The financial Implications of Bird Strike to Airline and Private Operators

By

Friday Yusuf

Abstract

Bird strike is a big problem in the aviation industry, which has great financial impacts on the industry. This study analyses the cost implication of bird strike on airlines using a model linear regression, general linear regression and k-means cluster algorithm. Data was sourced from Research & Development Wing of the Federal Aviation Administration and Tundata.BirdStrikes_Data . Data was analysed using SAS visual statistics. Bird strike had the highest direct financial implication on Fedex Express ($5,130,814), but more aircraft out of service hours were lost by private owners. Bird strikes had the least financial implication ($435) and the least aircraft service hours loss on military aircrafts. This study provides more insight into the sector of the aviation industry that are affected more by airstrikes and the most suitable model for predicting future cost.

Introduction

The rate of bird strikes is increasing worldwide. Bird strikes result in bad operational effects, financial loss, aircraft damage, and fatalities. The financial implications of bird strike on the aviation industry globally, is gross, and its impact on the business and safety of the industry have been extensively recognized by civil aviation. It is estimated that approximately US$1.28 billion is lost annually from damages and delays to commercial transport aircraft operators due to airstrike by birds (Allan & Orosz, 2001). It is believed that between 1912 and 2002, there was about Fifty-five (55) fatal bird strike occurrences, resulted in 276 human fatalities and destroying 108 aircraft with a damage cost beyond US$1.2 billion (Thorpe, 2012). The risk of bird strikes is higher when aircrafts are below 3000 ft attitude (Thorpe, 2012). Therefore, commercial, private, and military aircraft are mostly endangered during departure and arrival which are the critical stage of flight that takes place within the airport environment. From 1973 and 2000, 35 crew and 42 aircrafts were lost by United States Air Force (USAF) because of birds strike (Kelly, 1999)

Linear regression is one of the most extensively used technique for analysing multifactor data. This technique became popular due to its ability to conceptually use an equation to express the relationship between a variable of interest and a set of predictor variable (Montgomery, Peck & Vining, 2012). Clustering is an unsupervised Machine Learning tool that enables the grouping of data points. Clustering provides valued insights in data sets by displaying the groups that the data points fall into (Xu, & Wunsch, 2008). Data points in the same group often have similar properties, while data points in different groups should have highly dissimilar properties.

This study analyses the cost implication of bird strike on airlines using a model linear regression, General linear regression and k-means cluster algorithm. These models acknowledge the influence of precise strike features on the cost of the strike to be observed and measured. The hypothesis that aircraft type, attitude level when the strike occurs, number of bird strikes, damage caused by the strike, impact of the bird strike on the flight, number of engines on the plane that was struck by the bird, and airline operating the flight that was struck.

Data

Reported wildlife strikes from Research & Development Wing of the Federal Aviation Administration and Tundata.BirdStrikes_Data SAS Visual Analytics data source (BirdStrikes_Data dataset) from airlines, private operators, military airports, pilots, and other sources. The full BirdStrikes_Data contains information on 39 different variables associated with each strike and 140,000 observations. Two types of costs are reported in the BirdStrikes_data: bird strike associated with damage costs and other non-damage costs such as flight, delay or schedule disruptions, flight cancellation, aircraft stoppage. Though the non-damage costs are reported not often in the BirdStrikes_data and not much is known about the extensiveness of existing reports of the no damage costs (Aaron et al.). Therefore, the analysis focus is narrow to modeling direct, bird strike-related damage costs.

Table 1 variables in the bird strike dataset

Variable	Classification	Data Type	Meaning
Type: Aircraft	Category	Character	Type of Aircraft
Altitude bin	Category	Character	Altitude level when bird strike occurred
Wildlife: Number struck	Category	Numeric	Number of birds strikes for this event
Effect: Impact to flight	Category	Character	Impact of the bird strike on the flight
Effect: Indicated Damage	Category	Character	Damage caused by the bird strike
Aircraft: Number of engines?	Category	Character	Number of engines on plane that was struck
Aircraft: Airline/Operator	Category	Character	Airline operating the fight that was struck
When: Phase of flight	Category	Character	What phase of the flight when aircraft struck
Cost: Aircraft time out of service (hours)	Measure	Numeric	Distribution of cost when aircraft time out of service
Cost: Total $	Measure	Numeric	Distribution of total cost during bird strike
Wildlife: Size	Category	Character	Frequency of wildlife size during bird strike
Wildlife: Number struck	Numeric	Category	Number of birdstrikes for this event

The analysis is focused on bird strikes related to cost, measures for the size, and the numbers of birds involved in the strike. Attentive to those variables which are correlated to the predictive value of damages.

