task and the C15 output statistic, the confidence interval halfwidth decreased by 56.23 percent. To put this improvement into perspective, the number of replications was calculated in order to achieve the same level of confidence with the original model using equation 2.15 from Chapter 2. In order to achieve the same confidence interval halfwidth as the control variates model, the user would need to run approximately 115 replications in order to achieve the same confidence in the C15 output statistic, versus 21 replications with the control variates model.

 For the second experiment with the response variable replaced by C24, the stepwise regression determined, like the previous experiment with C15 as the predictor, that a single control was a predictor of the response variable, C24. In this case the predictor was control 10, Load Chaff Dispenser. The results from this experiment are shown in Figure 4.6:

\vert Base -- C24	
95% CI: 67.5334843 69.1293728	
CI halfwidth: 0.79794424	
Control Variates -- C24	
95% CI: 68.3349381 68.3421545	
CI halfwidth: 0.00360816	
Improvement: 99.55%	

Figure 4.6: Control Variates Results Summary, C24, Single Control

The improvement in the confidence interval halfwidth shown over the original model using control 10 as a predictor in the FX99 model is exceptional. The results indicate an extremely strong relationship between the load chaff dispenser control and the C24 output statistic. To achieve this same level of confidence with the original model, the user would need to make over a million additional replications with the original model. This is impractical, given the time constraints involved with the LCOM model runs.

While the results for the control variates experiment show a significant improvement in the confidence interval halfwidth, the stepwise regression revealed only one control with the alpha level of 0.05. Since typical, real-life LCOM models may incorporate many more variables with a much higher complexity, the possibility of a particular model possessing multiple controls is great. In order to demonstrate the technique using multiple controls, the stepwise regression was performed a second time for the model, using C15 as the response variable, with an alpha level of 0.15. In this case the stepwise regression concluded with 5 controls identified as predictors of the response variable, C15. The five controls identified are listed in Figure 4.7:

> Control 5 – UnMSRK Control 10 – Load Chaff Dispenser Control 14 – Task Selection Control 15 – Task Selection Control 20 – Task Selection

Figure 4.7: List of Five Controls Identified With 0.15 Alpha Level

The results for this experiment were not as favorable as the previous experiment with just one control, but still showed an significant improvement over the original confidence interval halfwidth. This is expected, since the alpha level was relaxed to 0.15, making the controls weaker predictors of the response variable, C15. The results for the multiple control variate experiment are summarized in Figure 4.8:

Base -- C15		
95% CI: 1.674593824 1.7692157		
CI halfwidth: 0.047310938		
Control Variates -- C15		
95% CI: 1.691849065 1.761883229		
CI halfwidth: 0.035017082		
Improvement:	259%	

Figure 4.8: Control Variates Results Summary, C15, Multiple Controls

In this case, the user would be required to make approximately 63 replications in order to achieve the same level of confidence with the original model. All calculations for the control variates experiments were made in Microsoft Excel. A copy of the full results can be seen in Appendix J.

5. Discussion

This section contains all of the conclusions from this research and recommendations for further research concerning the topic of variance reduction within LCOM. It contains conclusions summarized from the analysis of Chapter 4. Additionally, it concludes with recommendations for further research.

5.1. Conclusions

5.1.1. Common Random Numbers

The common random numbers experiment exhibited no improvement when using the C15 statistic as the output variable of interest, but exhibited limited improvement when using the C24 statistic. After investigating the configuration of random variate draws in the original model, the variance of the C24 output statistic was not uniformly reduced.

 Recall that the common random numbers theory relies on synchronization of random variate draws. Each function within the model would need a separate random variate draw that could be synchronized by designating a unique random number stream or starting seed. Furthermore, in a model such as LCOM where nearly every individual task has a significant impact on the output variables of interest, the individual tasks themselves must also be isolated so they can each be given a unique random number stream or starting seed. In the current configuration, LCOM often uses a single random variate source to generate input samples for multiple functions as well as dozens or even hundreds of tasks. Consequently, improvements observed using the common random

numbers method cannot be guaranteed. In this configuration, common random numbers is not a feasible method of variance reduction.

5.1.2. Antithetic Variates

The results for both the C15 output statistic and the C24 output statistic were unfavorable for variance reduction. In fact, just the opposite occurred. Both experiments concluded with an increase in the variance, rendering the use of antithetic variates unsuccessful in the search for a feasible variance reduction method.

 Like common random numbers, antithetic variates relies on random variate synchronization in order to induce a correlation. In CRN, this correlation occurs across multiple scenarios. In AV, the correlation is induced within replications of a single scenario. The AV method is unsuccessful in LCOM for the same reason the CRN method failed – the inability to synchronize random variates in LCOM used for the same purpose. Since several random variate draws generate input for multiple purposes within the model, the possibility for synchronization within the current framework does not exist. Also, like common random numbers, the results are simply due to randomness. Consequently, the increases in the confidence intervals using the AV method do not necessarily indicate a less effective model, but simply an ineffective method of variance reduction.

