VALIDATION OF THE ENLISTED GRADE MODEL GRADEBREAKS

Andrew O. Hall

Office of the Deputy Chief of Staff, G1
Military Strength Analysis and Forecasting, DAPE-PRS
300 Army Pentagon
Washington, D.C. 20310-0300, U.S.A.

ABSTRACT

This paper describes the validation of the Enlisted Grade model gradebreaks and describes a current application of simulation in operations research. The Enlisted Grade model is part of a suite of models used by Army strength analysts to forecast active Army strength to develop the Active Army Military Manpower Program and for use in the President's Budget and the Program Objective Memorandum. The Enlisted Grade model gradebreaks were tested by predictive validation, comparison to the Military Occupational Specialty Level System and face validity before acceptance. The validation of the Enlisted Grade gradebreaks was the final milestone in the acceptance of the Enlisted Grade model.

1 INTRODUCTION

The Strength and Manpower Analysis and Forecasting Division of the Army Deputy Chief of Staff, G1 runs a series of models to forecast the strength of the United States Army. The division consists of nine Army Operations Research Analysts and three civilian analysts. All the military analysts have advanced degrees in operations research, applied mathematics or a related field.

Our division's primary task is to produce the Active Army Military Manpower Program or AAMMP. This program is a seven-year projection of the Active Army's strength by officer and enlisted grade. We use the forecast to estimate the annual cost of our military manpower and submit this estimate as a part of the President's Budget and the Program Objective Memorandum (POM). The President's Budget details our near-term program while the POM details the manpower program for a six year period.

We create our projections using a suite of models called the Active Army Strength Forecaster 2 or A2SF2. This is our second suite of models and its three components replaced or are replacing our legacy systems. We have fielded the first two portions of the A2SF2 system, the Individuals Account (IA) and the Enlisted Grade (EG)

models. The IA model replaced our TTHS model which forecasts factors for trainees, transients, holdees and students. EG replaced ELIM/COMPLIP, Enlisted Loss Inventory Model/Computation of Manpower Program using Linear Programming, or ELIM. The ES model will replace the Military Occupational Specialty Level System (MOSLS) which adds the military occupational specialty (MOS) to our forecast.

The first suite of models was developed in the early 1970s by General Research Corporation, GRC, now a part of AT&T Government Solutions. The new suite of models is also being developed by AT&T Government Solutions in conjunction with Strength and Manpower Analysis and Forecasting Division (PRS). The validation of the EG gradebreaks was the last milestone in the acceptance of the EG model. The Enlisted Specialty (ES) model is in the final stages of development and the validation and verification process will begin in earnest this Fall.

2 THE A2SF MODEL SUITES

A change in variables of interest and technological model innovations motivated the creation of the new suite of models. The ELIM model focus on education level and test scores was relevant to the newly created all-volunteer force during the 1970s. As the Army moved into the 21st Century, a shift to modeling the force by grade was warranted. Advances in simulation, network, linear programming, data mining and optimization algorithms all suggested a need for new models. The advances in computational power from the 1970s alone justify the creation of the new models. Computing advances also made updating the terminal-based interface to a web-based graphical user interface (GUI) a natural choice.

The web-based GUI allows the analyst to control the models remotely. No longer is the analyst required to be physically in the Pentagon to start model runs. Currently the EG model has a six hour solve time, and the ability to start runs during the evening and weekends without traveling to the Pentagon has dramatically increased productivity. Also,

the point-and-click GUI offers an ease of use improvement over the 1970s terminal connection technology.

Both model suites are comprised of three different models. Each of the new models replaced its predecessor , but the differentiations between the models have blurred with the addition of the grade component to IA and EG. Currently, IA, EG and MOSLS are operating together awaiting the deployment of the ES model. The new models offer increased integration, incorporate additional variables, and provide more robust segmentation of output data.

