EMBRY-RIDDLE Aeronautical University
SCHOLARLY COMMONS

International Journal of Aviation, Aeronautics, and Aerospace

Volume 4 | Issue 1 Article 3

1-9-2017

Estimating Airport Operations at General Aviation Airports Using the FAA NPIAS Airport Categories

Mary E. Johnson
Purdue University, mejohnson@purdue.edu
Yue Gu
Purdue University, gu132@purdue.edu

Scholarly Commons Citation

Johnson, M. E., & Gu, Y. (2017). Estimating Airport Operations at General Aviation Airports Using the FAA NPIAS Airport Categories. *International Journal of Aviation, Aeronautics, and Aerospace, 4*(1). https://doi.org/10.15394/ijaaa.2017.1151

The number of annual aircraft operations (take-offs and landings) is a significant concern to airport owners and operators, state and local agencies, and Federal agencies in the United States. This operations data is used when developing airport master planning, conducting airport environmental research, forecasting economic impact, adjusting funding, measuring aviation performance. Operation counts are reported on the FAA Airport Master Record Form 5010 (FAA, 2016). In the United States, there are nearly 20,000 non-towered airports, and approximately 500 towered airports (FAA, 2016). Aircraft operations are counted at towered airports during tower hours. Aircraft operations counts at non-towered airports are estimated based on sample counts or other methods.

To find a reliable method for estimating aircraft operations at non-towered airports, several types of research have been conducted. In general, these methods that are being used can be classified into five types, including (1) asking the airport manager or personnel, (2) counting the aircraft activities year-round, (3) expanding a sample count into annual aircraft operations, (4) computing a ratio of IFR flight plans filed to total operations (IFPTO), and (5) multiplying the number of based aircraft at an airport with a predetermined number of operations per based aircraft (OPBA) (Muia, 2007; Muia & Johnson, 2015). In addition to these five estimating methods, three models for estimating the annual operations (OPS) have been developed using a sample of small airport data. These models have 29 variables (Black and Chimka, 2011), seven variables (Hoekstra, 2000), and eight variables (GRA, 2001) and use data such as the number of based aircraft and geographic location, the population in the surrounding area, and characteristics of surrounding airports. However, no method has been developed can take into account both economy and accuracy from the viewpoint of the airport manager. The OPBA method is attractive to airport managers because it is easy to do using the readily available number of based aircraft at the airport.

In 2012, the FAA published the *General Aviation Airports: A National Asset* and released criteria for categorizing general aviation (GA) airports into four new categories: National, Regional, Local, and Basic. The GA airports include general aviation airports, heliports, and seaplane bases. In 2014, new classification categories were used for 2,939 public-use general aviation airports in the United States and published in the *National Plan of Integrated Airport Systems (NPIAS) 2015-2019 Report* (FAA, 2014). These categories were revised and amended to include another category called ‘unclassified’ to describe airports that are in the *NPIAS* but do not fit into the four other categories (FAA, 2014). These *NPIAS* categories of GA airports have not been used in the previously published studies found during this research, and provide an opportunity to model and potentially improve an estimating method of annual operations at GA non-towered airports.

Literature Review

In Muia (2007), most respondents to the study's survey that distributed to all 50 state aviation agencies, selected airports, and metropolitan planning agencies (MPOs) reported that they simply ask the airport manager or personnel to estimate the number of the annual aircraft operations. However, this method is also the most inaccurate one among five methods studied in Muia (2007). Muia (2007) noted that in the *1994 Aircraft Activity Counter Report* of Texas Department of Transportation, the airport managers could greatly underestimate the aircraft operations for their airport compared to using the method that expands a sample count into an estimate of annual aircraft operations (Muia, 2007).

The most accurate method currently in use is to count all aircraft activities at an airport year-round, but this approach is also costly and time-consuming. While towered airports can obtain the counts of operations that occur during tower operating hours, tower counts are not available for non-towered airports. Also, technologies such as acoustic or pneumatic counters and video cameras may be used all year, but are costly to operate and to post-process to obtain annual operations counts (Muia & Johnson, 2015). As a substitute, a method that expands the number of sample counts into an estimate of annual aircraft operations may be used to decrease the cost and time but may lose accuracy and precision. A sampling method may use the same counting technology that may be utilized when counting the aircraft activities year-round, such as acoustical counters, airport guest logs, fuel sales, pneumatic counters, video image detection counters, and visual observation counters, to sample aircraft activities for one to two weeks in each of four seasons and then extrapolate the sample into an annual estimate of operations (Muia & Johnson, 2015). Full details of the statistical extrapolation of sample counts into annual operations estimates are found in Ford and Shirack (1985). However, the statistical extrapolation of sample counts is also considered costly and time-consuming. (Muia, 2007; Muia & Johnson, 2015).

The IFPTO method is a relatively simple method that uses a ratio of instrument flight plans filed (IFP) to total operations (OPS). The IFPTO ratio is calculated by dividing an average annual number of total IFR flight plans by the mean annual number of total General Aviation operations in the same year at small towered airports. This IFPTO is then used with the IFP filed at a non-towered airport to estimate annual operations. Although it is an economical means to estimate OPS, this method as implemented in the 2015 study was shown to lack consistency (Muia & Johnson, 2015).

The final method used to estimate the aircraft operations is multiplying based aircraft of an airport with a predetermined number of operations per based aircraft. In the FAA Order 5090.3C, *Field Formulation of the National Plan of Integrated Airport Systems (NPIAS),* the FAA provides operations per based aircraft (OPBA) factors for four different categories of airports: rural GA airports, busier GA airports, busy reliever airports, and airports in unusual circumstances (FAA, 2000). The OPBA numbers are shown in Table 1.

Table 1

Airport Categories and Corresponding NPIAS OPBA Numbers

Airport Category	OPBA
Rural GA airports with little itinerant traffic	250
Busier GA airports with more itinerant traffic	350
Busy reliever airports	450
Unusual circumstances (e.g., busy reliever with high itinerant operations).	750

Note. OPBA is Operations Per Based Aircraft. The OPBA of the different categories is from FAA (2000).

