In a given fiscal year, the United States Marine Corps accesses approximately 30,000 enlisted personnel into its ranks. This labor supply of recruits is classified into various Military Occupational Specialties (MOSs) according to the forecasted requirement for new personnel into a particular MOS. The Classification Plan is the primary initial training input into the Training Input Plan, which allocates all training resources for Training and Education Command. The current Classification Model is based on a steady-state Markov Model that estimates the first-term inventory of each initial training MOS inventory of personnel. A performance comparison was made against a transient Markov Model that solves for an optimal classification plan over the course of a four-year planning horizon. First, the validity of the steady-state assumption is tested and found to produce a variance of annual targets for each MOS throughout the Future Years Defense Plan that is prohibitively high. Next, a comparison of each models’ ability to forecast annual attrition by MOS between the years 2001 and 2011 is tested. Results indicate that the transient model produced a more accurate forecast for 5,321 out of 7,379 design points (approximately 72% of the observations). The transient model achieved a Mean Absolute Proportional Error that was on average 14 percentage points smaller than that of the steady-state model. In over 25% of the cases, this difference exceeded 20 percentage points. Based upon this improved performance, it is recommended the Marine Corps adopt the enhanced transient Markov Model as the foundation for forecasting its annual Enlisted Classification Plan.
FORECASTING THE MARINE CORPS’ ENLISTED CLASSIFICATION PLAN:
ASSESSMENT OF AN ALTERNATIVE MODEL

Erik P. Hovey
Major, United States Marine Corps
B.S. Business, University of Idaho, 1998

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
March 2012

Author: Erik P. Hovey

Approved by: Chad W. Seagren
Thesis Advisor

Bill Hatch
Second Reader

Bill Gates
Dean, Graduate School of Business and Public Policy
ABSTRACT

In a given fiscal year, the United States Marine Corps accesses approximately 30,000 enlisted personnel into its ranks. This labor supply of recruits is classified into various Military Occupational Specialties (MOSs) according to the forecasted requirement for new personnel into a particular MOS. The Classification Plan is the primary initial training input into the Training Input Plan, which allocates all training resources for Training and Education Command. The current Classification Model is based on a steady-state Markov Model that estimates the first-term inventory of each initial training MOS inventory of personnel. A performance comparison was made against a transient Markov Model that solves for an optimal classification plan over the course of a four-year planning horizon. First, the validity of the steady-state assumption is tested and found to produce a variance of annual targets for each MOS throughout the Future Years Defense Plan that is prohibitively high. Next, a comparison of each models’ ability to forecast annual attrition by MOS between the years 2001 and 2011 is tested. Results indicate that the transient model produced a more accurate forecast for 5,321 out of 7,379 design points (approximately 72% of the observations). The transient model achieved a Mean Absolute Proportional Error that was on average 14 percentage points smaller than that of the steady-state model. In over 25% of the cases, this difference exceeded 20 percentage points. Based upon this improved performance, it is recommended the Marine Corps adopt the enhanced transient Markov Model as the foundation for forecasting its annual Enlisted Classification Plan.
# TABLE OF CONTENTS

I. INTRODUCTION ........................................................................................................1  
   A. BACKGROUND ..................................................................................................1  
   B. OBJECTIVES ..................................................................................................1  
   C. RESEARCH QUESTION AND METHODOLOGY ...................................2 
   D. SCOPE AND LIMITATIONS .......................................................................2 
   E. ORGANIZATION OF THESIS .....................................................................3  

II. LITERATURE REVIEW ...........................................................................................5  
   A. INTRODUCTION OF PREVIOUS STUDIES .............................................5  
   B. PAST NAVAL POSTGRADUATE SCHOOL THESIS ..............................5  
      1. Forecasting Enlisted Attrition in the United States Marine Corps by Grade and Years of Service ....................................................5  
      2. Analysis of the U.S. Marine Corps’ Steady State Markov Model for Forecasting Annual First-Term Enlisted Classification Requirements .........................................6  
      3. Forecasting Marine Corps Enlisted Attrition Through Parametric Modeling ...............................................................7  
   C. CENTER FOR NAVAL ANALYSIS REPORT ...........................................8  
      1. Endstrength: Forecasting Marine Corps Losses Final Report .........8  

III. MODEL OVERVIEW .................................................................................................9  
   A. INTRODUCTION ............................................................................................9  
   B. PRINCIPLES AND ASSUMPTIONS IN THE MARKOV CHAIN ..........9  
      1. States .....................................................................................................9  
      2. Markovian Property ..........................................................................9  
      3. Stationary Transition Probabilities ....................................................9  
   C. NOTATION ....................................................................................................10  
   D. CURRENT STEADY-STATE MODEL ...........................................................10  
   E. PROPOSED ENHANCED CLASSIFICATION PLAN MODEL ..........12  
   F. SUMMARY ....................................................................................................13  

IV. ANALYSIS OF MODELS ........................................................................................15  
   A. OVERVIEW OF THE ANALYSIS..............................................................15  
      1. Purpose of Analysis ..........................................................................15  
      2. Method of Analysis ...........................................................................15  
      3. Endstate of Analysis .........................................................................15  
   B. STEADY-STATE ASSUMPTION .................................................................16  
   C. INPUTS TO THE ANALYSIS .................................................................17  
      1. Description of GAR Data ..................................................................17  
      2. Transition Probabilities ..................................................................17  
      3. Target Population by MOS ...............................................................18  
      4. Distributions of Accessions ...............................................................18  
   D. TESTING OF THE STEADY-STATE ASSUMPTION ............................19  
      1. Coefficient of Variation ..................................................................20
LIST OF FIGURES

Figure 1. Transition Probabilities Between the Various States of Potential YCS ........11
Figure 2. NPS Enlisted Accessions ...........................................................................16
Figure 3. Endstrength ...............................................................................................17
Figure 4. Rifleman MOS ...........................................................................................20
Figure 5. Summary of distribution of Coefficient of Variance of GAR targets across
          MOSs for FY04–11 ........................................................................................21
Figure 6. CV Distribution Comparison .....................................................................22
Figure 7. Summary of Distribution of Variance in Time Lags Across MOSs ..........23
Figure 8. Example GAR Requirements Over the FYDP ........................................24
Figure 9. Models’ MAPE Distribution ......................................................................28
LIST OF ACRONYMS AND ABBREVIATIONS