2.1 TTHS and IA

The TTHS model and IA model work similarly. Both forecast the number of soldiers not available to be assigned to units. This number of soldiers is necessary to calculate the operating strength: the number of soldiers assigned to Army units. The IA model improved upon the TTHS forecast by adding a grade component to the individuals account. Both models calculate factors for each sub-account of TTHS by grade. The IA model, however, creates a different factor by grade for each sub-account, since some sub-accounts are dominated by grade. While medical holdees and prisoners, two sub-accounts of holdees, are not dominated by grade, trainees are all in the rank of E3 and E4. The TTHS model would calculate different rates for each sub-account, but neglected the grade component. By adding the grade to IA, we forecast trainees at a different rate for E3 versus E9.

Our forecasts begin with the IA model creating factors for each component of the individuals account. We evaluate 84 months of data to determine factors for each component of TTHS by grade. The A2SF2 suite requires the creation of these factors before the EG model can run.

2.2 ELIM and EG

The ELIM and EG models were both designed to forecast the active Army strength in aggregate. The ELIM model focused on the education level, test scores and term of enlistment to forecast the aggregate active Army strength. The EG model has the grade of the soldier as a central variable. Before the fielding of EG, MOSLS added both grade and specialty to the forecast. Currently, EG forecasts strength, operating strength, gains and losses by grade, gender, month of service and term of service. The ES model will add the specialty component to A2SF2.

After IA has developed factors for TTHS, the EG model forecasts the number of soldiers by month for an eight-year projection. Rates for gains, losses, promotions and re-enlistments are developed from 84 months of data and used to simulate the Army's enlisted population over an 8 year time horizon. Once the model has completed the aggregate modeling by grade, the strength by grade is passed though the factors created by the IA model to de-

termine operating strength. Operating strength is defined as the aggregate strength by grade minus the number of soldiers in the individuals account by grade.

The EG model aggregate forecast was the first validation project upon receiving the model. Validation of the EG aggregate forecast allowed us to retire the ELIM model. Next, we focused on validating the EG grade component, although we will not be able to retire MOSLS until ES has been deployed and validated.

2.3 MOSLS and ES

The MOSLS model added both grade and specialty component to the forecast. In addition to the multiple grades in the model, over 260 MOSs are simulated in MOSLS. Currently, the specialty modeling is completed with MOSLS. Although the aggregate strength results from EG are used as the initial conditions, MOSLS does not normalize to EG during its model run. With ES, the model will normalize to EG and strengths by grade will match in both models.

Both MOSLS and ES are stochastic optimizations. The models have both an optimizer and simulator that are iterated until a convergence tolerance is reached. MOSLS simulates losses and optimizes promotions and reclassifications. The simulator predicts stochastic behavior and the optimizer recommend actions that can be controlled by policy.

ES uses an optimization that matches operating strength by MOS and grade to the Army's force structure allowance. The optimizer prescribes promotions, accessions and reclassifications. Next, the ES model uses a simulator to adjust for the changes in behavior due to different promotion, accession, and reclassification programs. After the simulator runs, the output is normalized to EG. The model alternates runs of the optimizer and the simulator until they converge.

The ES model will be fielded during Fall of 2004 and will add the specialty component to the new suite of models. MOSLS will continue to be run in conjunction with EG to provide military occupational specialty level detail in the interim.

3   THE EG MODEL

The EG model is formulated as a network flow model based on conservation of flow and constraints designed to model Army personnel policies. Twelve different networks are solved using a linear programming model. Female and male soldiers are separated into networks for first term soldiers and re-enlistees. The first term soldiers are further divided into five additional networks for initial terms of service of two, three, four, five and six years. Each node in the network model represents a months of service and grade combination with variables for gender, grade, month of service, time and term. Grade has values of E3 though E9. Months

of service (MSV) can range from 1-434 and time is the month of the projection, 1-96 months.