5.1.3. Control Variates

Unlike the previous two methods, control variates capitalizes on an existing correlation between random inputs and a particular output variable of interest. In this experiment using the FX99 model, control variates performed extremely well. In all cases using both the C15 output statistic and the C24 output statistic, control variates

produced a significant improvement in the output variance. A significant reduction in the output variance equates to a more accurate estimate of the mean. Additionally, with a halfwidth goal in mind, a reduction in variance can reduce the amount of time the user spends performing additional replications in order to achieve that specified halfwidth goal. In one case, as shown in the CV experiment using the C24 output statistic, the variance reduction achieved by the control variates model was so significant that over a million replications would be necessary to achieve the same halfwidth with the unmodified LCOM model.

5.2. Recommendations for Further Research

5.2.1. Improvements to LCOM – Implementation of Control Variates

 Implementing a permanent option for the control variates method in LCOM is relatively simple for a particular model. However, the controls may change from model to model. Therefore, the potential controls for each model must be captured and arranged in the same manner as the matrix found in Appendix G, listing potential controls and the output variable of interest. This can be performed fairly easily by a postprocessor. It would first require the postprocessor to capture the random variates and their known expected value, and subtracting the observed mean from the expected value to obtain the difference. Then, the stepwise regression must be performed in order to identify the actual predictors of the response variable. A stepwise regression could be performed by a postprocessor or a statistical software package such as Minitab.

Once the controls are identified, the calculations in Appendix A must be performed in order to calculate the new, improved confidence interval. These steps can

also be performed by a postprocessor – one for a single control and a separate postprocessor for multiple controls, since the methodology is slightly different.

5.2.2. Improvements to LCOM – Random Number Generator

As discussed in Chapter 2 of this research, the LCG used for random number generation in the LCOM model could be considered inadequate given the relatively small period, or stream length, and the small number of random number streams offered by the LCG. With 33 random variate draws, the modeler ideally should have at least 33 different streams of numbers in order to dedicate a unique stream to each particular point of consumption. In LCOM, multiple tasks often sample from the same random variate draws, so ideally the tasks and functions would be separated and isolated. Then, a unique stream or seed value should be dedicated to each random variate draw used for a particular purpose. Since the LCG in SIMSCRIPT II.5 only has 10 options to choose from, other options for random number generation should be explored.

Several alternatives to LCGs exist in the random number generation world. One of these alternatives is found in the simulation modeling software Arena Version Seven and later, called a combined multiple recursive generator. The generator is titled, MRG32k3a by L'Ecuyer (L'Ecuyer et al. 2002). This generator, based on research by L'Ecuyer, Simard, Chen, and Kelton, "differs in that (1) it involves two separate component generators that are then combined, and (2) the recursion to get the next values looks back beyond just the single preceding value" (Kelton et al. 2004). The cycle lengths are much longer than LCGs – instead of burning the entire stream of numbers in a matter of minutes, the new generator will take an ordinary personal computer 2.78×10^{40} millennia. Furthermore, the number of separate streams improves from 10 unique

random number streams to 1.8×10^{19} streams (Kelton et al. 2004). Fortunately, L'Ecuver and his colleagues have developed a package that allows a simple implementation in various languages such as Java, C, and C++. An experienced programmer could easily explore the implementation of L'Ecuyer's multiple recursive generator in LCOM, allowing for synchronization of all points of consumption for common random numbers.

5.2.3. Further Variance Reduction Applications to LCOM

This research effort explored the application of each of the three variance reduction techniques – common random numbers, antithetic variates, and control variates – in isolation. However, the possibility exists for combining multiple variance reduction techniques and applying them to LCOM in order to achieve an even greater level of variance reduction observed with a single technique, as was done by Bednar (2005).

Furthermore, the techniques were applied to a single, hypothetical model developed for educational aid. While the basic logic of the control variates model remains the same, the technique should be applied to real-world models and analyzed for similar effectiveness.

5.2.4. Impact to the LCOM Optimizer

After the techniques are implemented into real-world models with some level of improved variance reduction, the next question is, how do the techniques affect the performance of the optimizer? Can the control variates method improve the performance of the optimizer? If not, can the control variates method be applied to the optimizer itself? In theory, a successful variance reduction application to LCOM would mean a more effective, predictable optimizer performance. However, the actual effectiveness should be tested and analyzed to measure the improvement, if any exists.

