

## Predicting Undergraduate RPA Training (URT) Student Performance

Jonathan Farmer, Robert Martin, Sung O, Eve Schoenrock, and John Miller

Department of Operations Research  
United States Air Force Academy  
Colorado Springs, CO 80840

Corresponding author's Email: [eveschoenrock2018@gmail.com](mailto:eveschoenrock2018@gmail.com)

**Author Note:** Jonathan Farmer, Robert Martin, Sung O, and Eve Schoenrock are senior cadets at the United States Air Force Academy, collectively working on this project as part of a year-long Operations Research capstone course. The authors extend their thanks to Lt Col Miller, the project advisor, as well as the client, the 558th Flying Training Squadron. The views expressed in this paper are those of the authors and do not necessarily reflect the official policy or position of the Air Force, the Department of Defense, or the U.S. Government. Distribution A: Approved for Public Release, Distribution Unlimited.

**Abstract:** This project is in conjunction with the 558th Flying Training Squadron, the only Undergraduate Remotely Piloted Aircraft (RPA) Training (URT) squadron. The 558th is currently seeing a trainee failure rate of approximately 6%. The project team aims to predict student performance at URT using econometric regression strategies analyzing initial student data for selected trainees and identifying at-risk students early in the training timeline. The 558th seeks to enhance the performance of their trainees by identifying student trends indicative of success and failure prior to URT, enabling them to provide attention and resources to students in need. The team's goal is to provide the 558th with information that will help them reduce the failure rate at URT without requiring additional funding. It was determined that GPA, PCSM, STEM, and Private Pilot's License (PPL) status to be significant indicators of student success.

**Keywords:** Undergraduate RPA Training, GPA, PCSM, STEM, Private Pilot's License

### 1. Introduction

The 558th Flying Training Squadron (FTS) is a United States Air Force unit assigned to the 12th Flying Training Wing at Randolph Air Force Base. Primarily, the 558th focuses on training Remotely Piloted Aircraft (RPA) operators to fly the MQ-1 Predators, MQ-9 Reapers, and RQ-4 Global Hawks.

The Air Force's training of RPA pilots is essential to mission success. Being the only URT squadron, the 558th continuously aims to improve its training processes. To combat this issue, the 558th has asked our team to analyze initial student data to determine characteristics that are most indicative of success or failure in URT. The intent behind this study is to improve RPA pilot effectiveness following URT by enabling instructors to focus attention on at-risk students and tailoring training to each student's anticipated needs.

The primary goal of this paper is to capitalize on current URT infrastructure and processes, improve RPA trainee success rates, and build the best pilots to support the mission of the Air Force.

#### 1.1 Problem Statement

The 558th Flying Training Squadron (FTS) stood up in 2011 as the only Undergraduate RPA Training (URT) squadron. They started with the traditional Undergraduate Pilot Training program and modified that program to suit their needs. Currently, URT utilizes a T-6 simulator for their training events, which does not translate well to operating a Remotely Piloted Aircraft (RPA) and is limiting pilots' advantage. Without asking for additional funding, the 558th is aiming to improve URT by targeting at-risk students and helping them early. The 558th is looking to enhance the performance of its trainees by identifying characteristics indicative of success and failure in URT and focusing its efforts on students that are predicted to fail. By providing these at-risk students with more attention and individualized training, the failure rate of URT should decrease, thereby providing the Air Force with more capable RPA pilots without asking for additional funding.

## 1.2 Related Work

Despite RPA pilots being essential to the mission, there has yet to be much research done specifically on the RPA training pipeline or success factors for drone pilots. Caretta (2013) conducted a study predicting the validity of pilot selection instruments for remotely piloted aircraft. He first highlights the surging demand and the need for RPA pilots; he also indicates that the traditional manned aircraft pilot training might not be a sufficient template for RPA pilots. This analysis utilized the Air Force Officer Qualifying Test (AFOQT) pilot and the Pilot Candidate Selection Method (PCSM) composite scores as predictors and graduating Undergraduate RPA Training (URT) completion as the outcome. The dataset contains 139 officers between 2009 and 2012. Caretta (2013) examined the correlations of the predictors and URT completion. The results confirmed a sufficiently strong correlation from both predictors, using Hotelling's t-test. This study is of particular interest because PCSM score is a predictor in our dataset for the 558th Flying Training Squadron. Caretta, joined by Rose and Barron (2015), also did a study in 2015 on the enlisted counterparts, the sensor operators. While this is not directly about RPA pilots, it is related to the people that work closely with them. In addition to this new insight, the authors use a different regression technique, stepwise regression analysis, which could aid in the analysis of our data.