In addition to these factors and sample count methods, annual general aviation operations (OPS) estimates and operations per based aircraft (OPBA) estimates were modeled by regression analysis for a sample of non-towered and towered airports. There were four previous research efforts found that attempted to develop an effective regression model for estimating GA operations at non-towered airports by using data from small, towered airports (Hoekstra, 2000; GRA, 2001; Black and Chimka, 2011; Muia and Johnson, 2015). While Hoekstra, GRA, and Black and Chimka used OPSBA for general aviation operations per based aircraft, Muia and Johnson used OPBA for that term. In this paper, OPBA is used for operations per based aircraft in all cases, even when the original authors used OPSBA. OPBA includes both GA local and GA itinerant operations.

Three studies used the same data for OPS and BA (Hoekstra, 2000; GRA, 2001; Black and Chimka, 2011). Using fiscal year 1999 Terminal Area Forecast data (TAF) for annual operations and based aircraft, Hoekstra sampled data from 127 small towered airports to develop a regression model of OPS and OPBA using stepwise regression and reported R^2 values for an OPS regression equation of 72.96% and OPBA regression equation of 25.56% (Hoekstra, 2000). Hoekstra's equations are list in Table 2. These equations were tested by comparing the GA annual operations estimated by the regression equations versus estimates based on sample counts that were extrapolated to annual counts for 129 small non-towered airports in nine states (Hoekstra, 2000). The equations "tended to produce higher annual operations estimates than the state estimates at these airports" (GRA, 2001, p. 3). In GRA (2001), the researchers improved the Hoekstra's OPS and OPBA regression equations by adding independent variables for airport characteristics, demographics, and geographic features to the existing Hoekstra model.

The GRA models for OPS and OPBA shown in Table 2 were created based on a dataset containing GA operations at a sample of 127 small towered and 105 non-towered GA airports, with a towered airport having a value of 1 in a dummy variable and non-towered airports having a value of 0. These are the same airports used by Hoekstra except that 24 non-towered airports were removed because they were not in the TAF and the population data could not be calculated (GRA, 2001). The GRA models were developed using GA airports in the Terminal Area Forecast (TAF), flight school data from the FAA Flight Standards Service's VITALS database containing Part 141 schools, population within 100 miles, 50 miles, and 25 miles from each airport from U.S. Census, and some categorical and geographical of these GA airport from TAF (GRA, 2001). Table 3 identifies the variables in the models shown in Table 2. It is important to note that the models use the square of based aircraft, thereby incorporating a polynomial in the model to achieve an improved fit based on R^2 values reported.

In GRA (2001), the researchers did not deeply pursue the OPBA estimating equation since the R^2 of the model is much lower than the R^2 for the OPS model, as shown in Table 2. The researchers also stated that for many of the non-towered airports, the "based aircraft may be a more reliable estimate than annual operations" (p.19-20).

The variables in Table 3 represent different characteristics of airports. Some of the data required in equations are easy to collect such as the number of based aircraft, towered airport or non-towered, the number of flight schools, or FAA region. Some of the variables are harder to obtain estimates for, such as the ratio

DOI: https://doi.org/10.15394/ijaaa.2017.1151

of the population within 25 miles to the population within 100 miles, the numbers of based at all GA airports within 100 miles, and personal income in the county.

Table 2

Regression Equations for Estimating General Aviation Operations at Non-Towered Airports

Source	Equation	R^2
Hoekstra	OPS = 813.5 + 417 BA + 0.80 PCI –0.63 BA2 – 11,683 WST – 21,752 AAL –7,072 FAR139 + 4.0 EMP	72.96%
Hoekstra	OPBA = 581.3 – 138.5 BAE100 – 125.9 WST – 326.1 AAL + 113.1 R12	25.56%
GRA	OPS = -571 + 355 BA – 0.46 BA2 – 40,510 %in100mi + 3,795 VITFSnum + 0.001 Pop100 - 8,587 WACAORAK + 24,102 Pop25/100 + 13,674 TOWDUM	74.30%
GRA	OPBA = 595.2 – 164 BAE100 – 325 %BA100 - 107 WST – 244 AAL + 0.00002 Pop100 – 0.00002 Pop50	30.72%

Note. The equations and R^2 of estimating model from Hoekstra (2000) and GRA (2001).

In 2011, Black and Chimka recreated the GRA study and redeveloped OPS models based on the same airports and data used by GRA. The study added explanatory variables, including second order terms. The equation using the full model with all 29 variables showed a modest improvement in R^2_{adj} (77.5%) compared to the recommended reduced model of eight variables developed by GRA with an R^2_{adj} (72.9%). R^2_{adj} was not published in the GRA report but was provided by Black and Chimka (2011) in their recreation of the GRA study.

Using the approximately 3,300 airports in the 2010 Terminal Area Forecast, Li and Trani (2014) developed two translog regression models for forecasting the GA local and GA itinerant demand at the airport-level. The explanatory variables in the regression models include airport characteristics and social-economic and demographic factors of each county where the airports are located. The R^2_{adj} for estimating GA itinerant demand is 41.2%, while the R^2_{adj} is 53.1% for estimating GA Local demand. In the study, they showed that the "relative fuel price – fuel price compared with disposable income per capita" – is a significant determinant of airport level GA demand (Li & Trani, 2014, p 2). Further, the demand of GA local and GA itinerant will respectively decrease 4.3% and 5.2% when the relative fuel price increases 10%.