Using SAS Visual Statistics to do the data profiling to get a thorough knowledge of the structure, quality, and contents of the BirdStrikes_data, there are significant numbers of missing observations in most of the variables. Multicollinearity problems within the variables, assessing the model with the largest available size, and the tightest set of variables seems to be a suitable plan for tackling any remaining multicollinearity.

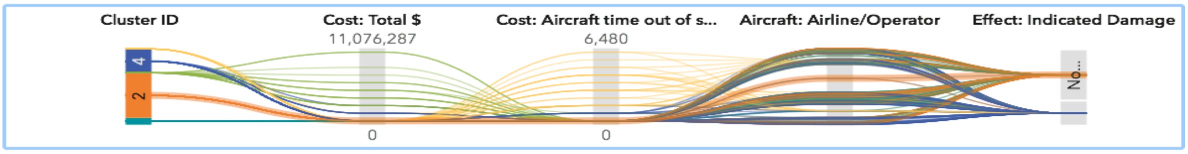
Methodology

The analysis comprises of three statistical models, cluster algorithm, linear regression, generalized linear regression (GLM). These analyses were carried out to build models for forecasting the cost or financial implication of bird strike to airline operators based on the selected predictor variables. Statistical software SAS Visual Statistic Environment was used for all data analysis.

Cluster analysis is the task of grouping a set of values in such a way that values in the same group called a cluster that is similar to each other than to those in different groups. In this analysis four variables were assigned to the model, cost total, time out of service, impact effect damage, and airline/operators. SAS visual statistic uses a k-means clustering approach by default it generates 5 clusters, and 7,383 observations were used. Eleven variables selected on data section were used for the linear and generalized linear regression, informative missingness was selected to impute the missing value in the model and 99,404 observations used.

Results

A total of $11,076,287 and 6,480 hours were lost due to bird strikes (fig 1). The cluster model shows the total cost by clustering based on time out of services, airline/operator, damage and non-damage caused, and the centroid (fig 1). Bird strike had the highest direct financial implication on Fedex Express ($5,130,814), but more aircraft out of service hours were lost by private owners (fig 1). Bird strikes had the least financial implication ($435) and the least aircraft service hours loss on military aircrafts (2 hours; fig 1). This may be because data on military airlines are not easily accessible (Thorpe, 2012). It shows that non-damage cost an airline operator about $1,903, 12 hours out of services, and cost the military operator about $435 and 2 hours out of service. The model shows the damage cost to the airline/operator, cluster 3 shows that FedEx Express suffered $5,130,814 cost of damage and 153 hours out of services. Cluster 4 shows the cost of damage for a business operator $63,290, 110 hours out of service while cluster 5 which privately owned aircraft score damage cost of $114,559, and 2,640 hours out of service.



Centroids | Cluster Summary | Model Information | Within Cluster Statistics | Iteration History | Frequency | Standardization | Interval information

Cluster ID	Cost: Total $	Cost: Aircraft time out of service (hours)	Aircraft: Airline/Operator	Effect: Indicated Damage
1	1,903	12	BUSINESS	No damage
2	435	2	MILITARY	No damage
3	5,130,814	153	FEDEX EXPRESS	Caused damage
4	63,290	110	BUSINESS	Caused damage
5	114,559	2,640	PRIVATELY OWNED	Caused damage

Centroids | Cluster Summary | Model Information | Within Cluster Statistics | Iteration History | Frequency | Standardization | Interval information

Cluster ID	Observations	Min centroid-to-observation	Max centroid-to-observation	Nearest Cluster	Centroid Distance	Average Distance	Within Distance
1	664	4.3070962E-6	0.9808641489	2	0.5007778601	0.2028489035	134.69167192
2	4,338	0.0000298159	2.278233082	1	0.5007778601	0.2075707377	900.44186018
3	28	1.4565955843	1185.0002681	4	861.36527638	190.8828064	5344.7185791
4	2,243	0.0008257686	175.01576705	1	0.6991883987	1.8333420013	4112.186109
5	110	0.1288172483	115.81875586	4	48.751721343	11.340362173	1247.439839

Figure 1. Cluster summary

Linear Regression and Generalized Linear Regression Analysis

Linear regression is the correlation between the independent input variable and dependent output variables. In this analysis “Cost: Total $” was selected as a response variable for both linear regression and generalize linear regression were added in the model as continuous or classification effect, interaction effect using impact flight effect and damage, informative missingness was selected due to the significant amount of missing values. The idea of informative missingness is one way to justify missing values in the variables. For the linear regression model the likelihood ratio test against a null hypothesis it has a p-value 0.00001 which is close to zero, settling the correctness of estimating of the model for the analysis. The linear regression model fits the data well. The p-values show that the interaction effect is statistically significant in the model with an R-square value of 0.0287.