Though this may seem like a specific problem, there are a few different strategies to go about these analyses. Jenkins (2021) utilizes a machine learning strategy to mitigate the waste of time and resources spent on dropouts in RPA training. Jenkins examines specialized undergraduate pilot training (SUPT) candidate data and the randomized tree machine learning technique and has shown to achieve 94% accuracy in predicting candidate success. His results applied to our project very heavily because he found that the type of degree attained by candidates and their commissioning source were the most influential factors when attempting to determine success in the initial stages of training. This provides the team with different perspectives on potential strategies to be implement as well as numerous factors to keep in mind when cleaning and selecting data.

## 1.3 Data

The 558th FTS provided data in the form of PDFs documenting all students that graduated Initial Flight Training (IFT) and were moving onto URT. Relevant data from these PDFs were transferred to an excel spreadsheet with one row afforded to each unique student. Additionally, we received excel files separated by class documenting students' final URT rank. These excel files were combined to include all students (observations) in one excel document. To finalize our dataset, we combined the new excel files we had created into one excel file including initial student characteristics and the final URT rank of each student. The first and last names of each student serve as unique identifiers for our data.

The data most closely represents time series data because it includes observations gathered over time, where the time unit is the class indicator of an observation. Initial student characteristics that are observed in the dataset include: PCSM, flight hours, GPA, name, STEM degree status, gender, class, commissioning source, PPL status, prior status, race, and highest degree earned. Our dependent variable throughout analysis is the final URT rank score of each student, which is recorded after completion of URT.

## 2. Methodology

The team three different regressions to estimate the effect of different initial student characteristics on URT class rank. The first regression we run is ordinary least squares (OLS), which seeks to minimize the sum of squared residuals by fitting a line through all the available data. This is the most basic form of our analysis, and it is represented as equation (1).

$$R_i = \alpha + \beta C_i + \varepsilon_i \quad (1)$$

The variable,  $R_i$ , indicates a student's ranking codified as a score upon graduation from URT. Note that this regression does not include students that failed URT. The coefficient,  $\alpha$ , represents the intercept of the best fit line, and  $\varepsilon_i$  indicates the inclusion of error terms into the model. The various values of the  $\beta$  coefficient are of the most interest because they indicate which variables are initially most valuable in determining URT class rank. The variable vector,  $C_i$ , includes the following variables: GPA, PCSM score, flight hours, STEM, gender, class, commissioning source, PPL status, prior enlisted status, race, and degree type. STEM, PPL status, and prior-enlisted status are encoded as dummy variables with values of 0 or 1. Gender, class, commissioning source, race, and degree type are encoded in the data and used as dummies in the model. We find that GPA, PCSM, STEM, PPL status, and prior-enlistment status are significant regressors from equation (1), and we move forward

into establishing equation (2) with fewer independent variables of interest.

$$L_i = \alpha + \beta X_i + \Omega Q_i + \varepsilon_i \quad (2)$$

Equation (2) is a logit function that has  $L_i$  as the dependent variable representing the probability that a student graduates in the top fourth of their class. The  $\beta$  coefficient values indicate which variables are of interest that are included in vector  $X_i$  – GPA, PCSM, STEM, flight hours, PPL status, and prior status. The vector  $Q_i$  indicates variables that are included as controls – female, class, commissioning source, race, and degree type – which have the  $\Omega$  coefficients. These control variables are not expected to be significant, however, it is important to include them to prevent omitted variable bias in our analysis. The value in equation (2) comes in determining the probability that a student graduates in the top fourth of their class using the logit equation, getting  $z$  from the right-hand side of equation (2) and plugging that into the function,  $L_i$  where  $L_i = 1/(1+e^{-z})$ .  $L_i$  will result in the desired probability.