CV  Coefficient of Variation  
FY  Fiscal Year  
FYDP  Future Years’ Defense Plan  
GAR  Grade Adjusted Recapitulation  
KISS  Keep it Simple, Stupid  
MAGTF  Marine Air-Ground Task Force  
MAPE  Mean Absolute Percent Error  
MCT  Marine Combat Training  
MOE  Measure of Effectiveness  
MOS  Military Occupational Specialty  
NEAS  Non-End of Active Service  
PMOS  Primary Military Occupational Specialty  
TIG  Time in Grade  
TIS  Time in Service  
TOE  Term of Enlistment  
TTT  Time to Train  
YCS  Years Completed Service
THIS PAGE INTENTIONALLY LEFT BLANK
ACKNOWLEDGMENTS

I would like to thank Chad Seagren for his unwavering support throughout the process of writing this thesis. The Marine Corps and Marine officers at NPS are much better off because of his service and dedication to teaching…General Lejeune would be proud. In addition, I would like to acknowledge the support provided to the Marine officers of the GSBPP by Commander Bill Hatch, USN (Ret). His desire to set us up for success in our future billets is appreciated and speaks to his character of leadership – one that continues to develop subordinates through education and training.

My family has continued to provide invaluable support to me over the past two years while at NPS. My wife, Elaine, and our beautiful children continue to serve as the source of my motivation, dedication, and inspiration.
THIS PAGE INTENTIONALLY LEFT BLANK
I. INTRODUCTION

A. BACKGROUND

As an organization of over 200,000 personnel, the Marine Corps recruits and brings into its workforce approximately 30,000 new personnel on an annual basis. This workforce is distributed across over 200 Military Occupational Specialties (MOSs). This distribution is a function of the total requirement for the MOS by rank, current MOS strength, new accessions to the MOS, and attrition from the MOS. Therefore, this labor supply of recruits is assigned into various MOSs due to the forecasted demand for new personnel into that particular MOS. The goal of this forecast is to ensure an enlisted force structure exists to meet manpower resource requirements.

The demand for new accessions per Fiscal Year (FY) is tied to forecasting attrition behavior. The plan that determines the number of Marines to assign to each MOS is referred to as the Enlisted Classification Plan. This classifications plan then serves as an input to a number of successive plans. From the classification plan, the Marine Corps develops requirements for its Recruiting Command’s accession and initial training (recruit training) requirements, and its MOS training requirements for its multiple MOS-producing schools.

B. OBJECTIVES

Because of the sheer magnitude and impact that a classification plan has to the subsequent planning and the long-term health of an MOS, the importance of an accurate classification plan cannot be overemphasized. Further, in this era of fiscal austerity and future downsizing, the efficiency and precision of any forecasted number is worthy of analysis and scrutiny.

Given that desired endstrength is a function of beginning (current) endstrength + (plus) accessions – (minus) attritions, the number of personnel attritions must be forecasted to the highest degree of accuracy possible. In this formula, the desired and current personnel endstrengths are known quantities. The accessions required in a future fiscal year are predicated upon the attritions. Since the actual level of attrition is an
unknown variable, it must be forecasted. The current model used to predict attrition is based upon a steady-state inventory of personnel. However, it is believed that an enhanced, transient model could not only reduce misclassifications of MOSs, but also better account for dynamic changes to desired endstrength.

C. RESEARCH QUESTION AND METHODOLOGY

The Marine Corps currently uses a steady-state Markov model to forecast the first-term inventory of each unique initial training MOS and subsequent accession requirement. As previously stated, accurate and precise forecasting is always important; however, the degree of required accuracy and resulting efficiency required of manpower planners will be ever increasing in the Marine Corps’ drawdown and future Department of Defense’s fiscally constrained environment. The purpose of this thesis is to assess the validity of a Classification Plan derived from a transient Markov Model.

Primary Research Question: Would a transient Markov Model serve as a more accurate forecasting tool to support the development of the Marine Corps’ Enlisted Classification Plan? If more accurate, what is the relative magnitude in the difference in accuracy (i.e., how much better of a predictor is the transient model)?

Through simulation, a comparison of the current attrition model and a proposed enhanced attrition model is conducted. This comparison includes a review of each models’ predicted attrition behavior.

D. SCOPE AND LIMITATIONS

The enlisted manpower and personnel analysts and planners make career force decisions (promotion, retention, training, etc.) based upon existing and forecasted personnel resources by specific MOS. These career force decisions are driven from the requirements by grade and specific MOS as outlined in the Grade Adjusted Recapitulation (GAR). As previously noted, the variable of attrition must be forecasted as a basis for these career force decisions. Therefore, the more accurately forecasted the level of attrition is as an input to this decision system, the less variability the system as a whole. Further, a forecast model that supports the ability to account for changes in
desired endstrength would not only prove useful in the short-term and future draw down of personnel endstrength, but also be flexible enough to support any future changes to desired endstrength.

Although Marine Corps’ future endstrength requirements are not conclusively known, it is generally accepted that its endstrength will be no greater than 182,200 in the upcoming years (Amos, 2012). This roughly 10% decrease in overall personnel endstrength and subsequent reshaping of its force will have an effect across all MOSs. Additionally, technological advances, equipment changes, and unforeseen operational requirements contribute to the management of manpower and the subsequent personnel effects inherent in the Marine Corps’ manpower and personnel system.

E. ORGANIZATION OF THESIS

Chapter II of this thesis consists of a literature review and commentary on those works that have previously looked at (either directly or indirectly) elements of the Marine Corps’ Enlisted Classification Plan. Chapter III presents an overview of the current steady state Markov Model that is used by HQMC planners and introduces an enhanced Markov Model. Chapter IV presents an analysis and comparison of the current and transient models. Lastly, Chapter V presents a summary, conclusions, and recommendations for the developers of the Marine Corps’ Enlisted Classification Plan.
II. LITERATURE REVIEW

A. INTRODUCTION OF PREVIOUS STUDIES

Previous studies and papers discuss attrition and retention in the military as a binary (yes or no) outcome. This literature typically looks at various demographic and other observable variables (test scores, physical fitness scores, etc.) and attempts to predict behavior through multivariate regression analysis.