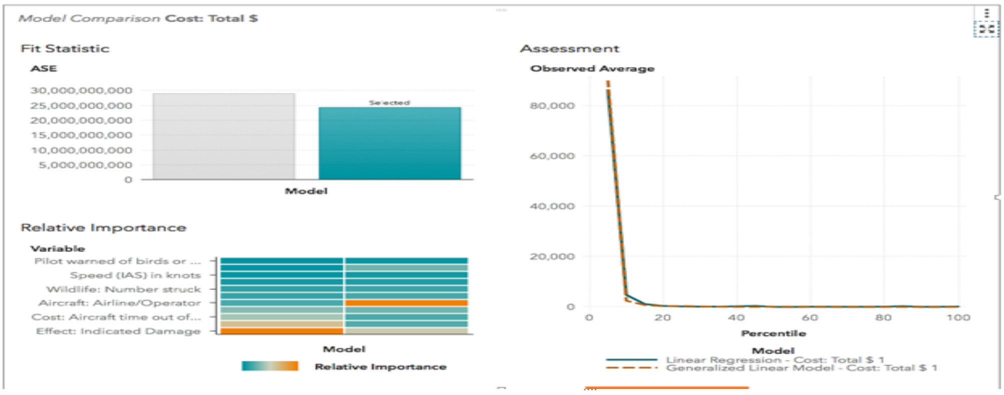
Dimensions | Overall ANOVA | Fit Statistics | Parameter Estimates | Type III Test | Assessment | Assessment Statistics

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	-46465.4	176766.4	-0.26286	0.79266
Aircraft: Type	13408.59	7599.404	1.764427	0.07766
Aircraft: Type Airplane	6603.762	6921.284	0.954124	0.34002
Aircraft: Type C	-142152	171221.9	-0.83022	0.40642
Aircraft: Type Helicopter	0	.	.	.
Aircraft: Airline/Operator ABSA AEROLINHAS BRASILEIRAS	200379.7	241458.5	0.829872	0.40661
Aircraft: Airline/Operator ABX AIR	196390.8	170970.6	1.148682	0.25069
Aircraft: Airline/Operator ACM AVIATION	148075.9	209198.3	0.707826	0.47906
Aircraft: Airline/Operator ACTION AIRLINES	200046.5	241462.3	0.828479	0.40740
Aircraft: Airline/Operator ADI SHUTTLE GROUP	191331	175312.3	1.091373	0.27511
Aircraft: Airline/Operator AER LINGUS	196308.3	175825.4	1.116496	0.26421
Aircraft: Airline/Operator AERO AIR	199603.1	209171.5	0.954256	0.33996
Aircraft: Airline/Operator AERO CHARTER AND TRANSPORT	158811.5	241480.8	0.657657	0.51076
Aircraft: Airline/Operator AERO COSTA RICA	197903.1	241462.6	0.819601	0.41245
Aircraft: Airline/Operator AERO INDUSTRIES	191107.6	241452.4	0.791492	0.42866
Aircraft: Airline/Operator AEROFLOT	199493.5	241465.4	0.826178	0.40870
Aircraft: Airline/Operator AEROLINEAS EJECUTIVAS	124874.4	209235.6	0.596813	0.55063
Aircraft: Airline/Operator AEROLITORAL	200972.1	241457.7	0.832328	0.40523

Figure 2 Linear Regression Parameter Estimates

Model Comparison

Linear regression is more suited for this analysis because it has better accuracy than the generalized linear regression (fig4).



Statistics Variable Importance Relative Importance

Selected	Model	Visualization Type	Number Of Observations	Percentile	Predicted Average	
Yes	Generalized Linear Model - Cost: Total $ 1	Generalized Linear Model	99,404	5	90,457.5150	24,307,647,
No	Linear Regression - Cost: Total $ 1	Linear Regression	99,404	5	91,793.0157	28,974,483,

Figure 4. Comparison between general linear regression model and the linear regression model.