Equation (3) expands upon equation (1) to include fixed effects which serve as a greater source of control because it takes into consideration differences in classes, which is considered a unit of time since classes appear one after the other.

$$R_{ic} = \alpha + \gamma_c + \beta C_{ic} + \Omega Q_{ic} + \varepsilon_{ic} \quad (3)$$

Like  $R_i$  of equation (1),  $R_i$  of equation (3) represents the estimated rank score of an individual graduating from URT. The control variables from equation (2) are included as such in equation (3), and the  $C_{ic}$  variables remain the variables of interest, focusing on the  $\beta$  coefficients. The main addition to equation (3) that makes it our preferred model is the introduction of time fixed effects. The variable,  $\gamma_c$ , represents time-fixed effects, relying on class as the unit of time. This is a reasonable inclusion because classes appear one after the other, with no two classes training simultaneously. The presence of these fixed effects controls for variation across classes that might impact our estimates by inspecting each class individually prior to estimating regression results.

## 2.1 Assumptions

Primarily, we assume that our data as given is accurately surveyed. Much of the data is self-reported by individual students, which brings concerns of measurement error. However, these students are commissioned officers and have little incentive to misrepresent any data points as they are already accepted into the RPA pipeline. We are assuming that this historical data is random enough to predict future student performance - we assume that students are randomly distributed across classes, and this random distribution will continue into future classes, additionally assured by client.

We assume that missing ranks indicate students that either failed or dropped out of URT. For other missing data values, we keep observations in our dataset. If a student has a name, they will be included in analysis, but regression techniques omit observations that lack values of independent variables used in the model.

Regarding the chosen models, there are several assumptions we make that enable us to perform analysis on the data. For all OLS models, we assume our estimations are linear in parameters, we have random sampling of observations, the conditional mean is zero, there is no multi-collinearity, and spherical errors.

## 2.2 Descriptive Statistics

The data comes from the 558th FTS, documenting initial student data of incoming RPA pilots transitioning to the RPA Instrument Qualification (RIQ) portion of URT. We observe variables including the following: final class score, PCSM scores, flight hours, GPA, STEM degree indicator, gender, class number, commissioning source, private pilot's license status, prior enlistment status, race, and type of degree. From this information, we create a pass variable to indicate whether a student passes RIQ and a variable for high rank, which indicates if a student graduated RIQ in the top fourth of their class. Note that categorical variables are still modeled as dummies, but their meaning is better understood as categories.

It is immediately clear that this data is unbalanced, leaning heavily toward Caucasians, males, individuals with a bachelor's degree, and OTS graduates. This information is taken into consideration when analyzing the results of our model. If a variable appears to be significant, but it is rarely observed, we will attribute the significance to poor sample variation and assume that it is not relevant in determining the final URT class rank as advised by project advisor.

## 2.3 Model Inputs

Model 1 is an OLS regression estimating URT final rank score. The independent variables used to estimate the dependent variable are as follows: GPA, PCSM, flight hours, STEM as a dummy variable, gender as a dummy variable, class as a dummy variable, commissioning source as a dummy variable, PPL status as a dummy variable, prior enlistment status as a dummy variable, race as a dummy variable, and degree type as a dummy variable. Each estimation represents an individual because data is gathered on the individual level.

Model 2 is a logit regression estimating the probability of a student earning a high rank, being in the top fourth of their URT graduating class, a dummy variable equal to 1 if true and 0 if false. The independent variables used to estimate the dependent variable are as follows: GPA, PCSM, flight hours, STEM as a dummy variable, gender as a dummy variable, class as a dummy variable, commissioning source as a dummy variable, race as a dummy variable, degree type as a dummy variable, prior enlistment status as a dummy variable, and PPL status as a dummy variable. Variables that are not statistically significant after running this model are considered controls to prevent omitted variable bias. The independent variables are used to calculate  $z$ , which is used in the logit function,  $1/(1+e^{-z})$ , to determine the probability of a given individual graduating in the top fourth of their class.

Model 3 is an OLS regression with fixed effects estimating URT final rank score. Fixed effects cover the time entity of class because classes take place consecutively over time. The independent variables used to estimate the dependent variable are as follows: GPA, PCSM, flight hours, STEM as a dummy variable, gender as a dummy variable, commissioning source as a dummy variable, degree type as a dummy variable, prior enlistment status as a dummy variable, PPL status as a dummy variable, and race as a dummy variable. Variables that are not statistically significant after running the model are considered controls. This method controls for all factors that might vary over time between URT classes because of the presence of fixed effects.

## 2.4 Model Outputs

Our primary model outputs came from an econometric regression that outlined the influence of initial student data on their overall score. We found several variables tied URT score worth further investigation and provided a level of significance important to our study. All models provided us with insight into the different student data that affects URT final scores.