Table 3

Variables Used in GRA's Regression Equations

Variable	Description
OPS	Annual GA Operations at an airport
OPBA	Annual GA Operations per Based Aircraft (BA) at an airport
BA	Total Based Aircraft at an airport
BA^2	Total Based Aircraft squared
%in100mi	Percentage of based aircraft among based aircraft at GA airports within 100 miles
VITFSnum	Number of FAR141 certificated pilot schools on airport
Pop100	Number of people within 100 miles
Pop50	Number of people within 50 miles
Pop25	Number of people within 25 miles
Pop25/100	Ratio of Pop25 to Pop100
TOWDUM	1 if towered airport, 0 otherwise
WACAORAK	1 if state is CA, OR, WA, or AK, 0 otherwise
BAE100	1 if airport-based aircraft is 100 or greater, 0 otherwise
%BA100	Term used interchangeably with %in100mi
WST	1 if airport is located in FAA Western Region, 0 otherwise
AAL	1 if airport is located in Alaska, 0 otherwise
R12	Located in FAA Region ANE or AEA
PCI	Personal Income of County (Billions of $)
FAR139	Certification for Carrier Service
EMP	Non-Agricultural Employment of County

Note. The variables and their descriptions from GRA (2001), except GRA used OPSBA where OPBA is employed in this paper.

In Muia and Johnson (2015), the researchers used regression analysis to test the consistency of OPBA estimates and if OPBA is affected by climate, population and the presence and number of flight schools. The TAF operations data were used. The models were developed using a dataset with information on small towered airports because there are no tower counts available for non-towered airports. The selection of the 205 small, towered airports in the study was based on three criteria: non-hub public use airport with FAA VFR tower or FAA contract tower, fewer than 10,000 enplanements annually, and fewer than 730 air carrier operations per year. Using TAF data on small towered airports as a proxy for non-towered airports, regression models were developed using best subsets regression and principles of parsimony, such selecting models with the fewest number of variables and highest the R^2_{adj}. The statistics considered an R^2_{adj}, Mallow's C_P and S. Mallow's C_P is an indicator of possible multicollinearity. *S* is the standard deviation of the model; or model error. If there were two or more models with very close R^2_{adj}, then model selection also considered the accessibility of the data to a typical GA airport manager. For instance, on the OPS reduced model with five variables, the R^2_{adj} is 65.3%; for the full model with 14 variables, the R^2_{adj} was 64.6%. Five variables are easier to access by the typical airport manager as compared to 14 variables. Both full and reduced regression models for estimating OPBA and OPS were developed. However, the OPBA method was shown to be not practical and consistent in the context of that study. Importantly, in that modeling effort, after the variables OPBA and BA were shown to have a non-linear relationship, the variables were transformed using $\log_{10}$ to linearize the model. The resulting R^2_{adj} was 50.4% for log10 OPBA for the reduced regression model. The R^2_{adj} was 65.3% for OPS in the reduced regression model with five variables.

Airport Data Sources

The models for operations estimates rely on accurate operations and based aircraft data. Accurate operations counts are problematic for non-towered airports. Two sources of operations and based aircraft data are the FAA Form 5010 and the TAF. However, each of the sources has weaknesses. The data in FAA Form 5010 are reported by airport managers. The methodology for estimating the number of operations is not consistent and may have inaccuracies, while the BA data is considered accurate (Muia, 2007). The operations data reported in TAF also have the same problem because the "operations at non-FAA airports are taken from FAA Form 5010 reports" (FAA, n.d, p. 3.). Since non-FAA airports are those airports without a tower, the researchers selected Form 5010 as the source for this study.

Parsimony and Regression Validity

In Vandekerckhove, Matzke, and Wagenmakers (2015), parsimony is defined as a principle of model selection states that "all other things being equal, simpler models should be preferred over complex ones, or greater model complexity must be bought with greater explanatory power" (p. 24). By parsimonious, the fewest number of explanatory variables are used. Also, given a choice of equally explanatory value the variables that are easiest to collect are used.

According to Moore (2000), the assumptions required for valid regression inference are 1. For any fixed value of explanatory variables, the response varies according to a normal distribution; 2. The observations are independent of each other; 3. The mean response has a straight-line relationship with the explanatory variable; 4. The standard deviation of responses is the same for all values of explanatory variable – constant variance.

General Aviation Airport Categories

According to the FAA Modernization and Reform Act (2012), a general aviation airport is “a public airport that is located in a State and that, as determined by the Secretary does not have scheduled service, or has scheduled service with less than 2,500 passenger boardings each year” (p. 26). GA airports share the same aeronautical functions serving public interest included aeromedical flights, remote population/island access, air taxi/charter services, self-piloted business flight, agricultural support, tourism and access to special events, and other different types of general aviation services (FAA, 2012a).

In 2012, the FAA published the report of *General Aviation Airports: A National Asset* and revealed new categories criteria which is developed according to the differences of the numbers of based aircraft, including both jets and propeller aircraft, at airports. As shown in Table 4, GA airports are divided into four categories which are National, Regional, Local, and Basic GA airport. Each type of airport has its average based aircraft number and serves different functions. According to the *National Plan of Integrated Airport Systems* (*NPIAS*), there are 2553 GA airports in the United States, which includes 16 National airports, 276 Regional airports, 1152 Local airports, 874 Basic airports, and 235 airports which cannot be classified (FAA, 2014). This set of new categorization criteria may provide an opportunity to improve the method that multiplying based aircraft of an airport with a predetermined number of operations per based aircraft (OPBA), by generating an accurate, efficient, and economical estimation method for aircraft operations at non-towered airports.

DOI: https://doi.org/10.15394/ijaaa.2017.1151

Table 4

Number of US airports in General Aviation Airport Categories

Number of Airports[a]	*NPIAS* Airport Category[b]	Based Aircraft[b]	Level of Activity[b]
16	National	Averaging about 200 total based aircraft, including 30 jets	Very High
276	Regional	Averaging about 90 total based aircraft, including 3 jets	High
1,152	Local	Averaging about 33 based propeller-driven aircraft and no jets	Moderate
874	Basic	Averaging about 10 propeller-driven aircraft and no jets	Moderate - Low
235	Unclassified		

Note. [a] The Number of the airports in each category are from. FAA (2014). [b] The airport categories and their criteria are from FAA (2012a).