The network flow model provides for conservation of flow, with each soldier having the option of advancing to the next node at current grade, being promoted, changing networks or being eliminated from the model, Figure 1. Once the eight year optimization is complete, we are interested in variables of strength, losses, gains, and promotions. The enlisted strength can be partitioned by any of the variables as we analyze the losses, gains, strength and promotions.
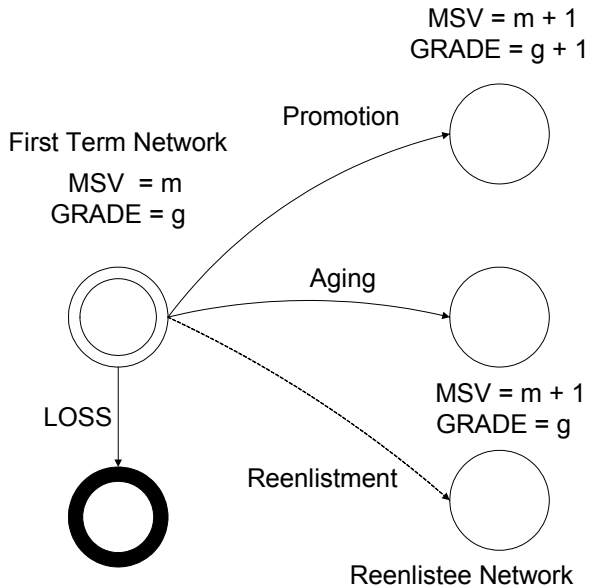

Figure 1: Node Level Flow in EG

4 VALIDATION OF THE GRADEBREAKS

The validation of the EG gradebreaks proceeded by making a few assumptions. The major policy tool the gradebreaks is used for monthly is to prescribe promotions. The promotions set that is prescribed are from E5 to E9. The complement of this set of prescribed promotions are the promotions to E4 which are simulated by historical time-in-service rates. All privates are aggregated at the E3 level so promotions to E2 and E3 are not modeled. The validity of the model will be determined by its ability to forecast the grades of E5 to E9. The two variables chosen as the focus of the analysis were strength and losses. From these two variables all the other variables can be derived.

Three different validation techniques were selected in the analysis of the EG gradebreaks (Sargent, 2003). First, the projection versions from EG and MOSLS were compared to history. Second, statistics from seven-year projections from MOSLS and EG were compared, and lastly we assessed the face validity of the EG gradebreaks.

4.1 Predictive Validation

We compared three projection versions from both MOSLS and EG from December 2002, January 2003 and February 2003 with the historical data from the first twelve months of the projection. Table 1 contains the losses for E9 by month and the projections from the EG and MOSLS models. Losses were evaluated by grade for each of the three projection versions. The paired T test and Welch test were used to create confidence intervals for the EG and MOSLS losses (Law and Kelton, 1991). EG losses were not statistically different from history, Table 2 and Table 3. After examining strength, operating strength, TTHS, promotions and losses, we focused our validation testing on strength and losses.

Table 1: E9 Losses EG Dec 02 Final

E9 LOSSES	EG	MOSLS	ACTUAL
200301	51	37	45
200302	53	45	52
200303	34	39	21
200304	33	38	42
200305	45	36	39
200306	53	32	53
200307	57	36	60
200308	43	43	52
200309	63	53	57
200310	50	39	21
200311	31	44	34
200312	26	30	50
Mean	45	39	44

Table 2: Paired T Test for E9 Losses

T	I	Xi-X1	1/2 length	90% CI	
Data	1	0.00	23.82	-23.82	23.82
EG	2	1.21	18.46	-17.25	19.67
MOSLS	3	-4.50	32.02	-36.52	27.52

Table 3: Welch Test for E9 Losses

Welch	I	Xi-X1	DF	1/2 length	90% CI	
Data	1	0.00		23.82	-23.82	23.82
EG	2	1.21	34.57	8.46	-7.24	9.67
MOSLS	3	-4.50	16.58	7.26	-11.76	2.76

Another example of the forecast and data is found at Table 4. This example shows the aggregate strength of E6. This is one of the larger populations, much larger than the population of E9. The confidence intervals at Table 5 and 6 again show no statistical difference from history for both the EG and the MOSLS projection. The E9 losses and the E6

populations displayed here are samples of the statistical results and show an example with a small population and an example with a larger population. The results across the ranks of E5-E9 were very similar for losses and strength data.