Using variables identified in the analysis phase of our model as independent variables in OLS regressions, we determine the impact that each variable has on URT class rank. We looked for statistically significant coefficients on the independent variables of interest and the signs on these coefficients, indicating their type of effect on URT class rank.

Model 1 is a simple OLS Regression with 5 significant variables seen in Table 1. For a 1 unit increase in each of these variables, the score of the trainee is expected to increase by the coefficient associated with that variable (i.e., a 1-point increase in GPA is correlated with a 3.138 increase in that student's final URT Score).

Table 1. Model 1 Significant Output

Dependent Variable: Score	Coefficient	95% Confidence Interval
GPA	3.138 {0.623}	[1.915, 4.362]
PCSM	0.096 {0.012}	[0.071, 0.120]
STEM	1.018 {0.485}	[0.067, 1.969]
PPL	-2.602 {0.893}	[-4.354, -0.850]
Prior	1.343 {0.574}	[0.215, 2.470]

\*Variables included on the left-hand side are significant at the 5% level, all other variables (Commissioning Source, Race, and Type of College Degree) are considered controls; standard error is included in brackets.

Model 2 is a logit model with only 3 significant variables. Rather than the coefficients being explanatory towards a student's score, the coefficients represent an increase or decrease in that student's  $z$ -value, which is input into the logit equation.

Table 2. Model 2 Significant Output

Dependent Variable: High Rank	Coefficient	95% Confidence Interval
GPA	0.973 {0.240}	[0.503, 1.442]
PCSM	0.028 {0.005}	[0.018, 0.037]
PPL	-1.380 {2.100}	[-2.102, -0.658]

\*Variables included on left-hand side are significant at the 5% level, all other variables (STEM Degree, Flight hours, gender, commissioning source, race, level of degree, and prior enlisted experience) are considered controls; standard error is included in brackets.

Model 3 is a fixed effects OLS Regression. It is like model 1, with coefficients, confidence intervals, and standard error being nearly identical, suggesting that fixed effects are not necessary.

Table 3. Model 3 Significant Output

Dependent Variable: Score	Coefficient	95% Confidence Interval
GPA	3.138 {0.623}	[1.915, 4.362]
PCSM	0.096 {0.012}	[0.071, 0.120]
STEM	1.018 {0.485}	[0.067, 1.969]
Prior Enlistment	1.343 {0.574}	[0.215, 2.470]
PPL	-2.602 {0.893}	[-4.354, -0.850]

\*Variables included on left-hand side are significant at the 5% level, all other variables (Flight hours, gender, commissioning source, type of degree and race) are considered controls; standard error is included in brackets.

## 2.5 Model Analysis

Based on the results obtained from all four models estimated, we conclude that there is a multitude of factors that impact success and failure in URT. The R<sup>2</sup> values of each of the models estimated in this paper do not exceed 0.2, indicating that outside of these models, there is approximately 80% of the variation in URT final rank score and probability of high rank that is left unexplained. The use of these models is beneficial to get a baseline understanding of a student's future performance in URT, but results should be viewed critically with an understanding that some students will surprise and exceed or fall below their estimated performance.

Model 1 has an R<sup>2</sup> value of 0.166, suggesting that the simple OLS regression explains 16.6% of the variation in final URT rank score. The main student characteristics driving this explanatory power are GPA, PCSM, STEM, PPL status, and prior enlistment status because they each have a p-value less than 0.05.

Model 2 has a Pseudo-R<sup>2</sup> value of 0.1162, indicating that the logit regression explains 11.62% of the variation in students' obtaining a high rank in URT. The characteristics driving this explanatory power are GPA, PCSM, and PPL status.

Model 3 has an R<sup>2</sup> value in the range of 0.1485 to 0.2027, suggesting that the OLS regression with class-based fixed effects explains 14.85% to 20.27% of the variation in the final URT rank score. The characteristics driving these results are GPA, PCSM, STEM, PPL status, and prior enlistment status – the same as those driving the explanatory power of Model 1.

The low explanatory power of our models is concerning because it shows that most of the characteristics affecting RPA pilot success and failure are unobserved, and potentially unobservable. However, the true value of these models will be shown in the trend of failure rates of URT – if employment of these models drives decisions in individualized training at RIQ and failure rates decrease over time, then these models will be proven successful. Valuable insight can be drawn from these models as showing factors that may not affect performance can be just as powerful as factors that do.

