Transportation Research Record
2022, Vol. XX(X) 1–10

DOI: 10.1177/ToBeAssigned
journals.sagepub.com/home/trr
SAGE

An Improved Framework for Classification of Flight Phases of General Aviation Aircraft

Qilei Zhang[1] and John H. Mott[1]

Abstract

Flight data mining enables airport owners, operators, and governmental entities to explore more intelligent management strategies; in particular, cost-effectively obtaining accurate operational data is beneficial for general aviation airports and their associated communities. The current data collection modus operandi, however, does not meet future needs, as aircraft operations are counted manually or estimated by sampling methods. The increasing traffic flow and limited available personnel at most general aviation airports make it unrealistic to continue using traditional methods to analyze aircraft operational statistics; therefore, a customized approach is needed to address this problem. Since different flight phases have different levels and types of impacts on the environment, acquiring information related to the duration of each flight phase at the airport and within its surrounding airspace is critical to the assessment of emissions and noise pollution from aircraft. The primary goal of the research is to provide quantified inputs for the environmental evaluation model, such as the Aviation Environmental Design Tool (AEDT). This paper demonstrates a programmed framework that successfully achieves satisfactory performance in solving flight phase identification problems by testing the synthetic flight data as well as validating the empirical ADS-B data. The experimental results suggest that the proposed methods achieve promising classification accuracy, leading to feasible deployment in airport operations.

Keywords

ADS-B, flight identification, AdaBoost, probability distribution, classification, machine learning

There exists a large number of general aviation (GA) airports in the United States with limited or nonexistent air traffic control facilities (*1*). Both government regulators and airport owners need accurate operational statistics to facilitate airport planning, environment impact estimation, and funding allocation, as well as economic impact forecasting (*2, 3*). Federal Airport Improvement Program (AIP) grants are one of the primary sources for airport development, especially for small airports (*4*). The funding amount is justified by demand or passenger volume and granted to airports accordingly. Currently, however, operational statistics such as the number of operations (take-offs and landings) are collected by air traffic personnel at towered airports and estimated by sample counts or other methods at nontowered airports (*5, 6*). In the meantime, with the implementation of regulations (*7*) starting from January 2020, Automatic Dependent Surveillance-Broadcast (ADS-B) equipment is a mandated requirement to operate in most controlled airspace in the national airspace system. This new policy offers an opportunity to obtain previously unknowable details of flight operations. Because ADS-B contains a significant amount of data related to flight operations, these data may be considered a multivariate time-series sequence, as ADS-B receivers collect a series of measurements of multiple flight status attributes. The process of classifying empirical flight data, known as flight phase identification, can facilitate the more accurate estimation of flight operations.

Broad Impact

The quantification of the flight phase duration contributes to various studies, including the measurement of aircraft emissions, noise pollution impact, and airport capacities. Consider the emission problem as an example. Only by accurately obtaining the duration of different flight phases is it possible to understand the impact of different types of emissions, which, in turn, can facilitate further environmental evaluation for airports and their associated communities. Such evaluation could determine the characterization of the distribution of concentrations of airborne trace elements in the vicinity of the airport with general

[1]School of Aviation and Transportation Technology, Purdue University, West Lafayette, IN

Corresponding author:
Qilei Zhang, zhan3599@purdue.edu

aircraft operations. For example, leaded aviation gasoline contributes the most significant proportion source of lead (Pb) emissions, exposing at-risk groups to adverse health impacts, particularly neurological effects in children (*8*). U.S. Environmental Protection Agency (EPA) estimates that 16 million people live within one kilometer of airports where the aircraft operate on leaded fuel, and three million children attend school within one kilometer of these facilities (*9*). Given the significant impact of the potential health risks, evaluating exposure emissions such as lead from piston-engine emissions would help inform policy.