Research Questions

Airport managers are required to report operations using the FAA Form 5010. Airports with towers have air traffic control (ATC) personnel that count OPS during tower hours. Non-towered airports must rely on other data to develop estimates of operations. The demands on airport managers' time are such that preclude the managers from counting every operation as it occurs. GRA (2001) is shown in the FAA Aviation Data and Statistics website, as the method for estimating operations at GA non-towered airports. The FAA APO-85-7 describes the method to expand operations sample counts into a statistical annual estimate (Ford & Shirack, 1985). This sample counting is taken for two weeks each season. A total of eight weeks of sample counting is time-consuming and costly. Therefore, an easy to use and easy to understand estimation method is needed. The GA airport categories developed in 2012 provides a rationale to analyze groups of airports having similar characteristics. To develop an effective and efficient method, the three research questions were posed.

RQ1: Are BA, OPS, and OPBA significantly different in the five *NPIAS* categories for the general aviation airports?

RQ2: Are the *NPIAS* general aviation airport categories helpful for estimating the OPBA of the U.S. GA airports?

RQ3: Are the *NPIAS* general aviation airport categories helpful for estimating the OPS of the U.S. GA airports?

Method

Data

The list of GA airports and their identities, cities, and states located, airport categories based on the new categories were collected from *NPIAS* database in July 2016 (FAA, 2014). The data of towered or non-towered, the number of based aircraft, and GA operations including GA local and GA itinerant, at an airport, were collected from the FAA Form 5010 in July 2016 (FAA, 2016). According to the FAA *NPIAS* report (2014), there are 2,553 GA airports in the United States, while only 2,544 of these airports could be found in the FAA Form 5010 database.

Additionally, the data of the number of flight schools at an airport were collected from the AOPA Training and Safety database in July 2016 (AOPA 2016). The flight schools listed on AOPA website included both Part 61 and Part 141 schools. There were 770 airports with at least one flight school of the 2,544 airports.

The data collected were conditioned according to two criteria that both must be satisfied: 1) airports with at least one based aircraft are included, and 2) airports with at least one GA operation. Otherwise, the airports were excluded. The conditioned data of 2,284 GA airports were used in this study.

Procedure

In this study, a critical $\alpha = 0.05$ was selected by the researchers. The procedure to answer the three research questions was:

1. Collect data from the current FAA *NPIAS* and Form 5010 reports, and AOPA website.
2. Condition the data to include airports with at least on GA operation and at least one based aircraft.
3. Prepare descriptive the data summaries for BA, OPS, and OPBA of GA towered and non-towered airports in different categories.

4. To answer RQ1, compare BA, OPS, and OPBA of GA airports in the different *NPIAS* categories by using ANOVA, and Tukey Pairwise Comparisons test as appropriate.
5. Using the variables shown in Table 5, perform statistical analysis to answer RQ2.
    a. Perform Best Subsets Regression analysis to create a regression model for estimating OPBA at GA towered and non-towered airports.
    b. Determine a parsimonious regression model that can accomplish the desired level of explanation or prediction.
    c. Determine the validity and explanatory value of the chosen regression model by examining the *p*-value, R^2_{adj} of the model, and the compliance with the assumptions for regression inference.
    d. Transform the data if the created model violates the assumptions necessary for a valid model, and then repeat the a – c in this step.
6. Using the variables shown in Table 5, perform statistical analysis to answer RQ3.
    a. Perform Best Subsets Regression analysis to create a new regression model for estimating OPS at GA towered and non-towered airports.
    b. Determine a parsimonious regression model that can accomplish the desired level of explanation or prediction.
    c. Determine the validity and explanatory value of the chosen regression model by examining the *p*-value, R^2_{adj} of the model, and the compliance with the assumptions for regression inference.
    d. Transform the data if the created model violates the assumptions necessary for a valid model, and then repeat the a – c in this step.
7. Evaluate results and report conclusions.

Results

Data Analysis

Data was collected for GA airports using the sources in Table 5 and divided into towered and non-towered and into the five categories *NPIAS* categories. Summaries of the data for number of based aircraft (BA), annual operations (OPS) and operations per based aircraft (OPBA) at U.S. GA airports is shown in Table 6, Table 7, and Table 8. For all airport categories with sufficient *n*, the *p*-values resulting from the test for normality using the Anderson-Darling statistic were less than α = 0.05; therefore, the null hypothesis that the data are from a Normal distribution is rejected.

Table 5

Variables and Descriptions of Sources

Variable	Description	Source
OPS	Annual GA Operations at an airport	FAA Form 5010
OPBA	Annual GA Operations per Based Aircraft (BA) at an airport	OPS/BA
BA	Total Based Aircraft at an airport	FAA Form 5010
BA*2	Total Based Aircraft squared at an airport	*N/A
Log10_OPS	Log base 10 of OPS	*N/A
Log10_OPBA	Log base 10 of OPBA	*N/A
Log10_BA	Log base 10 of BA	*N/A
FS Y/N	1 if flight school (both FAR 61 and FAR 141) is present, 0 otherwise	AOPA
NFS	Number of FAR61 and FAR 141 certificated pilot schools on airport	AOPA
TOWDUM	1 if towered airport, 0 otherwise	FAA Form 5010
National	1 if airport is in National category, 0 for other categories	FAA NPIAS
Regional	1 if airport is in Regional category, 0 for other categories	FAA NPIAS
Local	1 if airport is in Local category, 0 for other categories	FAA NPIAS
Basic	1 if airport is in Basic category, 0 for other categories	FAA NPIAS
Unclassified	0 for other categories	FAA NPIAS

Note. *N/A in the column of Source indicates that this particular type of information is calculated using other variables.