B. PAST NAVAL POSTGRADUATE SCHOOL THESIS

1. Forecasting Enlisted Attrition in the United States Marine Corps by Grade and Years of Service

In his thesis, Tamayo (2011) set out to analyze enlisted attrition behavior through various time series forecasting techniques. Comparing the results of both moving average and weighted moving average models, Tamayo then applied the two relative measures of effectiveness (MOE). Specifically, Tamayo analyzed the Mean Square Error and the Mean Absolute Percent Error of the various models to serve as his MOEs. Tamayo’s modeling determined that the optimal weights associated with using a weighted-moving average model was nearly equivalent to that of a one-year moving average model.

To his credit, Tamayo included a sizable time horizon of study in his thesis (FY1987–2009). This timeline was significantly long enough to capture data prior to the personnel build up during Operation DESERT SHIELD/DESERT STORM and the accompanying personnel reductions, and the recent build up in endstrength in support of Operations ENDURING FREEDOM and IRAQI FREEDOM. Additionally, Tamayo acknowledges that the benefit and simplicity of using a one-year moving average model and its ability to support manpower planners’ ability to quickly estimate forecasts by grade and MOS.

To be sure, the Marine Corps organizationally supports and endorses the KISS (keep it simple, stupid) Principle. However, there is a cost with too much simplicity. Tamayo’s findings must be put into further context to gain a deeper appreciation for this cost-benefit analysis. Tamayo’s findings are correct for the analyzed data set on the
aggregate. However, his recommendations are nearly completely counter to the rankings (Tamayo, 2011, Table 3, p. 23) in those FYs, which were subject to sizable changes in consecutive years. The relevance of point in today’s environment is an understanding that the Marine Corps’ enlisted endstrength is likely to decrease by approximately 20,000 personnel in coming five FYs. (Amos, 2012) This volatility in endstrength is not a particularly well suited environment for a forecasting model that essentially states that the current situation very much mirrors the last observed situation.

2. Analysis of the U.S. Marine Corps’ Steady State Markov Model for Forecasting Annual First-Term Enlisted Classification Requirements

In his thesis, Nguyen (1997) proposed an alternative to the then-current method of forecasting continuation rates that were used by manpower planners as the basis for their forecasting. His claim stated that although the manpower planners were using equations in the steady state Markov Model properly, some components within the model needed to be reexamined.

Specifically, the manpower planners at the time were using a Markov Model that determined requirements using a weighted-average continuation rate. This weighted-average then assumed that the most recent attrition information was the most relevant and reliable, but it is also subject to greater relative over/under estimation due to this greater influence of the most recent years. In fact, Nguyen graphs continuation trends for PMOSs 0121 and 6521 (Cohorts 1986–94) to demonstrate this roller-coaster effect.

Nguyen’s thesis proposes that a revised steady-state model which employees an average continuation rate versus the weighted-average continuation rate would produce calculations that are more accurate and better forecast classification requirements for each Primary MOS (PMOS). Additionally, proposed changes to the spreadsheet formulations would lead to more precise (admittedly minor) results due to the effect of eliminating rounding errors that would otherwise be compounded with successive calculations.
The shortcoming to Nguyen’s model is that it assumes a constant total Grade Adjusted Recapitulation (GAR) requirement. This assumption is not a problem so long as total billets required for a given PMOS remains constant from year to year. However, in those scenarios that a GAR requirement changes from year to year in a particular PMOS and/or when overall endstrength results in changes to nearly all PMOSs, this model may not accurately account for such variations while still producing precise forecasts.

3. Forecasting Marine Corps Enlisted Attrition Through Parametric Modeling

In his thesis, Hall’s (2009) purpose was to apply a parametric modeling (specifically survival analysis) of historical data sets of enlisted personnel to develop a more efficient forecasting tool for manpower planners. Further, Hall’s intent was to tease out those variables that effect attrition behavior and focus on those contributing factors that lead to non-end of active service attrition (those Marines who did not complete their contracted term of enlistment).

Hall’s thesis specifically researched those contributing factors that influenced non-end of active service (NEAS) attrition. This NEAS attrition is an important element of the overall losses to the manpower system within a given FY. As Hall points out, personnel end strength is calculated at the end of each FY as:

\[
\text{Endstrength} = \text{FY beginning strength} - (\text{minus}) \text{losses} + (\text{plus}) \text{gains}.
\]

Although Hall’s end strength equation is generically correct, it is missing a key element that makes the equation more significant to manpower planners. The variable of capturing the net change to endstrength between consecutive FYs adds a depth and richness to the equation that is missing without its inclusion. Perhaps a more meaningful equation would include:

\[
\text{Endstrength} = \text{Beginning Strength} - \text{losses} + \text{accessions} +/\- \text{net change in FYs}
\]
C. CENTER FOR NAVAL ANALYSIS REPORT

1. Endstrength: Forecasting Marine Corps Losses Final Report

The purpose of this study is stated as a “recognition of the importance of correctly forecasting endstrength losses and gains and the severe consequences of incorrect estimates” (Hattiangadi, Kimble, Lambert, & Quester, 2005). Further, deviation from endstrength requirements results in one of two negative consequences for the service. First, carrying too many personnel into a future FY (over endstrength) constitutes an over expenditure in the MPMC account that must then come out of the O&M account. Second, carrying too few personnel forward may have impacts upon the operational readiness of units should they have fewer assigned personnel then necessary to properly execute their assigned missions. Citing a specific example, the authors note that, “The endstrength planners finished FY04 with an enlisted endstrength of 2,040 Marines above target.”

The authors state that at the beginning of their study the Marine Corps had “no institutionalized and document methodology for forecasting losses, (that) no one had made a systematic attempt to determine whether the (then) current combinations of methods could be improved, (and) no structured capability existed to run loss scenarios”.