The Aviation Environmental Design Tool (AEDT) came into being as the primary environmental model used to estimate aviation noise and exhaust emissions. The AEDT system "models aircraft performance in space and time to estimates fuel burn, emissions, and noise consequences" according to the description from Federal Aviation Administration (FAA) (*10*). The operational statistics, i.e., phase duration inputs for the landing and take-off (LTO) model, drive the system model estimates (*11*). Specifically, high hydrocarbon (HC) and carbon monoxide (CO) emissions account for a considerable portion of the taxi phase, but the highest nitrogen oxides (NO_x) emissions occur during take-off. Therefore, different pollutants are predominant during different flight phases (*12*). Hence, a feasible means to obtain statistics related to the duration of each flight phase will be helpful to construct summary information for aircraft operations and these environment impact estimates (*13*).

Literature Review

There are two main types of methods commonly used to sample historical aircraft traffic data at non-towered airports (*2*). One of them is to deploy air traffic counters by collecting acoustic, pneumatic, and visual image information (*2, 3*). Another is by soliciting information from airport personnel or applying statistical methods using airport guest logs or fuel sales (*2, 6*). Specifically, airport managers may sample traffic for two weeks for each of the four seasons and extrapolate that sample into an annual estimate (*3*). An FAA capstone program report (*14*) documented how these methods are applied in southeast Alaska airports. The literature mentioned above indicates that these two types of methods have problems of high cost and low accuracy.

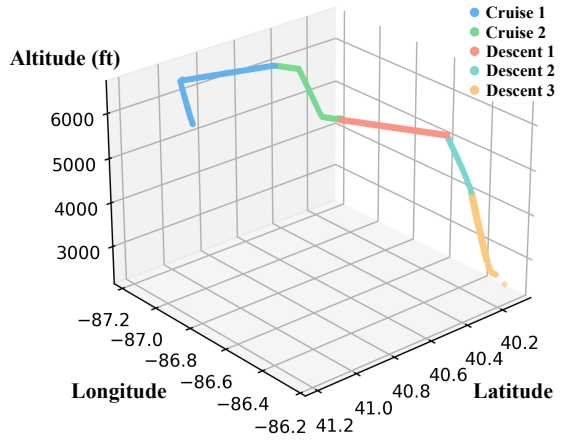
On the other hand, ADS-B data is known for its high-frequency update capability and low cost (*7*). However, the effort to analyze flight records derived from ADS-B transmissions can also be challenging for several reasons (*15*). First, due to the large amounts of cumulative data with low information densities, extracting the desired information from the dataset may be time-consuming. For example, data metrics calculation or visualization requires significant computing power. Consequently, manually processing and analyzing the rapidly accumulated ADS-B data may be unrealistic. Second, noise caused by dipole receiving antenna characteristics may lead to discontinuous sampled and lost data at certain orientations, especially at lower altitudes, such as the taxi phase. Higher gain antennas, although increasing the strength of the received signal, thereby allowing better information transfer, have a narrower antenna pattern, leading to some dead spots (*16*). These phenomena may increase the difficulty of precise classification. Lastly, ADS-B has some inherent limitations, such as providing ground speed (GS) instead of true air speed (TAS). These two quantities are identical only if the wind speed equals zero. Otherwise, the conversion of these two values under non-ideal conditions requires the introduction of wind speed data from a meteorological database. The reasons mentioned above constitute the challenge to effectively identify the corresponding flight phase classification from the ADS-B data.

Several recent research approaches attempt to solve the problem in different ways. Tian et al. (*17*) explored a method of dividing the flight phase based on a decision tree. Although the classifier achieved respectable results, the small scale of flight experiments leads to concerns over the construction of the decision tree. The stability of model parameters may be affected significantly by individual data; therefore, the model may be optimized only for specific datasets. Sun et al. (*18*) and Goblet (*19*) each propose a twofold method by clustering scattered data and labeling phases. Specifically, they first used unsupervised machine learning methods such as k-means (*20*), DBSCAN (Density-Based Spatial Clustering of Application with Noise) (*21*), BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies) (*22*), and TICC (Toeplitz Inverse Covariance-based Clustering) (*23*) to segment data into clusters. Subsequently, these data slices are labeled with correct flight phases by Boolean reasoning (*24*), decision trees (*25*), or fuzzy logic (*26*). Many other studies related to ADS-B identification data mining focus primarily on traffic flow classification (*27, 28*), specific maneuver pattern recognition (*29, 30*), and collision behavior detection (*31, 32*). Hence, developing a data-driven approach will facilitate the quantification of operations estimates for general aviation airports (*33*). This article proposes a promising accurate and efficient method to address the problem of classifying and labeling large quantities of flight data.