Table 6

Summary of Based Aircraft (BA) at U.S. GA Airports

GA Airport Categories	*n*	Mean	*SD*	95% of CI of Mean
Towered				
National	14	170.0	93.7	116.0 - 224.3
Regional	64	91.9	60.5	76.8 - 107.0
Local	14	53.7	31.1	35.7 - 71.7
Basic	4	14.8	9.3	0.0 - 29.5
Unclassified	1	8	N/A	N/A
Non-Towered				
National	2	12.0	N/A	N/A
Regional	210	76.8	52.6	69.7 - 84.0
Local	1,126	36.6	28.7	35.0 - 38.3
Basic	696	12.8	12.8	11.9 - 13.8
Unclassified	153	9.2	15.2	6.8 - 11.6

Table 7

Summary of General Aviation Operations (OPS) at U.S. GA Airports

GA Airport Categories	*n*	Mean	*SD*	95% of CI of Mean
Towered				
National	14	61,663	46,132	35,027 - 88,299
Regional	64	46,837	37,163	37,554 - 56,120
Local	14	42,759	42,300	18,336 - 67,182
Basic	4	11,619	14,394	-11,285 - 34,523
Unclassified	1	4,570	N/A	N/A
Non-Towered				
National	2	7,622	N/A	N/A
Regional	210	37,497	25,791	33,988 - 41,005
Local	1,126	16,764	14,811	15,898 - 17,630
Basic	696	6,949	7,364	6,395 - 7,502
Unclassified	153	5,290	7,165	4,151 - 6,428

Note. The information of GA towered unclassified airport and GA non-towered public airport are not listed because of the lack of data for analyzing.

The statistics for the one GA towered Unclassified airport, and the two GA non-towered National airports are not shown due to an insufficient number of data. The airport identifier codes of these airports are GOV, UTA, and ENV. The Grayling Army Airfield Airport (GOV) is a military and public use airport in Crawford County, Michigan. It is the only Unclassified GA airport with a control tower and is owned by United States Army. The Tunica Municipal Airport (UTA) and Wendover Airport (ENV) are non-towered GA National airports (FAA 2014; FAA 2016).

RQ1: Are BA, OPS, and OPBA significantly different in the five *NPIAS* categories for the general aviation airports? To answer the RQ1, the researchers conducted the one-way ANOVA by *NPIAS* categories to compare the data of BA, OPS, and OPBA. The results of the ANOVA showed that at least one of the true means of BA, OPS, and OPBA in different *NPIAS* categories are significantly different (p-value < 0.001, p-value < 0.001, p-value < 0.001). To discover which categories are different from other categories, the researchers performed the Tukey Pairwise Comparisons tests. As shown in Table 9, the BA in each *NPIAS* categories is significantly distinct from the other categories except for Basic and Unclassified which cannot be distinguished from each other. The OPS data in National, Regional, and Local *NPIAS* categories are significantly different from each other and different from Basic and Unclassified categories. In contrast, there is no significant difference between the OPS in Basic category and Unclassified category. The results of Tukey Pairwise Comparisons test for OPBA show that there are three groups (A, B, and C) that overlap each other in some categories. Local, Basic and Unclassified can be distinguished from each other. Unclassified and Regional can be distinguished from each other.

It was found that BA in the *NPIAS* categories was significantly different from each other, with the exclusion of Basic which cannot be distinguished from Unclassified. The same is true for OPS. The OPBA can be distinguished for three of the *NPIAS* categories (Local, Basic and Unclassified), and for the comparison between Unclassified and Regional.

DOI: https://doi.org/10.15394/ijaaa.2017.1151

Table 8

Summary of Operations Per Based Aircraft (OPBA) at U.S. GA Airports

GA Airport Categories	*n*	Mean	*SD*	95% of CI of Mean
Towered				
National	14	373.9	197.3	260.0 - 487.8
Regional	64	617.0	467.3	505.2 - 733.7
Local	14	783.0	444.2	526.6 - 1,039.5
Basic	4	1,351.2	1,525.1	-1,075.5 - 3,778.9
Unclassified	1	571.3	N/A	N/A
Non-Towered				
National	2	594.3	N/A	N/A
Regional	210	585.8	418.1	528.9 - 642.7
Local	1,126	540.9	616.5	504.9 - 577.0
Basic	696	703.7	1,219.4	613.0 - 794.3
Unclassified	153	1,024.5	1,412.8	799.8 - 1,249.1

Note: The information of GA towered unclassified airports and GA non-towered public airports are not shown due to the number of data points available.

Regression Analysis

RQ2: Are the *NPIAS* general aviation airport categories helpful for estimating the OPBA of the U.S. GA airports? To answer this question, a regression model was created by performing Best Subsets Regression analysis that used OPBA as the response variable and BA, BA*2, FS Y/N, NFS, TOWDUM, National, Regional, Local, and Basic as explanatory variables. After the analysis, five explanatory variables (BA, BA*2, NFS, Local, and Basic) were chosen in this Best Subsets Regression model. The regression equation is:

$$OPBA = 1100 - 9.94BA + 0.0232\ BA*2 + 63.8NFS - 279\ Basic - 272\ Local \quad (1)$$

The regression model is statistically significant (p-value < 0.001). The R^2_{adj} of 4.6% indicates a low explanatory value of the model. Figure 1, the residual plots for the OPBA Regression *Equation 1*, is used to examine the underlying statistical assumptions required for a valid ANOVA. The analysis of the residual plots indicated that the regression violated three of four assumptions, described in the introduction, which must be met to use regression. As shown in the left half of the residual plot, the mean response did not have a linear relationship with explanatory variables, and the residuals were not normally distributed. As shown in the right half of the residual plot, the dots in the graphs did not appear to be scattered in a random pattern which means the variance was not constant. To sum up, the regression model for estimating OPBA is not practical and is statistically invalid.