The authors present two alternatives to aid planners in their task of forecasting. The first alternative presented is an optimization tool based upon the methodology that enlisted planners in the Air Force use to determine weighting of historical data to forecast future personnel predictions. The second alternative presented is exponential smoothing. The technique of using exponential smoothing assigns the greatest weight to the most recent observation and relatively less weight to older observations. Additionally, the authors note that exponential smoothing can be adapted to account for seasonality and trend patterns within historical data observations.
III. MODEL OVERVIEW

A. INTRODUCTION

The foundations for the transition models analyzed in this thesis are based upon the Markov Chain model structure. The following is a brief discussion of the principles and notation associated with these models. Fundamentally, the purpose of these models is to “Serve in a practical manpower planning situation, (and to answer the question of) what should the recruitment numbers be in order to achieve a desired structure in a specified time?” (Bartholomew, Forbes, & McClean, 1991)

B. PRINCIPLES AND ASSUMPTIONS IN THE MARKOV CHAIN

1. States

The system consists of states. That is to say that the model is partitioned into defined locations within the system. Examples of states include paygrade, time in grade, years of service, etc. The states of interest in this research is years of completed service (YCS).

2. Markovian Property

Although the notation for this assumption follows in upcoming paragraphs, this property essentially states that the probability that the system transitions to a state \( j \) depends ONLY upon current state. This “memory-less” property lends itself to transitioning to the next state only depends upon current state location.

3. Stationary Transition Probabilities

The probability of transitioning between states is constant over time. Although the probability of transitioning from state to state is likely to experience some variation over time, these probabilities must be treated as constant to provide a baseline for comparison and to keep our models relatively simple and manageable to use by planners.
C. NOTATION

The following notation conventions are drawn from academic literature (Bartholomew, Forbes, & McClean, 1991).

Let $i =$ some initial state (years of service) and $j =$ some subsequent state, then

$$ p_{ij} = \Pr(\text{an element transitions from } i \text{ to } j \text{ in one timestep}) $$

We define $P = \{p_{ij}\}$ as the matrix of transition probabilities. For four-year Terms of Enlistments (TOE), a Marine can be in one of four states {0, 1, 2, or 3} YCS during his or her first term. Therefore, their associated $P$ is a 4x4 matrix. MOSs with five-year TOEs require an additional state and thus require a 5x5 transition matrix.

D. CURRENT STEADY-STATE MODEL

A steady-state model is currently used by manpower planners to determine the number of new accessions required to match long-term population requirements. The following figure graphically represents the transition probabilities between the various states of potential YCS. In this simple model, personnel either transition to the next state or attrite out of the system when YCS changes. In this model (depicting a four year TOE), Marines with fewer than 12 months of service are partitioned into state 0, those with more than 12 months but less than 24 months are in state 1, and so on. Since most enlistment periods are for a period of 48 months, the final state in this system is 3 (greater than 36 months of service, but less than 48 months). Upon completing 48 months of service, all personnel in state 3 transition into the attrition state.
Figure 1. Transition Probabilities Between the Various States of Potential YCS

Keeping with notation convention, let \( R \) be the total number of new Marines assigned to an MOS. Further, \( r \) is the distribution vector associated with MOS assignment over the given states (YCS). Although most Marines obtain their MOS in state 0 or 1, there is a proportion of Marines (especially those with a five year TOE and long MOS producing schools) who do not attain their MOS until a later state. Therefore, our \( r \) vector tells us that some proportion of each assessed FY cohort enters the system in state 0 and others from their cohort in state 1 or later. Therefore, \( R^*r \) is our row vector that tells us the quantity of each assessed cohort that is assigned to a given state.

Irrespective of initial beginning inventory of personnel, the current model produces a steady-state inventory of a constant number of personnel by state and required to be accessed (\( R \)) and distributed across states proportionality (\( r \)) each FY is captured with the following:

\[
n^* = Rr(1 - P)^{-1}
\]

The First-Term Inventory Planner must however plan for accessions that support the requirements stated in the GAR. Although the specifics of determining the target (\( T \)) are discussed in additional detail in the next chapter, the underlying goal of the planners is to minimize the difference in the number of Marines accessed and the target. Mathematically, this goal is represented with:
\[
\min \sum_{j=0}^{J-TOE} \left(n_j^*\right) - T
\]

\[s.t.: R > 0\]

where \( n^* = Rr(1-P)^{-1} \).

The model manager typically selects some future year GAR value for the target, usually the third out year. The rationale is that since the model assumes the system is in steady-state, that to the extent that the system is not in steady-state, then building to a year in the future provides the best chance of driving the system to steady state most effectively.

E. PROPOSED ENHANCED CLASSIFICATION PLAN MODEL

This volatility in the forecast thus lends itself to a transient, fixed inventory model that accounts for current inventory, annual accessions into the various states, and most importantly, allows for changes in accessions to meet changes in required endstrength. Building upon the notation previously discussed, the additional variable of year \( (y) \) is added to the equation. Solving for \( R \) in this formula then provides for the annual accession requirement (target).

\[
n(y) = n(y-1)P + R(y)r
\]

Within this model construct, planners aim to minimize the difference between the number of classified into a particular MOS’s targeted requirement \( n(y) \) and each of the build years of the FYDP that is associated with each GAR, given by \( T(y) \) (Seagren, (working draft)).

\[
\min_R \sum_{y=1}^{4} \left( \sum_{j=0}^{5} n_j(y) - T(y) \right)^2
\]

where: \( n(y) = n(y-1)P + R(y)r \) \( y = 1, 2, 3, 4 \)
In addition to allowing the accession quantity to change as endstrength (target) changes, the transient model can support additional constraints to limit the magnitude of change allowable to MOS target strengths in successive years. Although the tolerance levels on these constraints could be adjusted by the individual manpower planner, the intent of the constraints is to purposefully smooth out and limit variation in target requirements from year to year. The goal of minimizing these ebbs and flows is to create a manpower flow of personnel inside a system with reduced volatility. The effect of controlled builds and/or descents of personnel requirements aids in reducing or eliminating the proverbial “pig in the python” effect. Sharp increases and/or decreases in requirements have a detrimental effect on force shaping and both system effectiveness and efficiency.