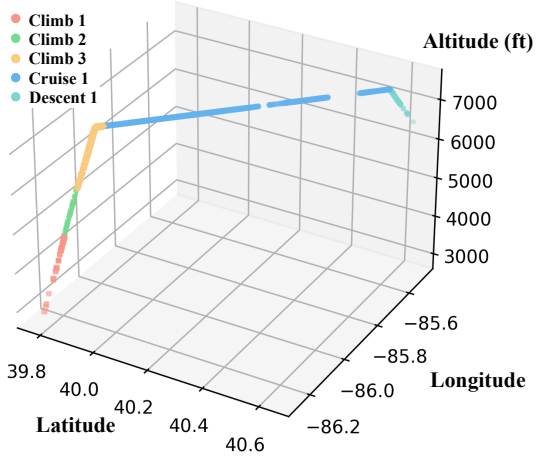
Data Preparation

The preliminary experiment selected two years of data (2019 to 2020) collected by ADS-B receiving equipment installed at the Purdue University Airport (KLAF). The data was then subsetted in a time range of interest, cleared of entries with missing values, and corrected for atmospheric pressure variations by obtaining the compiled Meteorological Terminal Air Report (METAR) data. Figure 1 shows two flight trajectories composed of data points. The motion of the

two aircraft is clearly divided into several segments marked by different colors according to the ICAO phase definitions (*34*). However, it also can be seen that the trajectory's pattern in Figure 1b is not as continuous as the one in Figure 1a. This common situation depicts the flaws in the ADS-B data, in that some data may be lost due to multiple factors such as antenna orientation, possibly leading to inaccurate statistics.



(a) This sample is extracted from the aircraft coded AA63F1 on August 17, 2020, from 01:07 to 01:51 in Eastern Standard Time.



(b) This sample is extracted from the aircraft coded ABB72E on August 16, 2020, from 21:53 to 22:25 in Eastern Standard Time.

Figure 1. Flight trajectory examples with empirical ADS-B data that are clustered and labeled by the TICC algorithm and fuzzy logic.

Because an empirical ADS-B dataset does not contain a means of validating phases of flight (*35*), there is no basis on which to compare the solution of each algorithm. It is also challenging to manually determine the corresponding label for such a large quantity of data. Therefore, synthetic data are introduced to evaluate the effectiveness of the algorithms, produced with an initial label, which can be stored to validate the accuracy of the models' judgment (*13*). Six different phases, i.e., *Taxi, Takeoff, Climb, Cruise, Descent, Approach*, are defined in this synthetic data generation process. These data record changes in the simulated operation of the aircraft as an ADS-B receiver does.

Every simulation contains an aircraft generated at simulation time t_0. A random acceleration vector is then assigned to this aircraft every second to accomplish a complete LTO cycle along the designated flight trajectory. Other attributes such as position (longitude, latitude), altitude, and GS can be calculated based on the acceleration and the previous position. Here, a simplified estimation is performed to reduce the computation. Specifically, the aircraft velocity vector, including speed and track, is constant during the fixed update interval, i.e., every second. As a result, the data series at each moment includes not only flight attributes such as location coordinate, ground speed, track, and heading, which are the same as those of the empirical ADS-B data, but also includes corresponding phase labels and derived information such as altitude and rate of change of velocity. The aircraft will remain associated with a consistent phase label and continuously slightly different acceleration vectors within a random duration. After the previous time slot, the acceleration vector will change drastically, and the label will advance to the next phase. Furthermore, normally distributed random error is appended to simulate the signal noise. Figure 2 depicts an example of the synthetic data in a three-dimensional plot. Finally, each set of simulation results are added to an integrated database. Two synthetic datasets were generated, marked as the training set (1,200 simulations with 2,590,889 data points) and the test set (200 simulations with 429,145 data points), respectively.