Table 9

Results of the Tukey Pairwise Comparisons Tests for BA, OPS, and OPBA Data

GA Airport Categories	*n*	Mean	*SD*	Group
BA				
National	16	150.38	102.62	A
Regional	274	80.32	54.76	B
Local	1,140	36.85	28.73	C
Basic	700	12.82	12.72	D
Unclassified	154	9.17	15.11	D
	OPS			
National	16	54,908	46,766	A
Regional	274	39,678	29,045	B
Local	1,140	17,084	15,662	C
Basic	700	6,976	7,485	D
Unclassified	154	5,285	7,104	D
OPBA				
National	16	401.5	203.5	A B C
Regional	274	593.1	429.4	B C
Local	1,140	543.9	615.1	C
Basic	700	707.4	1,219.7	B
Unclassified	154	1,021.5	1,402.4	A

Note. There is no significant difference between the categories that have the same group letter.

DOI: https://doi.org/10.15394/ijaaa.2017.1151

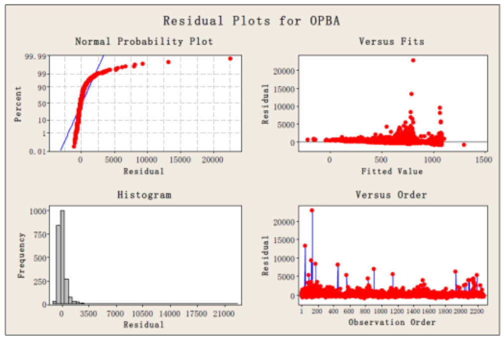

Figure 1. Residual plots for regression *Equation 1*.

There is a non-linear relationship between OPBA and BA shown in Figure 1. The researchers transformed data of OPBA and BA into log base 10 of OPBA (Log10_OPBA) and log base 10 of BA(Log10_BA) to create a linear relationship between Log10_OPBA and Log10_BA. Best Subset Regression analysis was performed to create a regression model for estimating Log10_OPBA. The explanatory variables used in this regression analysis are Log10_BA, BA*2, FS Y/N, NFS, TOWDUM, National, Regional, Local, and Basic. After the Best Subsets Regression analysis, a new regression equation was created:

$$\text{Log10\_OPBA} = 3.07 - 0.429\text{Log10\_BA} + 0.752\text{TOWDUM} + 0.0591\text{NFS} + 0.148\text{Local} + 0.189\text{National} + 0.31\text{Regional} \quad (2)$$

The regression *Equation 2* is statistically significant (p-value < 0.001). The R^2_{adj} is equal to 12.2% which is significantly increased over *Equation 1*, but still very low. In Figure 2, the residual plots for OPBA regression model, are used to examine the underlying statistical assumptions required for the regression. The analysis of the residual plots indicated that the regression satisfied all four assumptions that must be met to use the ANOVA regression. The slight non-linear relationship between the response mean and explanatory variable shown in the normal probability plot of the residuals is acceptable based on the “fat pencil test” heuristic that the points in normal probability plots are normally distributed “when all of the points in normal plot are covered by the imaginary pencil” (Rossi, 2010, p. 127). Although the new regression model for estimating Log10_OPBA is statistically valid, it is still impractical because of the low R^2_{adj}. Therefore, while the *NPIAS* categories for Local, National, and Regional contributed to a valid linear regression model, the new model does not explain enough of the variation in Log10_OPBA to be useful.

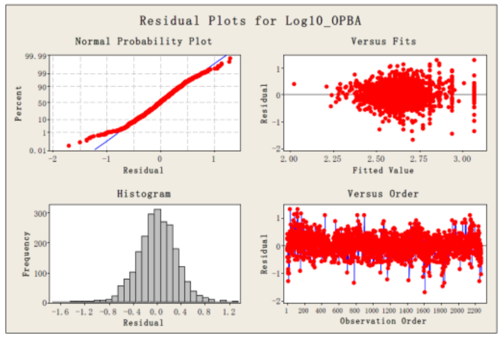

Figure 2. Residual plots for regression *Equation 2*.

RQ3: Are the *NPIAS* general aviation airport categories helpful for estimating the OPS of the U.S. GA airports? To answer this question, a regression model was created by performing Best Subsets Regression analysis where OPS is the response variable and BA, BA*2, FS Y/N, NFS, TOWDUM, National, Regional, Local, and Basic are explanatory variables. After the Best Subsets analysis, seven explanatory variables (BA, BA*2, TOWDUM, FS Y/N, NFS, Local, and Regional) were chosen in this regression model. The regression equation is listed *Equation 3*.

$$OPS = 2842 + 315BA + 0.38BA*2 + 7875TOWDUM - 6747FS\ Y/N + 7701NFS + 2454Local + 10197Regional \quad (3)$$

The regression model is statistically significant (p-value < 0.001). The R^2_{adj} is equal to 50.9% that shows a much higher level of explanatory value than *Equation 2*. The residual plots for OPS regression model are shown in Figure 3. The analysis of the residual plots indicates that the regression violated two assumptions that must be met to use ANOVA in regression. As shown in the left half of the residual plot, the mean response did not have a linear relationship with explanatory variables. Second, as shown in the right half of the residual plot, the dots in the graphs did not appear to be scattered in a random pattern which indicates that the variance was not constant. The created regression model for estimating OPS has a considerable explanatory power but is statistically invalid because of the violations of two statistical assumptions required for regression.

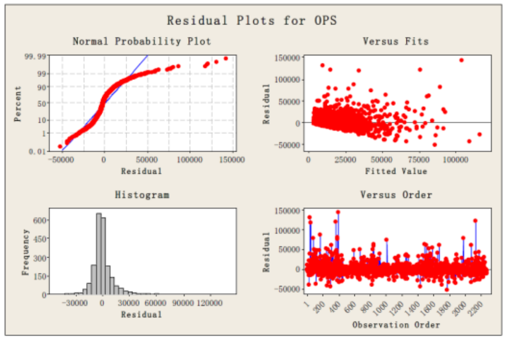

Figure 3. Residual plots for regression *Equation 3*.