## 3. Results

In this section, we consolidate our results and provide recommendations based on our findings. We also include an explanation of our analysis to enhance our explanation of results and recommendations.

Overall, models 1 and 3 each account for about 15% of the variation in URT success, which is not ideal but still provides valuable insight. The difficult nature of predicting trainee performance in URT reflects on the low predictive power of initial student characteristics.

Across models, we anticipated commissioning source, degree type, GPA, PCSM, STEM, and PPL status to be significant regarding final student rank score. Both commissioning source and degree type contributed minimally to our results and were thenceforth classified as controls. GPA, PCSM, STEM and PPL were the most significant predictors for trainee success across models. According to our preferred model, model 1, 1-point increase in GPA accounts for a 3.14 increase in URT score. For most of the models, having achieved a STEM degree positively affects URT score as well. In model 1, a STEM degree increased URT score by 1.02 points.

Despite the presence of some significant initial student characteristics, initial student characteristics do not possess strong predictive power for student performance in URT. An increase in GPA results in a slight increase in URT score, but the relationship does not have a sharp slope.

The relationship between PCSM and final URT rank score has a slightly steeper relationship than that of GPA and final URT score. The scales of these two independent variables are different – PCSM having many more numerical possibilities. The estimated increase in final URT rank score from a 1-point increase in PCSM is 0.096. Across 99 values, the final URT rank score has the possibility of increasing by 10 points. However, there is significant variation around the best fit line, and we should be wary in drawing concrete conclusions on student success based solely off PCSM score.

The significant initial student factors - GPA, PCSM, STEM, and PPL status – found from models 1 through 3 are valuable in determining a baseline assumption for student performance in URT. However, there are a multitude of other factors that impact student performance more significantly than the initial student characteristics that are observed in our data. We call the missing source of variability in final URT rank score the “it” factor – whether this is passion, gaming skills, or something else, we are not sure. We believe the “it” factor to be a crucial missing piece in our analysis, and we urge the 558th to evaluate their trainees in attempt to discover this potential driving force of success.

#### **4. Conclusions and Future Research**

In this section, we consolidate our results and provide recommendations based on our findings. We also include a brief explanation of our analysis to enhance our explanation of results and recommendations.

Ultimately, GPA, PCSM, STEM and PPL status are the most significant initial student characteristics that impact trainee success because of their consistent presence as significant variables in all models. URT student success and failure coincides with their status in these areas - characteristics that instructors can use as identifiers for students that may require more individualized training prior to the beginning of training. Use of these models will enable URT instructors to get a head start on struggling students and decrease overall failure rates in URT.

Additionally, after speaking with a recent high-performing graduate of URT as well as the Director of Operations (DO) and squadron Innovations officer we confirmed the importance of the “it” factor in RPA training success. Variables accounting for motivation, athleticism, ability to handle stress/criticism, ego, competitiveness, situational awareness, and overall well-roundedness would be incredibly predictive of student success in URT. Unfortunately, there are no true numerical values we can use to measure these effects. In the future, it would be beneficial to investigate these potential “it” factors to determine the driving force of success at URT. In the meantime, our analysis makes it clear that anyone can succeed in URT.

Future research will discuss how RPA training is outdated and will be revamped in the next ten years to be more specialized. With this baseline research, the future holds the possibility to not only predict student success and failure, but to build a living model to project how students are going to perform as training progresses, adapting every day based on student training factors and historical performance. When this happens, additional analysis must be conducted to account for the newness of these future training techniques, and analysts will find numerous factors to be influential in URT student success.

## 5. References

- Carretta, T. R. (2013). Predictive validity of pilot selection instruments for remotely piloted aircraft training outcome. *Aviation, Space, and Environmental Medicine*, 84(1), 47–53. <https://doi.org/10.3357/ase.3441.2013>
- Carretta, T. R., Rose, M. R., & Barron, L. G. (2015). Predictive validity of UAS/RPA Sensor Operator Training Qualification measures. *The International Journal of Aviation Psychology*, 25(1), 3–13. <https://doi.org/10.1080/10508414.2015.981487>
- Jenkins, P. R., Caballero, W. N., & Hill, R. R. (2022). Predicting success in United States Air Force pilot training using Machine Learning Techniques. *Socio-Economic Planning Sciences*, 79, 101121. <https://doi.org/10.1016/j.seps.2021.101121>