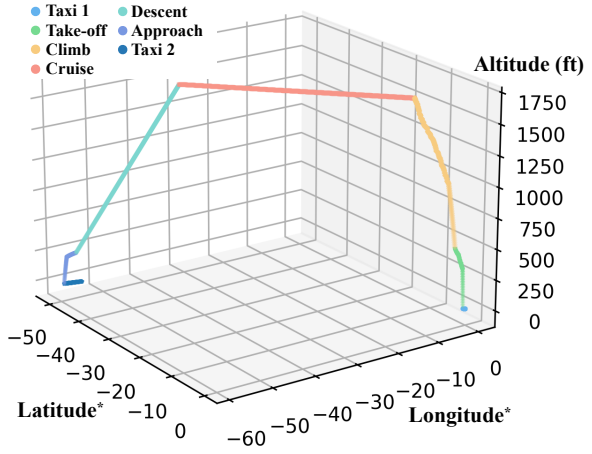


Figure 2. Flight trajectory example with synthetic ADS-B data. All points are generated with initial phase labels. The latitude* and longitude* here do not represent the actual measurement but only the relative movement on the plane.

Methodology

Two different methods are presented individually in this section to address the classification problem. Additionally, these two methods are integrated to further enhance model performance.

AdaBoost

AdaBoost (*36, 37*), short for *Adaptive Boosting*, is a meta-algorithm, also known as an ensemble learning technique. The main idea is to combine weak learners, the accuracy of which is slightly better than random guessing in a binary classification problem, to form a "strong learner" that can more effectively classify "difficult" examples. This type of machine learning algorithm generally achieves better predictive performance than could be obtained from any of the constituent learners alone (*38, 39*). To address the problem of phase identification, a single-phase label is expanded into multiple phases' binary judgment ("0" for "false", "1" for "true"). For example, if a data point belongs to "climb", it will be assigned 1 in the phase attribute *Climb*, and 0 in the other phase attributes, e.g., *Taxi* and *Cruise*. Four flight attributes (altitude, GS, altitude change rate, and GS change rate) are selected as universalized input $\mathbf{x}$ since they are the fundamental and representative identifiers in phase classification. In other words, the phase transition occurs in the dynamic changes of these four variables. The selection of the input feature vectors is also based on the consideration of simplification to merge ADS-B data with other data sources, thereby making up for the lack of some missing ADS-B data in the future. Six binary phase classification models are then trained by leveraging the *scikit-learn* Python machine learning library's *AdaBoostClassifier* module (*40*) to apply the AdaBoost algorithm.

Algorithm 1: AdaBoost Learning Algorithm

Input: Training set S with four attributes, decision tree classifier as base learner A, initial weight $D_t = \{D_1(1), \cdots, D_1(N)\}$;

Initialization: $D_1(i) = \frac{1}{N}, \forall i$;

for t = *1* **to** T **do**

  Obtain a classifier: $h_t \leftarrow A(S, D_t)$;

  Calculate importance factor β_t for h_t: $\beta_t = \frac{1}{2} \ln \frac{1-\epsilon_t}{\epsilon_t}$, where weighted error of h_t is: $\epsilon_t = \sum_{n=1}^{N} D_t(n)\mathbf{I}(h_t(x_n) \neq y_n)$;

  Update weights as follows: $D_{t+1}(n) = \frac{D_t(n)\exp(-\beta_t y_t h_t(x_n))}{Z_t}$

  The denominator Z_t ensures all D_{t+1} sum to 1, where $Z_t = \sum_{n=1}^{N} D_t(n)\exp(-\beta_t y_t h_t(x_n))$

end

return the final classifier $H(x) = \text{sgn}(\sum_{t=1}^{T} \beta_t h_t(x))$.