Appendix A: Control Variates Derivation

The following derivation is quoted directly from Bednar's thesis, Feasibility of Variance Reduction in the Thunder Campaign Model (2005):

Consider the case were there is only a single control. Now, assume there is a mean response of interest from the simulation called μ_Y for which *Y* is an estimator. Also, assume there is another output variable, *X*, that is correlated with the *Y* response and has an expected value μ_X that is known. Since *X* is correlated with the *Y* variable, it is known as the control variable. Now consider the controlled estimator *Y*(*b*) given in equation $((A.1)A.1)$ $((A.1)A.1)$ where *b* is a constant.

$$
Y(b) = Y - b(X - \mu_X)
$$
\n^(A.1)

Note that *Y*(*b*) is an unbiased estimator of μ_Y by equation (A.2).

$$
E[Y(b)] = E[Y] - E[b(X - \mu_X)]
$$
\n(A.2)

$$
E[Y(b)] = \mu_Y - b(\mu_X - \mu_X) = \mu_Y
$$
\n(A.3)

The variance of *Y*(*b*) is given in equation (A.4).

$$
Var(Y(b)) = Var(Y) + b^2 Var(X) - 2bCov(Y, X)
$$
 (A.4)

With a little manipulation of equation (A.4) it can be shown that the variance of *Y*(*b*) is smaller than the variance of Y if equation (A.5) holds.

$$
2b\text{Cov}(Y,X) > b^2\text{Var}(X) \tag{A.5}
$$

In observing equation (A.5), it is apparent that if the variables *X* and *Y* are independent, then the Cov $(X, Y) = 0$ and it follows that there can be no reduction in variance of *Y*.

To find the value of *b* that minimizes equation (A.4) the derivative with respect to *b* is found.

$$
\frac{\partial \text{Var}(Y(b))}{\partial b} = 2b\text{Var}(X) - 2\text{Cov}(Y, X) \tag{A.6}
$$

From equation (A.6) the minimum point is found by setting the derivative to zero. Equation (A.7) is the candidate for *b* that minimizes equation [\(A.4\)](#page-63-0)(A.4) thus labeled β .

$$
\beta = \frac{\text{Cov}(Y, X)}{\text{Var}(X)}\tag{A.7}
$$

To verify this is a minimum, the second derivative is found.

$$
\frac{\partial^2 \text{Var}(Y(b))}{\partial b^2} = 2\text{Var}(X) \tag{A.8}
$$

Since the variance is always non-negative, then equation (A.7) is the value for *b* that minimizes equation $(A.4)$. Combining equation $(A.4)$, equation $(A.7)$, and using some simple algebra yields equation (A.9).

$$
Var(Y(\beta)) = (1 - \rho_{XY}^2)Var(Y)
$$
 (A.9)

Where ρ_{XY} is the correlation coefficient between X and Y. Var $(Y(\beta))$ is the minimum variance. The controlled observations (A.10) are averaged (A.11) to obtain an unbiased estimator of μ_{γ} .

$$
Y_i(\beta) = Y_i - \beta (X_i - \mu_X), i = 1, ..., K
$$
\n(A.10)

$$
\overline{Y}(\beta) = \frac{1}{K} \sum_{i=1}^{K} Y_i(\beta)
$$
\n(A.11)

where K is the sample size.

Since the value of β is unknown, it must be estimated. An estimate of β can be found by substituting the sample quantities into equation (A.7) This solution is the least squares solution for β . The least squares solution is also the maximum likelihood solution with the assumption of joint normality between *X* and *Y*. Equation (A.12) estimates β .

$$
\hat{\beta} = \frac{\sum_{i=1}^{K} (Y_i - \overline{Y})(X_i - \overline{X})}{\sum_{i=1}^{K} (X_i - \overline{X})^2}
$$
\n(A.12)

The point estimator of μ_Y is estimated by equation (A.13).

$$
\hat{\mu}_{Y}\left(\hat{\beta}\right) = \frac{1}{K} \sum_{i=1}^{K} Y_{i}\left(\hat{\beta}\right)
$$
\n(A.13)

Using regression theory an interval estimate for μ_Y can be obtained. By making the assumption of joint normality for *X* and *Y*, the conditional distribution of *Y* given *X* will be normal by equation $(A.14)$

$$
Y \mid X = x \sim \mathcal{N}(\mu_Y + \beta(x - \mu_X), \sigma_E^2)
$$
 (A.14)

where

$$
\sigma_{\varepsilon}^{2} = \sigma_{\varepsilon}^{2} (1 - \rho_{XY}^{2})
$$
 (A.15)

and

$$
\sigma_Y^2 = \text{Var}(Y) \tag{A.16}
$$

Since the values of the control variable *X* and it mean μ_X are known, then it can be seen that the conditional mean of *Y* given *X* has two terms. The terms are broken into the parameter to be estimated μ_Y and a correction term. To get the μ_Y term, the corrections need to be subtracted out as in equation (A.10). From equation (A.14), equation (A.17) can be formed.

$$
Y_i = \mu_Y + \beta (X_i - \mu_X) + \varepsilon_i, 1 \le i \le K
$$
\n(A.17)

where ε_i are the residuals and are of the form in equation (A.18).

$$
\varepsilon_i \sim \mathcal{N}\big(0, \sigma_\varepsilon^2\big) \tag{A.18}
$$

Since the values of μ_Y and β are unknown, the method of least squares can be applied to solve for them. μ_Y will be the intercept and is normally distributed as in equation (A.19).

$$
\hat{\mu}_{Y}\left(\hat{\beta}\right)_{i} \sim N\left(\mu_{Y}, \sigma_{\varepsilon}^{2} s_{11}\right) \tag{A.19}
$$