Table 4: E6 Strength EG Feb 03 Final

E6 Strength	EG	MOSLS	ACTUAL
200303	56532	56805	56751
200304	56487	56732	56717
200305	56930	57184	57177
200306	56800	56992	57250
200307	56737	56821	57372
200308	56550	56637	56964
200309	56553	56762	57532
200310	56336	56399	53909
200311	56241	56264	53941
200312	56610	56709	53962
200401	56456	56444	53952
200402	56283	56277	53989
Mean	56543	56669	55793

Table 5: Paired T Test for E6 Strength

T	I	Xi-X1	1/2 length	90% CI	
Data	1	0.00	2949.65	-2949.65	2949.649
EG	2	749.96	1935.29	-1185.33	2685.254
MOSLS	3	875.84	1995.17	-1119.33	2871.014

Table 6: Welch Test for E6 Strength

Welch	I	Xi-X1	DF	1/2 length	90% CI	
Data	1	0.00		2949.65	-2949.65	2949.65
EG	2	749.96	37.17	1084.52	-334.56	1834.49
MOSLS	3	875.84	33.95	1063.54	-187.70	1939.38

4.2 Comparison to Other Models

The EG model and the MOSLS model were both designed to provide gradebreaks for the manpower program for the time horizon of the POM. The POM always covers six years and this year's project is POM 06-11. I compared the April 04 EG and the April 04 MOSLS projections. Figure 2 and Figure 3 provide examples of E7 and E6 strength forecasts for both EG and MOSLS. The EG forecast has greater variability, especially in the near term. However, the manyears, a measure of the average strength, was very similar in the EG and MOSLS projections. Comparing the EG to the previous standard, MOSLS, the percent error in the senior grades was small, see Table 7. The greatest percent error was in E9, the smallest grade, and the manyears projections had less that one percent error for E6-E8. The manyears number is an important metric that is used to allocate dollars to the manpower program in the President's Budget and the POM.
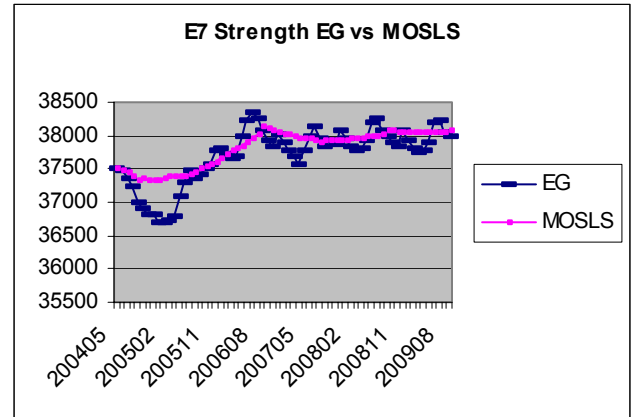
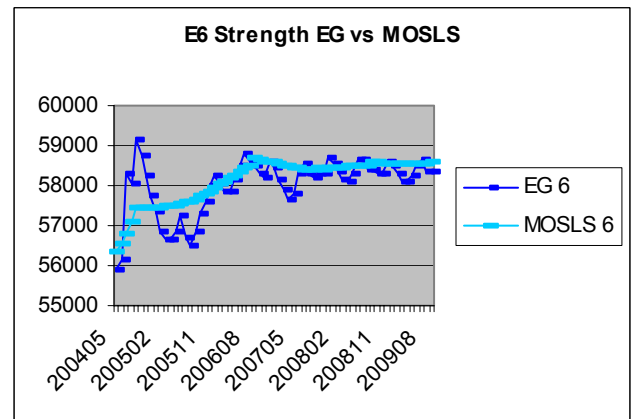

Figure 2: E7 6 Year Strength Projection

Figure 3: E6 6 Year Strength Projection

Table 7: Percent Error in 6 Year Projection

	Monthly	Manyears
E5	-0.01%	0.64%
E6	-0.37%	-0.21%
E7	-0.10%	-0.29%
E8	0.71%	0.99%
E9	4.05%	3.82%

4.3 Face Validity

PRS developed a spreadsheet tool to present the monthly gradebreaks information including losses, promotions, gains, strength, operating strength, manyears, authorizations, TTHS, and percentage fills. This information and recommendations are presented to the Directorate of Military Personnel Policy (DMPP). We examined the results from each EG run for the first year until we were satisfied that the gradebreaks were accurately managing enlisted operating strength. The gradebreaks were then used in recommendations to DMPP and in this year's POM submission.