The complete learning process of AdaBoost is shown in Algorithm 1. Here, the base algorithm can be formalized as

$$\min \sum_{i=1}^{N} D_t(i) L(\mathbf{x_i}, y_i). \tag{1}$$

The input data is $\{(\mathbf{x_i}, y_i)\}$ and $L(\mathbf{x_i}, y_i)$ denotes the loss function. $D_t(i)$ represents the weight of the loss function at time step t. $h_i(x)$ is a weak classifier whose accuracy is slightly greater than 50%, i.e., error $\epsilon_t \leq 50\%$. The final classifier $H(x)$ can be expressed as $H(x) = \text{sgn}(\sum_1^T \beta_t h_t(x))$. It can be shown (*37*) that the total error rate of this output classifier is bounded as shown in Formula (2). Since $\epsilon < 0.5 \Rightarrow 2\sqrt{\epsilon_t(1-\epsilon_t)} < 1$, the total error converges quickly to 0 as t increases.

$$\frac{1}{N}\sum_{i=1}^{N} \mathbf{I}(H(x_i) \neq y_i) \leq \prod_{t=1}^{T} 2\sqrt{\epsilon_t(1-\epsilon_t)}. \tag{2}$$

Subsequently, to predict the phase, all the points in the data are traversed and trained models determine whether they belong to a particular phase. The output is six new columns filled with binary prediction. In most cases, the sum of these six numbers in every row is exactly "1" as every data point only belongs to one phase. Specifically, 5 (cols.) $\times$ 0 $(false)$ + 1(col.) $\times$ 1(*true*) = 1. The prediction result can then be determined by which phase column equals 1, i.e., "true" for this phase. For example, "1" in the column *Climb_Predict* and the other five all with "0" indicate that the prediction result is *Climb*. However, the phase prediction will be classified as *Unknown* if the sum is not equal to one. Consequently, Figure 3 depicts the increase of *accuracy* (defined as $\frac{\#\text{correctly predicted data items}}{\#\text{total data items}} \times 100\%$) for the test dataset as the size of the training dataset increases. The 60 scattered $\mathbf{x}$ marks in Figure 3 represent the number of the trained data series for the AdaBoost method, ranging from 5,000 to 2,500,000 series. The corresponding base e logarithm value is from 8.52 to 14.73. The final *accuracy* of AdaBoost stabilizes at roughly 98%. The other two methods displayed in Figure 3 will be described in subsequent sections.

Probability Distribution

The Probability Distribution method attempts to utilize the *a priori* probability by extracting the phase characteristic information from the training data. Specifically, data $x_{i,k} \in \mathbf{R}^{N\times K}$ in training set S with K attributes and N rows has a corresponding interval and probability value in every flight phase's probability density function (PDF) $f_{c,k}(x_{i,k})$. It is apparent that $\int_{-\infty}^{+\infty} f_c(x)dx = 1$ for a certain attribute k as the sum of the probabilities of all outcomes must equal 1. Here, k denotes the attributes such as altitude, and c represents the assigned classification code, such as "3" for cruise, "5" for approach, et cetera. For example, assume that an aircraft is operating at 800 ft at a certain moment. $f_{5,\text{alt}}(800)$ is greater than $f_{3,\text{alt}}(800)$ as shown