The value of s_{11} in equation (A.19) is the upper left hand entry in the matrix $(D^TD)^{-1}$ where D is of the form in equation $(A.20)$.

$$
D = \begin{bmatrix} 1 & X_1 - \mu_X \\ 1 & X_2 - \mu_X \\ 1 & X_3 - \mu_X \\ \vdots & \vdots \\ 1 & X_K - \mu_X \end{bmatrix}
$$
 (A.20)

To generate a confidence interval about $\hat{\mu}_{y}(\hat{\beta})$, σ_{ε}^{2} must be estimated. Since σ_{ε}^{2} represents the variability in *Y* given *X*, the formula for the residual mean square error is used.

$$
\sigma_{\varepsilon}^{2} = \frac{\sum_{i=1}^{K} (Y_{i} - \hat{Y}_{i})^{2}}{K - 2}
$$
 (A.21)

where

$$
\hat{Y}_i\left(\hat{\beta}\right) = \hat{\mu}_Y\left(\hat{\beta}\right) + \hat{\beta}\left(X_i - \mu_X\right), 1 \le i \le K
$$
\n(A.22)

It can be seen, from the above equations, that

$$
\frac{\hat{\mu}_Y(\hat{\beta}) - \mu_Y}{\left[\frac{\sigma_\varepsilon^2 s_{11}}{K - 2}\right]^{\frac{1}{2}}} \sim t_{K-2}
$$
\n(A.23)

has a Student-t distribution with *K*-2 degrees of freedom. Therefore, the confidence interval for μ_Y is given by

$$
\hat{\mu}_{Y} \pm t_{K-2,\left(1-\frac{\alpha}{2}\right)}\sqrt{\sigma_{\varepsilon}^{2} s_{11}} \tag{A.24}
$$

Now in simulations there are possibly more than one control for a response.

Therefore, equation (A.17) is modified to be

$$
Y_i = \mu_Y + \sum_{j=1}^{Q} \beta_j \Big(X_{ji} - \mu_{X_j} \Big) + \varepsilon_i, 1 \le i \le K
$$
 (A.25)

where,

$$
Q \le K - 1 \tag{A.26}
$$

Therefore, equation (A.21) is reformed into

$$
\sigma_{\varepsilon}^{2} = \frac{\sum_{i=1}^{K} \left(Y_{i} - \left(\hat{\mu}_{Y} + \sum_{j=1}^{Q} \hat{\beta}_{j} \left(x_{j} - \mu_{X_{j}} \right) \right) \right)^{2}}{K - Q - 1}
$$
(A.27)

and s_{11} is the upper left hand entry in the matrix $(D^TD)^{-1}$ where *D* is of the form

$$
D = \begin{bmatrix} 1 & X_{11} - \mu_{1X} & X_{21} - \mu_{2X} & \cdots & X_{Q1} - \mu_{QX} \\ 1 & X_{12} - \mu_{1X} & X_{21} - \mu_{2X} & \cdots & X_{Q1} - \mu_{QX} \\ 1 & X_{13} - \mu_{1X} & X_{21} - \mu_{2X} & \cdots & X_{Q1} - \mu_{QX} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & X_{1K} - \mu_{1X} & X_{21} - \mu_{2X} & \cdots & X_{Q1} - \mu_{QX} \end{bmatrix}
$$
(A.28)

Then the $100(1-\alpha)$ % confidence interval in equation (A.24) becomes

$$
\hat{\mu}_{Y} \pm t_{K-Q^{-1}\left(1-\frac{\alpha}{2}\right)} \sqrt{\sigma_{\varepsilon}^{2} s_{11}} \tag{A.29}
$$

Appendix B: Calculations, Paired Difference Test of Hypothesis for $\mu_d = (\mu_1 - \mu_2)$

Using FX-99 Scenarios with Unconstrained and Constrained Manpower

C15 (Overall Sorties per Aircraft per Day):

Absolute value of t (19.32) is greater than t_{0.025} (1.96), therefore we can reject the null

hypothesis that the two population means are equal. Furthermore, the confidence interval

does not include zero. We can infer that the mean for the unconstrained scenario exceeds the mean for the constrained scenario.

C24 (Mission Capable Rate):

Absolute value of t (11.11) is greater than $t_{0.025}$ (1.96), therefore we can reject the null

hypothesis that the two population means are equal. Furthermore, the confidence interval

does not include zero. We can infer that the mean for the unconstrained scenario exceeds the mean for the constrained scenario.