The result of the model implies that commercial airplane has a higher repair cost compared to another type, airline operator with the high cost of damage repair is US Airways, wildlife size involved in the strike damage is the large once, wildlife number struck in the strike are two to ten are involved in the strike, attitude the strike occur most are above 1000 feet, phase of flight during the strikes are during climb which is one of the crucial parts of the flight, and the speed of the aircraft was at 180 knots. Table 2.1 shows the parameter estimates.

Conclusion

In a wider term, this analysis correctly measures the choice and point of the bird strike issue in economic terms, by providing enhance information for bird strike damage cost value analyses and damage modification approaches. It cost the airline operator $37,948,803 for damage caused by bird strikes, which is a great loss for the airline and aviation industry at large.

Reference

Allan, J. R., & Orosz, A. P. (2001). The costs of birdstrikes to commercial aviation. In *2001 Bird Strike Committee-USA/Canada, Third Joint Annual Meeting, Calgary, AB* (p. 2).

Anderson, A., Carpenter, D. S., Begier, M. J., Blackwell, B. F., DeVault, T. L., & Shwiff, S. A. (2015). Modeling the cost of bird strikes to US civil aircraft. *Transportation Research Part D: Transport and Environment*, *38*, 49-58.

Dale, L. A. (2009). Personal and corporate liability in the aftermath of bird strikes: a costly consideration. *Human-Wildlife Conflicts*, *3*(2), 216-225.

Hedayati, R., & Sadighi, M. (2015). *Bird strike: an experimental, theoretical and numerical investigation*. Woodhead Publishing.

Metz, I. C., Ellerbroek, J., Mühlhausen, T., Kügler, D., & Hoekstra, J. M. (2020). The Bird Strike Challenge. *Aerospace*, *7*(3), 26.

Thomson, B. (2007). *A cost effective grassland management strategy to reduce the number of bird strikes at the Brisbane airport* (Doctoral dissertation, Queensland University of Technology).

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis* (Vol. 821). John Wiley & Sons.

Xu, R., & Wunsch, D. (2008). *Clustering* (Vol. 10). John Wiley & Sons.

Thorpe, J. (2012). 100 years of fatalities and destroyed civil aircraft due to bird strikes. In *30th Meeting of the International Bird Strike Conference June* (pp. 25-29).

Apendix:

Figure 5. Linear Regression

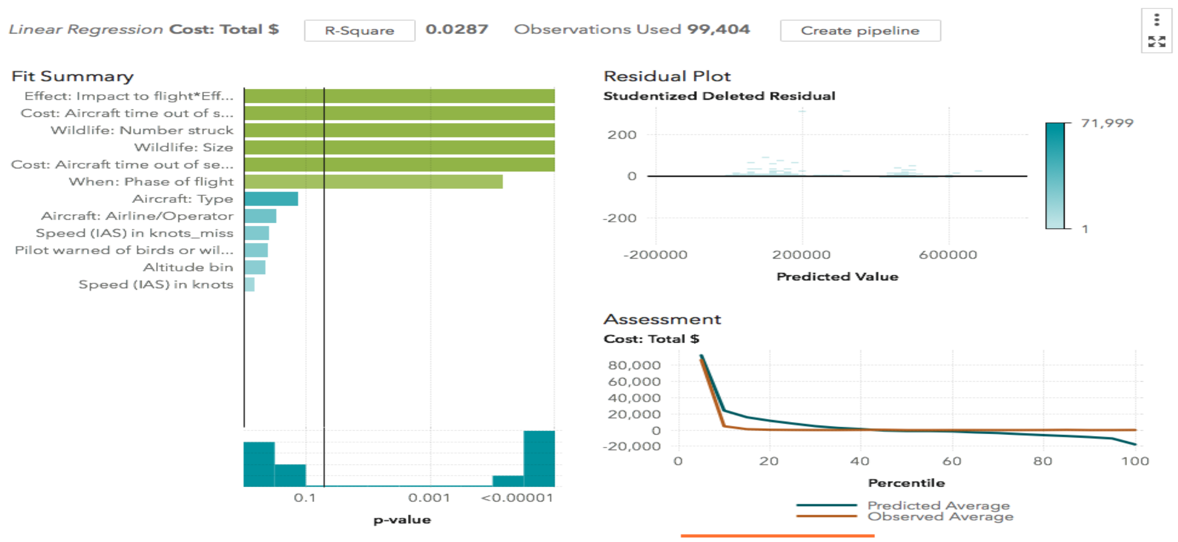


Figure 6. Generalized Linear Regression GLM