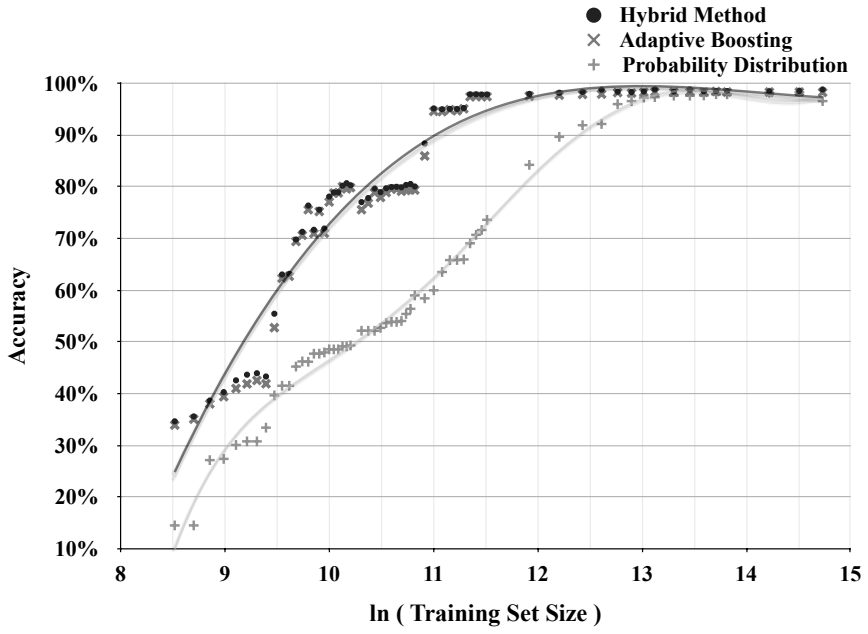


Figure 3. Scatter plot to show the increase of the *accuracy* for the test dataset as the size of the training dataset increases for all methods. The dependent variable has a natural log transformation applied. All three methods' trend lines are based on polynomial regression.

in Figure 4. It means that the aircraft is more inclined to be in the status "approach" rather than "cruise" at this altitude. In this case, all random variables of attributes are assumed to be independent so that their joint probability is the product of their marginal densities. The correlation matrix of four variables is provided in Table 1. The overall objective turns into finding maxima $c = \underset{c \in C}{\operatorname{argmax}} \prod_{k=1}^{K} f_{c,k}(x_{i,k})$ for the ith data item. Particularly, to utilize the continuity of chronological data, the regularization parameter λ is selected to encourage similar classification of neighboring data. As λ approaches infinity, the classification of the previous data item becomes more influential on the next item, i.e., they tend to be classified as the same phase. Therefore, the final objective function is formulated as:

$$\underset{c \in C}{\operatorname{argmax}} \prod_{k=1}^{K} f_{c,k}(x_{i,k}) + \lambda \prod_{k=1}^{K} f_{c,k}(x_{i-1,k}) \qquad (3)$$

Table 1. Correlation matrix of random variables.

	Alt	GS	Alt_rate	GS_rate
Alt	1.000	-	-	-
GS	0.484	1.000	-	-
Alt_rate	0.002	-0.007	1.000	-
GS_rate	0.002	0.010	0.123	1.000

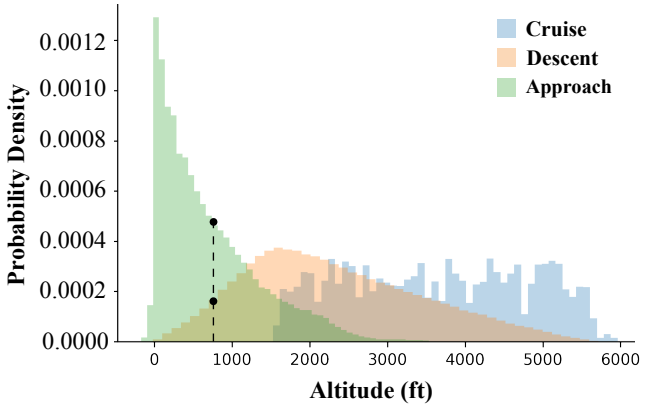


Figure 4. Histogram plot with 50 bins to demonstrate altitude probability distribution. The data is sampled from the training set containing 2,590,888 items.

Hybrid Method

To further improve the rate of correct prediction, the probability distribution method is integrated into the AdaBoost method to take advantage of both methods' strengths. Specifically, as mentioned previously, there are some *Unknown* data points in the AdaBoost prediction. If this is the case, the original attributes of this input will be transferred to the probability distribution method for further analysis. Consequently, the two methods' combination is employed to improve the solution performance. Figure 5

illustrates the basic process structure of the methodology with an overall framework and individual components.

Implementation and Results

Performance Metrics

In order to verify the effectiveness of the two methods, there are two perspectives for consideration, i.e., from the individual data point and from the duration of each phase. For the former, the evaluation focuses on the quality of classification. Therefore, Table 2 reports the precision, recall, and F-1 score for each phase. In particular, the summarized F-1 score for the entire prediction here averages the total true positives, false negatives, and false positives. This practice is usually referred to as *micro averaging* in multi-classification problems. In this case, the following equation always holds true: *micro-F1 = micro-precision = micro-recall = overall accuracy*. These metrics are all listed in Table 3 under the *Accuracy* column.