Appendix C: Merged Ouput Report Example for FX99 Model, Statistics C10 – C25

Appendix E: Screen Shot, Production Run Settings in LCOM Graphical User

Interface

Appendix F: Potential Control Observations for 21 Replications

			6		10		13	14	15	16	17	18	19	20	$CI5$ $C25$
	-1E-04 2E-04 1E-04 -3E-04 -0.001 7E-04 -5E-04 4E-04 -0.001 -8E-04 -4E-04 0.002						0.004 -2359		1115	2398	1145		2310 - 10801 654.2 1.72 67.91		
	5E-05 3E-04 -2E-04 5E-04 5E-04 1E-04 -2E-04 -6E-04 -0.001 -3E-04 4E-05 -5E-04 -0.004 -2279								1321	2183		-3965 956.6	6832	2529 1.68 68.76	
	-1E-04l-2E-05l-6E-04l-2E-05l 7E-04 l-1E-05l 4E-04 l-2E-04l 4E-04 l-0.002 l-0.002 l-0.004l 1999									2846 - 387.3 2464		1172	20016	503	1.72 68.14
	2E-04-2E-04-5E-05 1E-04-4E-04-3E-04-7E-04-0.001 7E-04-0.001 0.003 5E-04-0.007										3256 -3270 -3715 -2111 -603.2			26 902.4 1.56 69.53	
	2E-05 -2E-05 1E-04 -6E-05 8E-04 -3E-04 -3E-04 -3E-04 -2E-04 -0.002 -5E-06 -0.009									3750 -2246 -1623	169.1	-134	15635	$-218011.48170.55$	
6	4E-05-2E-04 2E-04 5E-04 -4E-04 -3E-05 -7E-04 -9E-04 3E-05 -2E-04 -0.003 -4E-04 3E-04 -3744 -69.94									6246	2326	1638	-6075		795 1.84 67.34
	-4E-05 -3E-04 2E-04 1E-04 -7E-04 -1E-06 -3E-04 -7E-04 0.001 -0.002 0.002 0.004 -0.005 -2818 -2280 2700 -819.8 -2517 -15022													$-51301.74166.6$	
	-3E-05 -3E-04 -2E-04 -2E-05 2E-04 -2E-04 -4E-04 -5E-05 0.001 3E-04 3E-05 -4E-04 -0.007 758.7 -4307 -4024 -4001 -490.8												775		1489 1.7 68.29
9	-3E-07 2E-04 1E-04 -0.001 3E-04 7E-04 5E-04 0.001 -0.002 7E-05 5E-04 -0.002 -9E-04 -1542 -2593 0.125 -312.6 2258 -1739 2845 1.62 68.86														
10	6E-05 -1E-04 -1E-04 -4E-05 -3E-04 -9E-05 1E-03 -1E-04 -0.003 -5E-04 0.005 5E-04 0.003 -1209 276.8 -5195										3565		2630-11637 -1692 1.84 66.97		
	5E-05 3E-05 -4E-04 -2E-04 -6E-06 -5E-04 2E-04 -0.001 -0.002 -0.002 0.001 -0.001 -0.015							2827		1570 252.7		6018-893.2	-2407		499511.76166.78
12	1 E-04 2E-04 2E-04 - 9E-06 - 1E-04 - 0.001 - 4E-04 - 9E-04 - 3E-04 - 1E-04 - 0.002 - 0.002 - 0.007									-306 -2495 525.3			4945 - 1336 - 25419		1993 1.6 68.97
13	-3E-05-6E-05 2E-04 6E-04 4E-04 0.001 4E-04 5E-04 -6E-04 2E-04 -0.002 8E-04 0.006 -3352 -89.86									4028			1338 2224 - 11490 - 1234 1.72 69.2		
14	-1E-06 1E-04 -4E-04 -3E-04 -2E-04 5E-04 1E-05 -9E-04 -4E-04 3E-05 9E-06 0.002 -0.017							879	1141		3193-725.1	-1623	12493		1567 1.8 66.27
15	-2E-05-1E-04 2E-04-6E-04 3E-04-2E-04 0.001 0.001 -0.002 5E-04 0.002 0.001 4E-04 636.4									2415 330.7	3201	-2507	-3399		1265 1.74 69.19
16 I	-3E-06 2E-04 -3E-04 2E-05 -2E-04 4E-04 -5E-04 -2E-04 -0.001 2E-04 -0.002 -0.001 -0.009 -3109								1030		3306-539.8	4621	3171	21631.78166.1	
17	-9E-05 3E-04 -6E-04 -2E-04 4E-04 8E-04 4E-05 5E-05 7E-04 -0.001 -0.001 0.002						-0.005 1726		5805	-7200 2420		1676	-8223	2433 1.76 67.91	
18	-8E-05 1E-04 4E-05 -1E-04 5E-04 4E-04 -6E-04 6E-05 4E-04 -6E-04 0.002 -0.002						$0.007 - 649.2$				3843 58.13 - 9082	2369	106831	$-135211.58169.26$	
19	-6E-06 8E-05 -1E-04 2E-04 -9E-04 4E-04 -2E-04 6E-04 0.001 -0.001 -7E-04 9E-04						0.011 -5158		1810	3210	2483	-1620	-1957	100811.9273.81	
20	3E-05-5E-05-2E-05 2E-04 2E-04-4E-04-3E-04 2E-04 1E-03-3E-04-8E-04 1E-03 0.009								$2585 - 1627$		1179 - 2430 129.9			-4812 553.2 1.9 65.04	
21	1683.1 = 169.6 = 169.6 = 169.6 = 169.6 = 169.6 = 169.6 = 169.6 = 169.6 = 169.6 = 169.6 = 169.6 = 169.6 = 169.6 = 169.6 = 169.6 = 169.6 = 169.6 = 178.6 = 178.6 = 178.6 = 178.6 = 178.6 = 178.6 = 178.6 = 178.6 = 178.6 = 178.6														169.48