To address the non-linear relationship between OPBA and BA shown in Figure 3, the researchers transformed data of OPS and BA into log base 10 of OPS (Log10_OPS) and log base 10 of BA (Log10_BA) to create a linear relationship between Log10_OPS and Log10_BA. Best Subset Regression analysis was performed again to create a regression model for estimating Log10_OPS. The variables used in this regression model analysis were as same as in the last Best Subsets Regression analysis except Log10_OPS replaced OPS and Log10_BA replaced BA. The model for Log10_OPS is:

$$\text{Log10\_OPS} = 3.07 - 0.571\text{Log10\_BA} + 0.752\text{TOWDUM} + 0.0591\text{NFS} + 0.148\text{Local} + 0.189\text{National} + 0.31\text{Regional} \quad (4)$$

The regression model is statistically significant (p-value < 0.001). The R^2_{adj} is 53% which is increased slightly over *Equation 3*. Figure 4, the residual plots for *Equation 4*, are used to examine assumptions required for the regression. The regression appears to satisfy all four assumptions for using regression except for a potential violation of the linear relationship between the response mean and explanatory variable. However, it does pass the “fat pencil test” as a heuristic estimate of normality as described in Rossi (2010). Therefore, the regression model for estimating Log10_OPS is practical and statistically valid.

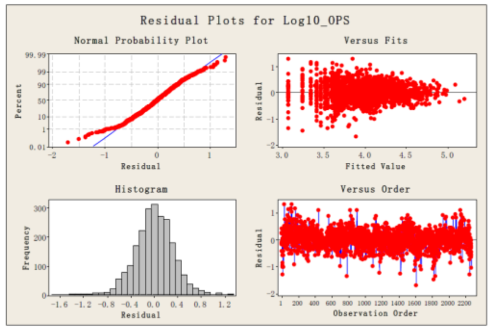

Figure 4. Residual plots for regression *Equation 4*.

However, the *p*-value for the variable TOWDUM in *Equation 4* is 0.059. In Best Subsets, the regression equations for the remaining five variables is:

$$\text{Log10\_OPS} = 3.07 - 0.571\text{Log10\_BA} + 0.0591\text{NFS} + 0.148\text{Local} + 0.189\text{National} + 0.31\text{Regional} \quad (5)$$

The regression is significant (p-value < 0.001) and the R^2_{adj} = 52.9%. The variables in the model are significant (p-value < 0.005). Residuals plots are shown in Figure 5 and are similar to Figure 4. Since the *Equation 5* is valid and contains the dummy variable that indicates the *NPIAS* categories, the use of the *NPIAS* categories have been shown to be helpful for estimating the OPS for the U.S. GA airports in this dataset.

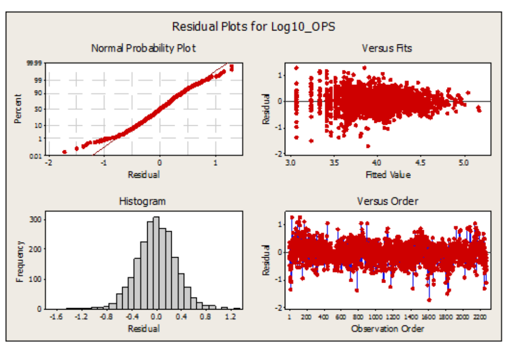

Figure 5. Residual plots for regression *Equation 5*.

DOI: https://doi.org/10.15394/ijaaa.2017.1151

Discussion

The BA and OPS in the *NPIAS* categories are different from each other, with the exclusion of Basic which cannot be distinguished from Unclassified. The OPBA can be distinguished for three of the *NPIAS* categories (Local, Basic and Unclassified), and for the comparison between Unclassified and Regional.

As summarized in Table 10, five equations were developed in this study using *NPIAS* GA airport categories, AOPA flight school data, and FAA Form 5010 airport data. The regression models were selected based on significant statistical regression and a high explanatory value with as few variables as is practical. Each of the five equations contains at least two GA airport categories indicator variables. The *NPIAS* categories were useful in developing statistically significant regression equations. Therefore, the new *NPIAS* categories are helpful for improving GA operations estimate using the Form 5010 data. Of the five equations formulated in this study, *Equation 5* has the highest explanatory value and meets regression assumptions using the fewest variables. *Equation 5* used five variables that can easily be obtained by an airport manager. Also, Table 11 simplifies *Equation 5* to be unique to each *NPIAS* GA airport category where the resulting equation needs only two data easily collected by an airport manager: number of flight schools and the number of based aircraft. While this set of equations does have a lower R^2_{adj}, it is far simpler when compared to the number and type of variables used by Hoekstra (2000), GRA (2001), and Black and Chimka (2011).

Limitations

The accuracy of the operation and based aircraft data collected from the FAA Form 5010 is a consideration. The GA operations data published in the Form 5010 were from the FAA database. As mentioned in the introduction, managers of non-towered airports may estimate the operations of the airports based on factors such as their experience or factors. Unfortunately, their estimations have been shown to differ from the actual operations in Muia (2007). Therefore, the regression models created using the operations data of Form 5010 may not present the realistic relationship between the operations and the explanatory variables. However, the TAF data also have issues of accuracy, especially for non-towered airports where the Form 5010 is used in the operations estimates.

Table 10

Equations for Estimating GA Operations at GA Airports using NPIAS Categories

Equation Number	Equations	R^2_{adj}
*1	OPBA = 1100 - 9.94BA + 0.0232BA*2 + 63.8NFS - 279Basic - 272Local	4.6%
2	Log10_OPBA = 3.07 - 0.429Log10_BA + 0.0752TOWDUM + 0.0591NFS + 0.148Local + 0.189National + 0.31Regional	12.2%
*3	OPS = 2842 + 315BA - 0.38BA*2 + 7875TOWDUM - 6747FS Y/N + 7701NFS + 2454Local + 10197Regional	50.9%
4	Log10_OPS = 3.07 + 0.571Log10_BA + 0.0752TOWDUM + 0.0591NFS + 0.148Local + 0.189National + 0.31Regional	53.0%
5	Log10_OPS = 3.07 – 0.571Log10_BA + 0.0591NFS + 0.148Local + 0.189National + 0.31Regional	52.9%

Note. * Models do not meet regression assumptions.