Monthly, the gradebreaks from the EG model run are used to determine promotions for E5-E9. Once a year, an-

nual recommendations for the number of soldiers to be selected on a promotion board for promotion to E7, E8 and E9 are developed. The DMPP experts in enlisted promotions have the responsibility to approve the number of promotions that will be allocated each month and allocated for each annual board. DMPP used both the MOSLS gradebreaks and the EG gradebreaks for several months until they believed our new system provided a realistic forecast of the needed promotions.

The acceptance of the face validity of the EG gradebreaks by the two groups of experts, the analysts in PRS and the policy makers in DMPP, made the model credible. Although the least scientific of the test procedures, the acceptance of the model by the DMPP was critical to completing the validation of the EG gradebreaks.

5 CONCLUSIONS

The validation of the EG gradebreaks was the final milestone in the acceptance of the EG model. The validation of the EG model's aggregate forecast had allowed the retirement of the ELIM model. With the new suite of models, part of the functionality of MOSLS was also assumed by EG and will be used to normalize the ES projections to a common gradebreaks.

The three tests conducted to validate the EG gradebreaks all indicated that the model was acceptable. The comparison of three projections of the MOSLS and EG gradebreaks against audited history tested the near-term uses of the model for promotion recommendations and selection objectives for promotion boards. The one year of history was a reasonable choice because promotions are adjusted monthly and waiting for six years of history was prohibitively costly.

Ideally, tests of the long-term projections could be compared against history to evaluate the performance of the model. However, the simulation is constrained by force structure documents that change frequently. The changes in force structure diminish the benefits from continuing to conduct extensive testing against past model forecasts.

The test of the EG model against the accepted standard of the MOSLS gradebreaks over the six year horizon of the POM tested forecasting abilities of the new model against a model that had served the Army well for over twenty years. The low relative error between MOSLS and EG indicates that EG will give a projection that is just as valid as the MOSLS projection. We would like to continue to fine tune the forecasting techniques used in EG to improve on the accuracy of the forecast, but a forecast that is not statistically different from the accepted model is a good benchmark to have achieved.

Lastly, the face validity of the model is key to its effective use in the Army G1. The DMPP and PRS must both be comfortable that the model will provide estimates for promotions and manyears that are at least as accurate as the prior estimates. The officers and noncommissioned officers of DMPP believe the promotion recommendations agree with their professional military judgment and they accept the validity of our model.

The EG model has been accepted by PRS and is used in production mode for managing the Active Army Military Manpower Program. The Global War on Terror has resulted in unique personnel dynamics. The effects of increased nationalism, stop-loss and repeated rotations to Iraq and Afghanistan may be unique to this period in history. One strength of the EG model is its reliance on history and the ability of the analyst to choose forecasting techniques and weighting for blocks of history. The analyst can use as much or as little of the 84 months of history as are needed to improve the forecast resulting from the simulation.

REFERENCES

Law, A. M. and W. D. Kelton. 1991. *Simulation Modeling and Analysis*. 2d ed. McGraw-Hill.
Sargent, Robert G. 2003. Verification and Validation of Simulation Models. In *Proceedings of the 2003 Winter Simulation Conference*, ed. S. Chick, P.J. Sanchez, D. Ferrin, and D.J. Morrice, 37-48. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

AUTHOR BIOGRAPHY

ANDREW O. HALL is an Operations Research Analyst in the Military Strength Analysis and Forecasting Division of the Deputy Chief of Staff, G1. He came to the Pentagon for his second tour as an Operations Research Analyst after serving four years as an Assistant Professor in the Mathematical Sciences Department at West Point. He has a B.S. in Computer Science from United States Military Academy (1991) and a M.S. in Applied Mathematics from the Naval Postgraduate School (1999). His research interests include applied probability, ordinary differential equations and simulation. He is a member of ACM and SIAM. His e-mail address is <Andrew.O.Hall@us.army.mil>.