Appendix G: 20 X 21 Matrix of Potential Controls and Output Variables

Base -- C15 CRN3 -- C15

Base -- C24				CRN3 -- C24		
Replication	unconstrained constrained		difference	Replication	unconstrained constrained	
1	68.26	61.74	6.52	1	67.28	61.29
$\overline{2}$	68.81	61.66	7.15	$\overline{2}$	68.18	65.17
$\overline{\mathbf{3}}$	66.37	65.16	1.21	\mathfrak{Z}	68.84	61.69
$\overline{4}$	67.19	61.81	5.38	$\overline{4}$	67.22	64.66
5	67.01	61.61	5.4	5	66.73	65.26
6	66.74	65.04	1.7	6	66.93	60.66
$\overline{7}$	65.29	60.77	4.52	τ	67	61.25
8	68.62	61.37	7.25	$8\,$	67.13	61.54
9	67.35	61.74	5.61	9	67.65	62.72
10	65.71	62.8	2.91	10	67.24	61.86
11	67.28	62.33	4.95	11	67.09	62.24
12	66.96	65.11	1.85	12	66.29	61.33
13	67.35	63.11	4.24	13	67.08	64.72
14	65.72	64.07	1.65	14	67.19	65.08
15	67.61	65.92	1.69	15	69.28	65.33
16	67.17	61.87	5.3	16	67.49	61.87
17	66.29	62.6	3.69	17	67.52	61.66
18	64.67	62.8	1.87	18	66.93	61.25
19	66.54	65.45	1.09	19	68.81	64.98
20	67.26	59.93	7.33	20	68.8	61.34
21	66.95	60.74	6.21	21	67.5	62.81
22	67.54	61.44	6.1	22	66.9	65.65
23	66.85	61.86	4.99	23	68.03	62.77
24	67.28	64.29	2.99	24	65.34	64.9
25	67.18	63.03	4.15	25	67.43	62.07
26	66.43	61.74	4.69	26	67.2	61.13
27	66.97	61.8	5.17	$27\,$	68.89	62.19
28	66.4	65.3	1.1	28	65.86	62.89
29	66.6	62.45	4.15	29	67.3	65.69
30	66.79	65.74	1.05	30	67.28	61.7
Mean	66.90633333 62.8426667 4.06366667			Mean		67.4136667 62.9233333
St Dev	0.875090077 1.68354702 2.03690855			St Dev		0.88028398 1.68821949
			11.1139593			
95% CI:	3.3470195694.78031376			95% CI:		3.8258381 5.15482857
	CI halfwidth: 0.716647098			CI halfwidth:	0.66449523	

Appendix I: Results for Antithetic Variates

C15 – Single Control:

C15 – Multiple Controls:

Ybar 1.7219048

C15 – Multiple Controls (cont.)