Table 11

Simplified Equations for GA Operations at GA Airports

NPIAS GA Categories	Equation 5 simplified for each *NPIAS* GA airport category
National	Log10_OPS = 3.259 – 0.571Log10_BA + 0.0591NFS
Regional	Log10_OPS = 3.38 – 0.571Log10_BA + 0.0591NFS
Local	Log10_OPS = 3.218 – 0.571Log10_BA + 0.0591NFS
Basic	Log10_OPS = 3.07 – 0.571Log10_BA + 0.0591NFS
Unclassified	Log10_OPS = 3.07 – 0.571Log10_BA + 0.0591NFS

https://commons.erau.edu/ijaaa/vol4/iss1/3
DOI: https://doi.org/10.15394/ijaaa.2017.1151

Conclusion

According to the analysis, the number of based aircraft and the number of operations are significantly different in the four primary *NPIAS* categories for the general aviation airports; the Unclassified category could not be distinguished from the Basic category. The operations per based aircraft for Local, Basic and Unclassified can be distinguished from each other, and Unclassified and Regional can be distinguished from each other. The *NPIAS* categories for general aviation airports are helpful in estimating GA operations using models for OPBA and OPS, but the models were not statistically valid. Transformations of the BA, OPS, and OPBA using log-based 10 were used in three models. The models using the transformed variables were statistically significant, but only the model for log 10 OPS had high explanatory value. Better regression models may be developed by gaining more accurate data or by understanding and locating any lurking variables that may be contributing to the variance of the residuals. In the aviation industry, the Terminal Area Forecast (TAF) data are believed to be more accurate than the data from Form 5010. However, the TAF data is flawed as well because it too relies on estimates and may use Form 5010 data for non-towered airports. Future research will be a focus on using the operations data for GA airports in the TAF database and for estimates other than OPBA or in addition to OPBA.

References

Aircraft Owners and Pilots Association (AOPA). (2016). *Flight school finder*. Retrieved from https://www.aopa.org/learntofly/school/index.cfm

Black, R. & Chimka, J (2011) Re-estimating and remodeling general aviation operations. *International Journal of Applied Aviation Studies, 11*(1), 47-56. Retrieved from https://www.academy.jccbi.gov/ama-800/Summer_2011.pdf

FAA Modernization and Reform Act, 49 Stat. § 132 (2012). Retrieved from https://www.gpo.gov/fdsys/pkg/PLAW-112publ95/pdf/PLAW-112publ95.pdf

Federal Aviation Administration. (2000). *Order 5090.3 C, Field formulation of the National Plan of Integrated Airport Systems (NPIAS).* Washington, DC: Federal Aviation Administration. Retrieved from https://www.faa.gov/regulations_policies/orders_notices/index.cfm/go/document.informtion/documentID/12754

Federal Aviation Administration. (2012a). *General aviation airports: A national asset (ASSET 1).* Washington, DC: Federal Aviation Administration. Retrieved from http://www.faa.gov/airports/planning_capacity/ga_study/media/2012AssetReport.pdf

Federal Aviation Administration. (2012b). *2012 ASSET Appendix A - Technical Analyses.* Washington, DC: Federal Aviation Administration. Retrieved from https://www.faa.gov/airports/planning_capacity/ga_study/media/2012AssetReportAppA.pdf

Federal Aviation Administration. (2014). *National Plan of Integrated Airport Systems (NPIAS) report.* Washington, DC: Federal Aviation Administration. Retrieved from http://www.faa.gov/airports/planning_capacity/npias/reports/

DOI: https://doi.org/10.15394/ijaaa.2017.1151

Federal Aviation Administration. (2016). *FAA Form 5010, Airport master record.* Washington, DC: Federal Aviation Administration. Retrieved from http://www.faa.gov/airports/airport_safety/airportdata_5010/

Federal Aviation Administration. (n.d.). *Forecast process for 2015 TAF.* Washington, DC: Federal Aviation Administration. Retrieved from https://www.faa.gov/data_research/aviation/taf/media/ForecastProcessfor2015TAF.pdf

Ford, M., & Shirack, R. (1985). *Statistical sampling of aircraft operations at non-towered airports* (No. FAA-APO-85-7). Washington, DC: Federal Aviation Administration. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA157095

GRA, Inc. (2001). *Model for estimating general aviation operations at non-towered airports using towered and non-towered airport data.* Washington, DC: Federal Aviation Administration. Retrieved from http://www.faa.gov/data_research/aviation_data_statistics/general_aviation/media/gamodel3f.doc

Hoekstra, M. (2000). *Model for estimating general aviation operations at non-towered airports.* Washington, DC: Federal Aviation Administration, Office of Aviation Policy and Plans.

Li, T., & Trani, A. A. (2014). A model to forecast airport-level general aviation demand. *Journal of Air Transport Management*. *40*, 192-206. Retrieved from http://www.sciencedirect.com/science/article/pii/S0969699714000957

Moore, D. S. (2000). *The basic practice of statistics.* New York: WH Freeman.

Muia, M. J. (2007). Counting aircraft operations at non-towered airports. *ACRP Synthesis of Airport Practice, 4*. Retrieved from http://www.trb.org/Publications/Blurbs/157794.aspx

Muia, M. J., & Johnson, M. E. (2015). Evaluating methods for counting aircraft operations at non-towered airports. *ACRP Report, 129.* Retrieved from http://www.trb.org/Main/Blurbs/172335.aspx

Rossi, R. J. (2010). *Applied biostatistics for the health sciences.* Hoboken, New Jersey: John Wiley & Sons.

Vandekerckhove, J., Matzke, D., & Wagenmakers, E. J. (2014). Model comparison and the principle of parsimony. In *The Oxford handbook of computational and mathematical psychology*. Retrieved from http://escholarship.org/uc/item/9j47k5q9

DOI: https://doi.org/10.15394/ijaaa.2017.1151