F. SUMMARY

As discussed, both models share the common characteristics as described in common manpower planning models (Bartholomew). The current steady-state model provides planners with a generally “good enough” solution. However, the dynamic state of the Marine Corps manpower system layered with the current political and economic environment calls for systems that pursue continuous improvement. Once planners account for the current inventory and revise MOS distribution vectors ($Rr$) to account for the differing MOS’ time-to-train (TTT), they employ a model that is more accurate and responsive to changes in the manpower system.
IV. ANALYSIS OF MODELS

A. OVERVIEW OF THE ANALYSIS

1. Purpose of Analysis

Prior to diving into the analysis of the different outputs that the two models produced, it merits a moment to reflect on the purpose of this research. As introduced in the opening chapter, the purpose of this research is to address the following questions:

1. Primary Question: Would an enhanced attrition model serve as a more accurate forecasting tool to support the development of the Marine Corps’ Enlisted Classification Plan?

2. Secondary Question: If more accurate, what is the relative magnitude in the difference in accuracy (i.e., how much better of a predictor is the enhanced model)?

At its heart, the purpose for this research and analysis is to serve as the foundation for model validation. Further, this analysis is intended to serve as an input to decision-makers as they consider potential changes to current processes.

2. Method of Analysis

First we assess GAR targets for each build year between FY2001 and FY2011 to determine the validity of the steady-state assumption. Next, we use historical data in that time period to build transition matrices and estimate attrition for each MOS and build year using both the steady-state model and the transient model.

3. Endstate of Analysis

The endstate to this analysis serves as an indication to the value of changing the methodology for developing the Classification Plan. The current model is less complex and there is certainly value to a process and system that is relatively simple to operated and execute. However, should additional complexity also produce additional clarity and more accurate estimations, then perhaps the “juice is worth the squeeze.” Ultimately, should the hypothesis that a transient model performs better be valid, then this analysis provides justification for the Marine Corps’ adoption of this transient Markov model.
B. STEADY-STATE ASSUMPTION

Fundamentally, the steady-state assumption asserts that the manpower system is static. Within the context of this analysis, this assumption translates to an environment in which annual accessions, individual MOS requirements, and overall endstrength remain constant over time, or at least sufficiently stable for the steady-state assumption to apply. Intuitively though, we understand that this environment is dynamic and subject to change as a result of any number of potential internal and/or external forces at work. Therefore, with the scenario in which either MOS targets and/or other inputs to the models changes, there is merit in questioning if the steady-state assumption is appropriate for planners. The following tables show the dynamic nature of both the number of accessions and the overall authorized endstrength from FY01–FY11.

Figure 2. NPS Enlisted Accessions
As the previous tables show, both accessions and endstrength numbers have experienced substantial variation over the past ten years of observations. Further, recent press releases from the Commandant of the Marine Corps project that endstrength requirements will continue to decline from the current endstrength of 202,200 to 182,100 by FY17 (Amos, 2012). At least taken together, it appears there exist reasons to question the validity of the steady-state assumption.

C. INPUTS TO THE ANALYSIS

1. Description of GAR Data

MOS target inventories are derived from the FY00–11 GARs. In addition to each FY’s base year, each FY forecasted manpower requirements by MOS for each of the upcoming years in the future years defense plan (FYDP). We compare future requirements (i.e., the FY06 build year from the FY04 GAR) with present-year requirements (FY06 requirements from the FY06 GAR) to assess the steady-state assumption.

2. Transition Probabilities

The data from which we have to build the models are all active duty Enlisted, first-term Marines from 1996 to 2008. In order to estimate transition probabilities, we categorize each Marine by MOS and YCS. So, for each category and build year, we have total number of Marines that began the year, the number that left during that year as an
attrite, the number that continued to the next year, and the number that accessed into that state during the next year. Additionally, both models use the same probability matrix as an input to the analysis (Seagren, (working draft)).

3. Target Population by MOS

Starting with FY00 and continuing through FY11, each years’ base GAR and accompanying FYDP requirements by rank and MOS were analyzed. From each of these FYs’ GARS, a four and five year term of enlistment (TOE) target number was developed. The four-year TOE figure was the sum of all E–4 and below requirements. The five-year TOE requirement was the sum of the E–4 and below population and 22% of the E–5 population requirement. The Enlisted Career Force Control Program within the Marine Corps is used to both actively shape the inventory of Marines by grade and MOS to meet the requirements of the Marine Corps and to stabilize retention in order to standardize promotion tempos across all MOSs. Further, the time in service (TIS) promotion targets to Sergeant (E–5) is four years. Therefore, on the aggregate (meritorious promotions aside), we assume that those with a four-year TOE is a Corporal (E-4) or below. In addition, promotion to Staff Sergeant (E–6) is targeted at the 8.5 year TIS mark. Since an individual Sergeant would then remain in that grade for 4.5 years, approximately 22% of Sergeants would have no more than 12 months’ time in grade (TIG) as a Sergeant (assuming that the population distribution of Sergeants was consistent across that 4.5-year period (MarAdmin 433/11, 2011).

4. Distributions of Accessions

Although most Marines attain their MOS within the first year of an individual’s enlistment a sizable portion of Marines accessed each year do not ultimately attain their MOS until into their second year of service. Prior to attaining their MOS all Marines attend recruit training for approximately three months. Following recruit training, Marines diverge into MOS producing pipelines (MOS dependent), each with varying timelines.

The key factor for consideration that manpower planners must account for is this population that has been accessed, but not yet attained an MOS. This dynamic of
accessed future gains to an MOS is partially mitigated with the assignment of training MOSs (0100 as an example for future 0111s). This relationship is moderately correlated, at best, due to classification errors, MOS school attrition, etc. It would almost be analogous to knowing that your spouse wrote a check, but you are not 100% certain of its dollar amount.