-0.001905	$-2.47E-06$	$-5.09E-07$	-3.317521	1.501129	0.019568	1.683E-06	7.14481E-08	3033514.524	621089.9242	105.5335545
-0.041905	2.09E-05	1.15E-05	-69.60758	41.63955	78.98196	2.479E-07	7.531E-08	2759219.804	987382.3112	3552458.688
-0.001905	1.27E-06	$-2.32E-06$	4.984787	4.797904	-0.268362	4.435E-07	1.48437E-06	6848758.755	6344854.958	19849.98407
-0.161905	$-5.89E-05$	-0.00013	627.2129	-582.3882	41.85223	1.325E-07	6.4558E-07	15007580.61	12939152.68	66821.77649
-0.241905	0.000203	7.34E-05	1056.521	-622.3533	-683.1308	7.062E-07	9.20425E-08	19075128.46	6618890.679	7974782.779
0.118095	5.16E-05	$-4.16E-05$	369.2509	46.89924	-17.83824	1.907E-07	1.24214E-07	9776387.309	157712.7719	22815.97394
0.018095	1.33E-05	2.16E-05	39.82036	47.18191	104.4864	5.376E-07	1.42174E-06	4842636.043	6798646.227	33341960.55
-0.021905	3.96E-06	1.77E-05	30.14747	-101.5169	18.50394	3.261E-08	6.54831E-07	1894195.444	21478303.36	713593.9117
-0.101905	3.57E-05	6.25E-05	-94.16964	-297.5895	224.2977	1.224E-07	3.76165E-07	853950.793	8527981.923	4844632.259
0.118095	3.42E-05	$-2.4E-06$	69.8812	5.948581	275.9043	8.367E-08	4.13645E-10	350151.4491	2537.242706	5458233.952
0.038095	1.24E-07	3.64E-05	-131.2326	-47.34395	-165.7385	1.062E-11	9.15004E-07	11867039.46	1544498.792	18928029.63
-0.121905	$-1.77E-05$	4.92E-05	37.98254	-343.9774	164.4519	2.115E-08	1.62603E-07	97079.28468	7961930.123	1819853.126
-0.001905	8.36E-07	1.41E-06	-5.207504	-0.794388	-3.576533	1.927E-07	5.48862E-07	7474426.93	173933.8038	3525681.723
0.078095	1.47E-05	$-4.5E-05$	-116.8797	-63.53621	-72.07986	3.529E-08	3.32005E-07	2239900.741	661901.807	851880.6209
0.018095	$-5E-06$	$-1.69E-06$	0.339896	-37.78117	-11.2336	7.646E-08	8.74132E-09	352.8286457	4359355.314	385397.9401
0.058095	1.05E-05	$-4.17E-05$	144.7097	-40.83826	-88.27161	3.293E-08	5.16064E-07	6204603.255	494143.799	2308665.329
0.038095	$-1.64E-05$	$2.04E-05$	-89.26848	-208.6622	-68.15683	1.864E-07	2.85376E-07	5491043.879	30001722.78	3200939.108
-0.141905	6.74E-05	$-9.18E - 06$	-4.483354	498.9755	-283.211	2.259E-07	4.18618E-09	998.1874018	12364154.37	3983146.627
0.198095	0.00018	0.000116	899.5053	-293.702	-72.1491	8.284E-07	3.4138E-07	20618609.2	2198193.022	132651.9763
0.178095	$-3.07E-05$	$-3.73E - 05$	-570.4432	348.1026	16.15524	2.974E-08	4.37632E-08	10259358.36	3820409.043	8228.533217
-0.021905	4.8E-06	1.11E-05	-65.10601	51.44678	-26.8781	4.795E-08	2.56791E-07	8834157.278	5516191.125	1505636.337
Beta(hat)5	Beta(hat)10		Beta(hat)14	Beta(hat)15	Beta(hat)20					
87.32128772		13.04777865	1.54923E-05	$-1.1933E-05$	$-6.1296E-06$					

C15 – Multiple Controls (cont.)

X

X(bar) 0.011669755 0.025545512 50618.0959 49673.31077 49356.57012

C15 – Multiple Controls (cont.)

Corr. Term	5.	10	14	15	20	sum	Y hat b	Y - Yhatb \sim 2	sigma γ 2 e
	-0.113556	-0.010605	-0.036551	-0.013309	-0.00401	-0.1780315	1.548834599	0.029297594	0.039349927
	0.04321	-0.003537	-0.035302	-0.015762	-0.0155	-0.0268915	1.699974614	0.000398985	
	0.057882	-0.023014	0.030976	-0.033964	-0.003083	0.0287962	1.755662369	0.001271805	
	-0.032053	-0.017601	0.050449	0.039021	-0.005532	0.0342844	1.76115053	0.040461536	
5	0.07311	-0.003159	0.058095	0.026797	0.013363	0.1682054	1.895071501	0.172284351	
61	-0.038402	-0.002519	-0.058008	0.000835	-0.004873	-0.102967	1.623899122	0.046699589	
	-0.064294	-0.022675	-0.04366	0.027211	0.031447	-0.0719718	1.654894385	0.007242966	
8	0.015499	0.003441	0.011754	0.051401	-0.009125	0.0729698	1.799835897	0.009967206	
9	0.030283	0.000885	-0.023884	0.030944	-0.017439	0.020789	1.74765511	0.016295827	
10	-0.025528	-0.006852	-0.018735	-0.003303	0.010373	-0.0440458	1.682820379	0.024705433	
	-0.000554	-0.019599	0.043801	-0.018735	-0.030615	-0.025702	1.701164117	0.003461661	
12	-0.012968	-0.001856	-0.004741	0.029768	-0.012216	-0.0020135	1.724852624	0.015588178	
13	0.038064	0.002549	-0.051923	0.001072	0.007562	-0.0026758	1.724190379	1.75593E-05	
14	-0.016674	0.0004	0.013618	-0.013613	-0.009604	-0.0258729	1.700993253	0.009802336	
15	0.023875	-0.005898	-0.009859	-0.02882	-0.007752	-0.0284543	1.698411894	0.001729571	
16	-0.016115	0.002255	-0.048158	-0.012293	-0.013261	-0.0875715	1.639294642	0.019797998	
17	0.03743	-0.014088	0.026735	-0.069269	-0.014914	-0.0341054	1.692760706	0.004521123	
18	0.041234	-0.007962	-0.010057	-0.045866	0.008286	-0.0143644	1.712501703	0.017556701	
19	-0.079745	-0.014741	-0.079915	-0.021597	-0.00618	-0.2021783	1.524687806	0.15627173	
20	0.014788	-0.004388	0.040054	0.019421	-0.003391	0.0664836	1.793349726	0.011374281	
21	0.018851	-0.000506	-0.055615	-0.031932	0.003574	-0.0656278	1.661238312	0.001502468	

C15 – Multiple Controls (cont.)