With the knowledge that the distribution of accessions (into classified MOSs) are seen over multiple stages (YCS) and the assignment of training MOSs is unable to provide a sufficiently reliable and clear picture of MOS health, manpower planners rely upon analyzing their populations of interest after they’ve been classified with a primary MOS. This distribution of assignment of an MOS over the YCS was then used as a proxy to account for the average time-to-train (TTT). Therefore, this TTT then accounted for the time an individual Marine attended recruit training, Marine Combat Training (MCT), and/or MOS producing school.

After reviewing the average TTT by MOS, the author then made the assumption (grounded in a personal knowledge of sample MOSs) whether to use the 4- or 5-year targets by MOS. Generally speaking, the TTT could be viewed as a bimodal distribution. Although ranging from a minimum TTT of .388 YCS (approximately 4.5 months) for assignment as a Correctional Specialist (5831) to a maximum of 2.49 YCS (approximately 30 months) for Middle East and Asia-Pacific Cryptological Linguists (2671 and 2673), the 4-year TTT observations were predominately in the .6–.9 YCS (approximately 7–10 months) and the 5-year TTT observations were in the 1.1–1.5 YCS (approximately 13–18 months) window.

D. TESTING OF THE STEADY-STATE ASSUMPTION

In this section, we examine the validity of the steady-state assumption at the level of the MOS. We assert that a necessary condition for a system to achieve steady-state is some amount of relative stability in the desired target of the population. The GAR provides target levels of each grade and each MOS for the present year and each of the six years throughout the FYDP.
This means that the target inventory level for a given MOS for a given FY has been crafted multiple times before the execution year. We examine the relative stability of these targets for given years through time to determine the extent of volatility.

Figure 4 illustrates this phenomenon for the Rifleman MOS. Notice that for any particular year, there exist multiple target levels, only one of which was in effect during the year of execution. This volatility is especially pronounced in the latter half of the decade as endstrength and accessions increased.

Figure 4. Rifleman MOS

**1. Coefficient of Variation**

One step in the analysis is to understand the relative variance in MOS target populations over the previous FYDP years’ forecast as they compare to their base year.
GAR (i.e., FY04–07’s projected FY08 target compared to the actual FY08 target). This variation (standard deviation) then divided by the average of those forecasts and actual target (mean for the MOS) produces a coefficient of variation (CV). The value to this measure is that it puts the magnitude of variation into a common perspective rather than a sheer number (Keller, 2009).

The following figure represents the summary of the relative CV for FY04–11’s GARs. As the bar graph shows, there has been a steady increase in the size of the CV over the past years. This increase in CV is an indication that more and more volatility has existed in recent manpower planning. Not surprisingly, this indication of volatility (increasing CV) has been parallel with the recent growth in the Marine Corps’ endstrength. Further, this greater volatility should persist as the Marine Corps is estimated to cut endstrength by 5,000 per year over the upcoming four FYs (Amos, 2012).

![Coefficient of Variance](image)

Figure 5. Summary of distribution of Coefficient of Variance of GAR targets across MOSs for FY04–11
The above output highlights the difference between relatively “good” (minimal volatility, cv_04 and cv_05) compared against “bad” (increased volatility, cv_10 and cv_11) year. However, during bad years the CV increased to between 9.5 and 10.5 percentage points and had a fourth of observations exceeding 12 percentage points.

Additional details on the CVs for each FY can be found in Appendix A.

2. Variability Over Time Lags

A common technique when employing the steady-state model is to select as a target the GAR requirements for the third year of the FYDP. The thought is that accessing the steady-state requirement for a future year will result in a smoother trajectory in achieving steady-state. In this section, we compare the relative changes between future targets at various lags (1 through 4 years) and the actual requirement as executed.
In other words, the FY01 GAR contains an FY04 future requirement. Comparing those targets with the actual FY04 GAR as executed is one of the comparisons included in the 3-year lags.

As shown in the following figure, the longer the lag (and further away the forecasted time period is from execution), the less the greater the variance between the future expected requirement and the requirement as executed.

Figure 7. Summary of Distribution of Variance in Time Lags Across MOSs

As Figure 7 indicates, at lags 3 and 4, greater than 25% of all MOSs have a difference of 10% or more. Additional details on the Lags for each time period can be found in Appendix B.

This variability over time is shown specifically for the 0311, Rifleman, MOS in Figure 8 below. The figure shows how the target population changes from the original base year GAR target thru the build years of the FYDP. A cross-section of MOSs is included in Appendix C for review.
These MOSs were selected by the author because they were not only larger in population size, but also representative of the Marine Corps’ Marine Air-Ground Task Force (MAGTF) and a combination of combat arms, combat support, combat service support, and aviation MOSs.

Figure 8. Example GAR Requirements Over the FYDP
3. Summary

Building upon the intuition that manpower planners operate in a dynamic environment, an examination of recent GARs reveals that the current steady state model is out performed by an enhanced model across a range of measurements. Further, the common practice of using a three-year lag/time horizon as the “gold standard” for planners is currently subject to mean forecast error of nearly 10 percentage points across all MOSs.

E. ATTRITION PREDICTION

In this section, we compare estimated attrition versus actual attrition for each model. Given the dynamic nature of the environment in which manpower planners must forecast, this section puts into context the relative comparison of how much better does the transient model perform over the steady-state model?

1. Notation and Mathematical Description

For each year between FYs 1996 and 2011, a comparison of the estimated attrition for the legacy and transient models is conducted. Actual data from those years were inputs to the transition matrices (P) and recruit distribution vectors (r). Additionally, we vary the window size for each model from one year to nine. For example, for the build year 2001, we have five prior years of history with which to build our model, so we consider windows of sizes from 1 year to 5. However, for build year 2010, we vary window size form one to nine. Recent research suggests that shorter windows of 1 to 3 years typically yield the best prediction of attrition behavior (Tamayo, 2011), but window size up to nine was used to ensure that any differences in the performance between the models did not depend on window size.

Given that the estimate of the transition probability from state i to state j is given by:

$$\hat{p}_{ij} = \frac{f_{ij}}{n_j}$$
where \( f_{ij} \) is the total flow of Marines from state \( i \) to state \( j \) in one time period, and \( n_j \) is the total number that began the time step in that state (Bartholomew, Forbes, & McClean, 1991). Let \( \mathbf{w} \) be the wastage rates and associated vector for a given MOS, build year, and window combination and is defined as follows:

\[
\mathbf{w} = \{1 - \hat{P}_{01}, 1 - \hat{P}_{12}, 1 - \hat{P}_{23}, 1 - \hat{P}_{34}\}.
\]

Note for a five year TOE, this vector has five elements. The dot product of this vector and an inventory vector yields an estimate of the number of Marines who attrite for that year. It is also important to note that \( \mathbf{w} \) is the same for both the legacy and transient models. The difference between the models is the inventory vector to which it is applied. For the legacy model, this inventory is the long term inventory vector \( \mathbf{n}^* \).

\[
a_t(t) = \mathbf{n}^*(t) \cdot \mathbf{w}(t)
\]

where \( t \) denotes the build year, which varies from FY2001 to FY2011. For the transient model, the inventory vector corresponding to the build year is used to for the attrition estimate.

\[
a_T(t) = \mathbf{n}(t) \cdot \mathbf{w}(t)
\]

The respective attrition estimates are then compared to the actual number of attrites for each year. Let \( a(t) \) be the number of attrites for build year \( t \). Then trans_MAPE is given by:

\[
trans_{MAPE} = \frac{a(t) - a_T(t)}{a(t)}.
\]
And ss\_MAPE is given by:

\[ ss_{\text{MAPE}} = \frac{a(t) - a_s(t)}{a(t)}. \]

Thus, for each build year, initial training MOS, and window, a total of 8847 design points, we have a value for trans\_MAPE and ss\_MAPE.

2. Number of Observations (Counts)

The most straightforward method of comparing the performance of the two models is by simply counting the number of times the transient model produces a better estimate than the steady state model. We find that the transient model produced a more accurate forecast for 5,321 out of 7,379 design points (approximately 72% of the observations).

Although the transient model produced more accurate forecasts over 70% of the time and this may be sufficient to address the primary research question, the differences in forecasting accuracy must be put into further context to get at the heart of addressing the secondary, and perhaps more important and relevant, research question.

3. MOE Comparisons

In this section, we examine the significance of the fact that the transient model tends to generate better estimates relative to the steady-state model. In particular, we seek to determine the magnitude of the improvement and consider whether the better performance is worth the extra requirements for input data and computation.

Figure 9 shows that across all MOS, build year, and window combinations, the transient model has a forecast error mean of 26 percentage points from the actual observations. These forecasts are put into additional perspective by reviewing the median and quartile distribution. With a median value of 12.2 percentage points, more than half of the over 7,300 forecasts were off from the actual by greater than 10 percentage points. Further, approximately one-fourth of the forecasts were off by greater than 25 percentage points (75% quartile).
However, an analysis of the steady-state model reveals that produced a forecast error mean of 40 percentage points. Further, over half of the observations exceeded a 21-percentage point error (median), and approximately one-fourth was over 44 percentage points off (75% quartile) from the actual observation.

Although the proceeding paragraphs are relevant and demonstrate that both models are subject to forecast error, the more useful analysis is to compare the relative
“goodness” of either particular model against the other. As previously discussed, we can do this comparison by comparing the difference between the transient outputs versus the steady-state output for each estimate, which is shown in the third column. Because the mean of this comparison is 0.14, it tells us that, on average, the transient model produces forecasts with MAPEs that are 14 percentage points less than the steady-state model. Although the median is approximately 8 percentage points, in a full one-fourth of the observations (75% quartile) of the transient model were greater than 20 percentage points more accurate at forecasting actual observations. As a measure of comparison, the transient model was 20 percentage points less accurate in only 2.5% of observations.

4. Summary

On all accounts of the various methods and measurements of model performance, the transient model is superior in performance to the current model. Further, when comparing the relative “goodness” of either particular model against the other, the analysis tells us that, on average, the transient model produces erroneous forecasts 14 percentage points fewer than the steady-state model. The dynamic environment that analysts and planners operate within must employ forecasting tools that are responsive to change yet produce the minimum potential forecast error. Further, a forecast model that supports the ability to account for changes in desired endstrength would not only prove useful in the short-term and future draw down of personnel end strength, but also be flexible enough to support any future change to desired endstrength.
V. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

A. SUMMARY

The Marine Corps is a manpower intensive organization with over 200,000 personnel. In fact, the Marine Corps recruits approximately 30,000 new enlisted personnel into its workforce annually. The number of new accessions each Fiscal Year is determined and classified as a function of attrition behavior.

Given that desired endstrength is a function of beginning (current) endstrength + accessions – attritions, it demands that the number of personnel attritions forecasted be as accurate as possible. In this formula, the desired and current personnel endstrengths are known quantities. The accessions required in the future fiscal years are predicated upon attrition. Since the actual level of attrition is an unknown variable, it must be forecasted.

The enlisted manpower and personnel analysts and planners make career force decisions (accessions, promotion, retention, training, etc.) based upon existing and forecasted personnel resources by specific MOS. These career force decisions are driven from the requirements by grade and specific MOS as outlined in the Grade Adjusted Recapitulation (GAR). As previously noted, the variable of attrition must be forecasted as a basis for these career force decisions. Therefore, the more accurately forecasted the level of attrition is as an input to this decision system, the less variability the system as a whole. Further, a forecast model that supports the ability to account for changes in desired endstrength would not only prove useful in the short-term and future draw down of personnel endstrength, but also be flexible enough to support any future change to desired end strength.

Although Marine Corps’ future endstrength requirements are not conclusively known, it is generally accepted that its endstrength will be no greater than 182,200 in the upcoming years (Amos, 2012). This roughly 10% decrease in overall personnel endstrength and subsequent reshaping of its force will have an effect across all MOSs.
Additionally, technological advances, equipment changes, and unforeseen operational requirements contribute to the management of manpower and the subsequent personnel effects inherent in the Marine Corps’ manpower and personnel system.