Dtranspose *D

Bibliography

- Aeronautical Systems Center. *ASC LCOM 2.6 Users Manual.* Wright-Patterson AFB OH: ASC/ENM, February 2004.
- Banks, Jerry et al. *Discrete-Event System Simulation* (4th Edition). New Jersey: Pearson Education, Inc. 2005.
- Bednar, Earl M. *Feasibility Study of Variance Reduction in the Thunder Campaign-Level Model.* MS Thesis, AFIT/GOR/ENS/05-01. Graduate School of Engineering and Management, Air Force Institute of Technology (AFIT), Wright-Patterson AFB OH, March 2005.
- Boughton, Gregory. *Legacy Air Force Simulation Model Enhanced by Optimizer.* Lecture, 74th MORS Symposium, ASC/ENMS, June 2006.
- Boyle, Edward. *LCOM Explained: Interim Technical Paper for Period May 1990 June 1990.* AFHRL-TO-90-58. Brooks AFB TX: Air Force Systems Command, July 1990 (AD-A224497).
- CACI. "CACI Simscript." Full description of software. [http://www.caci.com/asl/simscript_description.shtml. 28 December 2004.](http://www.caci.com/asl/simscript_description.shtml. 28 December 2004)
- Carrico, Todd and Patricia K. Clark. *Integrated Model Development Environment (IMDE) Support for Air Force Logistics.* Human Resources Directorate, Logistics Research Division, Wright-Patterson AFB OH, July 1996.
- Dawson, Kevin. *LCOM Process Reengineering.* Air Force Logistics Management Agency, Maxwell AFB, AL, March 2006.
- Dierker, Gregory J. LCOM Integrated Product Team, Systems Supportability Analysis Branch, Modeling, Simulation, and Analysis Division, Engineering Directorate, Aeronautical Systems Center, Wright-Patterson AFB OH. Personal Interviews. March 2006.
- Elhefny, Mohamed Refat *Variance Reduction Techniques With Applications.* Air Force Institute of Technology (AU), Wright-Patterson AFB, OH December 1983 (ADA138074).
- Erdman, Francis J. Aeronautical Systems Center LCOM Group Lead, Systems Supportability Analysis Branch, Modeling, Simulation, and Analysis, Division, Engineering Directorate, Aeronautical Systems Center, Wright-Patterson AFB OH. Personal Interviews. April – November 2006.
- Garcia, Robert and Joseph P. Racher Jr. *An Investigation Into a Methodology to Incorporate Skill Level Effects Into the Logistics Composite Model.* MS Thesis, AFIT/LSSR 29-81. School of Systems and Logistics, Air Force Institute of Technology (AFIT), Wright-Patterson AFB OH, June 1981.
- Kelton, David W. et al. *Simulation With Arena* (3rd Edition). New York: Marcel Dekker, Inc., 2004.
- L'Ecuyer, P., R. Simard, E. J. Chen, andW. D. Kelton. 2002. *An object-oriented randomnumber package with many long streams and substreams.* Operations Research 50 (6): 1073–1075.
- Law, Averill M. and David W. Kelton. *Simulation Modeling and Analysis* (Third Edition). New York: McGraw-Hill, 2000.
- Lehmer, D. H. *Random number generation on the BRL high-speed computing machines, by M.L. Juncosa*. Math Review. 15 (1954), 559.
- McClave, James T. et al. *Statistics for Business and Economics* (9th Edition). New Jersey: Prentice Hall, 2005.
- Minitab Release 14 Statistical Software. Minitab Help. LEAD Technologies, Inc, 2005.
- Pettingill, Kirk B. *An Analysis of the Efficacy of the Logistics Composite Model in Estimating Maintenance Manpower Productive Capacity.* MS Thesis, AFIT/GLM/ENS/03-11. Graduate School of Engineering and Management, Air Force Institute of Technology (AFIT), Wright-Patterson AFB OH, March 2005.
- Russell, Edward C. *Building Simulation Models With Simscript II.5.* California: CACI Products Co., 2000.

Vita

 Captain George Cole, III was born in Omaha, Nebraska. He graduated from Bossier High School in Bossier City, Louisiana in 1998. He spent the next four years at the U.S. Air Force Academy, graduating with a Bachelor of Science in Management in May of 2002 and earning a commission in the U.S. Air Force upon graduation.

Capt Cole spent 3 years as a C-17 aircraft maintenance officer in the $437th$ Aircraft Maintenance Squadron at Charleston Air Force Base, South Carolina. In 2004 he deployed to Ashgabat, Turkmenistan in support of Operation Enduring Freedom. In 2005 Capt Cole was selected to attend the Air Force Institute of Technology's Graduate of Logistics Management Program. Upon graduation, Capt Cole will be assigned to Air Mobility Command Headquarters at Scott Air Force Base, Illinois.

e-mail: alan.johnson@afit.edu