The Marine Corps currently uses a steady-state Markov model to forecast attrition behavior (retention by MOS and YCS) and subsequently determine its accession requirements. As the Marine Corps draws down in endstrength in today’s fiscally constrained environment, a precise classification plan cannot be overstated. The focus of this research was to evaluate an alternative model to reduce MOS misclassifications, and better account for dynamic changes to endstrength and targets. An analysis and comparison of each models’ predictive performance clearly points to the transient model as being more accurate across a range of performance measurements.

B. CONCLUSIONS AND RECOMMENDATIONS

1. Primary Research Question

Would an enhanced attrition model serve as a more accurate forecasting tool to support the development of the Marine Corps’ Enlisted Classification Plan? If more accurate, what is the relative magnitude in the difference in accuracy (i.e.- how much better of a predictor is the enhanced model)?

a. Conclusions

Analysis from the observed data set output noted that the transient model produced a more accurate forecast on 5,321 out of 7,379 occasions (approximately 72% of the observations). As a relative comparison, the transient model achieved a MAPE that was, on average, 14 percentage points smaller than that of the steady-state model. In over 25% of the cases, this difference exceeded 20 percentage points.

b. Recommendation

Based upon this improved performance, it is recommended that the Marine Corps adopt the transient Markov Model analyzed in this research as the
foundation for future forecasting of the annual Enlisted Classification Plan. This action reduces MOS misclassifications and more accurately account for changes in endstrength and/or GAR requirements.

C. FURTHER RESEARCH AND ACTION

1. Transient Model User Guide

Since implementation of a new model changes the current way of doing business, a user guide should be developed to serve as a desktop guide for the planners tasked to produce and analyze the annual Enlisted Classification Plan.

2. Attrition Rates by MOS and Gender

Literature has repeatedly pointed to the variable of gender as having a differing effect on attrition behavior. Building upon the established framework of the transient model, additionally precision would likely be obtainable if the model was able to account for the different attrition rates by MOS and gender.
APPENDIX A.
APPENDIX B.
APPENDIX C.
1371 - Combat Engineer
<table>
<thead>
<tr>
<th>FY 08</th>
<th>FY 09</th>
<th>FY 10</th>
<th>FY 11</th>
<th>FY 12</th>
<th>FY 13</th>
<th>FY 14</th>
<th>FY 15</th>
</tr>
</thead>
<tbody>
<tr>
<td>738</td>
<td>775</td>
<td>790</td>
<td>725</td>
<td>720</td>
<td>740</td>
<td>700</td>
<td>580</td>
</tr>
<tr>
<td>818</td>
<td>842</td>
<td>853</td>
<td>875</td>
<td>860</td>
<td>858</td>
<td></td>
<td></td>
</tr>
<tr>
<td>823</td>
<td>822</td>
<td>816</td>
<td>859</td>
<td>842</td>
<td>841</td>
<td></td>
<td></td>
</tr>
<tr>
<td>831</td>
<td>815</td>
<td>798</td>
<td>804</td>
<td>827</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>814</td>
<td>814</td>
<td>798</td>
<td>801</td>
<td>805</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>817</td>
<td>821</td>
<td>840</td>
<td>839</td>
<td>847</td>
<td></td>
<td></td>
</tr>
<tr>
<td>827</td>
<td>820</td>
<td>846</td>
<td>835</td>
<td>1008</td>
<td>1030</td>
<td></td>
<td></td>
</tr>
<tr>
<td>836</td>
<td>854</td>
<td>885</td>
<td>972</td>
<td>675</td>
<td>918</td>
<td></td>
<td></td>
</tr>
<tr>
<td>972</td>
<td>954</td>
<td>909</td>
<td>966</td>
<td>999</td>
<td>980</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>122</td>
<td>955</td>
<td>579</td>
<td>918</td>
<td>954</td>
<td>1005</td>
<td></td>
</tr>
<tr>
<td></td>
<td>914</td>
<td>935</td>
<td>912</td>
<td>950</td>
<td>990</td>
<td>1032</td>
<td></td>
</tr>
</tbody>
</table>
**6672 - Aviation Supply Specialist**

<table>
<thead>
<tr>
<th>FY00</th>
<th>FY01</th>
<th>FY02</th>
<th>FY03</th>
<th>FY04</th>
<th>FY05</th>
<th>FY06</th>
<th>FY07</th>
<th>FY08</th>
<th>FY09</th>
<th>FY10</th>
<th>FY11</th>
<th>FY12</th>
<th>FY13</th>
<th>FY14</th>
<th>FY15</th>
<th>FY16</th>
</tr>
</thead>
<tbody>
<tr>
<td>804</td>
<td>854</td>
<td>885</td>
<td>895</td>
<td>865</td>
<td>885</td>
<td>854</td>
<td>873</td>
<td>873</td>
<td>873</td>
<td>873</td>
<td>873</td>
<td>873</td>
<td>873</td>
<td>873</td>
<td>873</td>
<td>873</td>
</tr>
<tr>
<td>873</td>
<td>873</td>
<td>873</td>
<td>873</td>
<td>873</td>
<td>873</td>
<td>873</td>
<td>873</td>
<td>873</td>
<td>873</td>
<td>873</td>
<td>873</td>
<td>873</td>
<td>873</td>
<td>873</td>
<td>873</td>
<td>873</td>
</tr>
</tbody>
</table>
LIST OF REFERENCES


Seagren, C. W. ((working draft)). *An improved USMC enlisted classification plan model*. Monterey.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Fort Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Marine Corps Representative
   Naval Postgraduate School
   Monterey, California

4. Director, Training and Education, MCCDC, Code C46
   Quantico, Virginia

5. Director, Marine Corps Research Center, MCCDC, Code C40RC
   Quantico, Virginia

   Camp Pendleton, California

7. Professor Yu-Chu Shen
   Naval Postgraduate School
   Monterey, California

8. Dr. Chad W. Seagren
   Naval Postgraduate School
   Monterey, California

9. CDR Bill Hatch, USN (Ret)
   Naval Postgraduate School
   Monterey, California

10. Manpower Plans & Policy Division
    Enlisted Plans (MPP-20)
    Quantico, Virginia

11. Manpower Plans & Policy Division
    Integration & Analysis (MPP-50)
    Quantico, Virginia