If one treats data in the same category as an aggregate, Symmetric Mean Absolute Percentage Error (SMAPE) is chosen as an evaluation metric. This metric reflects the relative errors (deviation ratio) to the overall flight duration quantity. Formula (4) describes the measurement procedure. N is the total number of the aircraft in the test dataset, c still represents the number of phases, F_{nc} indicates the predicted time, and A_{nc} indicates ground-truthing phase duration; accordingly, ΔT_{nc} is the calculation deviation. Furthermore, predicted total time F_n and real total time A_n can be substituted by T_{ntotal} as all of them are equal.

$$\begin{aligned} \text{SMAPE} &= \frac{100\%}{N} \sum_{n=1}^{N} \frac{\sum_1^c |F_{nc} - A_{nc}|}{|A_n| + |F_n|} \\ &= \frac{100\%}{N} \sum_{n=1}^{N} \frac{\sum_1^c |\Delta T_{nc}|}{2 \cdot T_{ntotal}} \end{aligned} \tag{4}$$

Comparison

Table 3 reports the performance of the two methods and their combination. Additionally, the multi-class AdaBoost is investigated, and its result is given. Other traditional methodologies are also introduced as the baselines to demonstrate the performance boost. It is reasonable to see that the accuracy of the test set has a minor decrease compared to the training set. Nevertheless, the accuracy improvement of the proposed models is still significant relative to other methodologies. Besides, AdaBoost had better results in both metrics compared to Probability Distribution. Therefore, the former provides the primary means of judgment in the methods' combination, and the latter assists identification as described in the previous section. Moreover, although multi-class AdaBoost has a slight advantage over the binary classifiers' combination, the hybrid method yields the best performance as expected in the listed methods. In summary, the proposed methodologies achieve a strong classification ability when applied to the flight phase identification problem.

Empirical ADS-B Dataset Validation

The hybrid method has also been applied to the empirical ADS-B dataset to validate the feasibility of the trained model. Thirty aircraft operated from April 07 to April 09, 2020, have been randomly selected in the ADS-B dataset. Two examples per day are shown in Figure 6. The model accomplished satisfactory classification using synthetic flight data as a training dataset. Additionally, Figure 7 gives an enlarged display of the A64CAC aircraft's partial trajectory circled by red shown in Figure 6. The larger view explains why there are some trajectories mixing two colors. This phenomenon frequently occurs during the descending phase. When the aircraft descends in steps, the model only marks the changing neighbored points with a "*descent*" label, leaving other level flight points with a "*Cruise*" label. These steps typically last around one minute after descending 100ft. On the contrary, the "*Climb*" phase has a more straightforward path compared to the "*Descent*" phase.

Discussion

The results obtained by testing synthetic data verify the universality of the proposed methods, thus indicating the new framework to be a reliable and economical solution. The application experiment in the empirical ADS-B exhibits the generalization ability of the methodology. Its output may have the potential for widespread application, such as an accurate emission analysis for airports. As a result of the pre-trained model, the program uses limited computational resources in the identification process, and there is no human intervention in the judgment procedure. On the contrary, unsupervised machine learning methods such as k-means and DBSCAN usually initially cluster data and then execute identification steps (*18*). The performance of these methods relies heavily on expert knowledge since concrete identification logic rules are formulated manually. Parameters for the logic may need to be adjusted accordingly for the particular circumstances of each airport. As a result, these methods are not as accurate and efficient as the proposed model.

However, because both methods greatly depend on precisely labeled data, the lack of real high-quality training data may become a limitation of the proposed model as a supervised learning method. On the other hand, discontinuous or missing ADS-B signals, especially at low altitude stages, may increase the difficulty of the phase classification. This is because both methods require the derived attributes, such as the change rate of altitude, as inputs. Defective data may lead to more variation of the

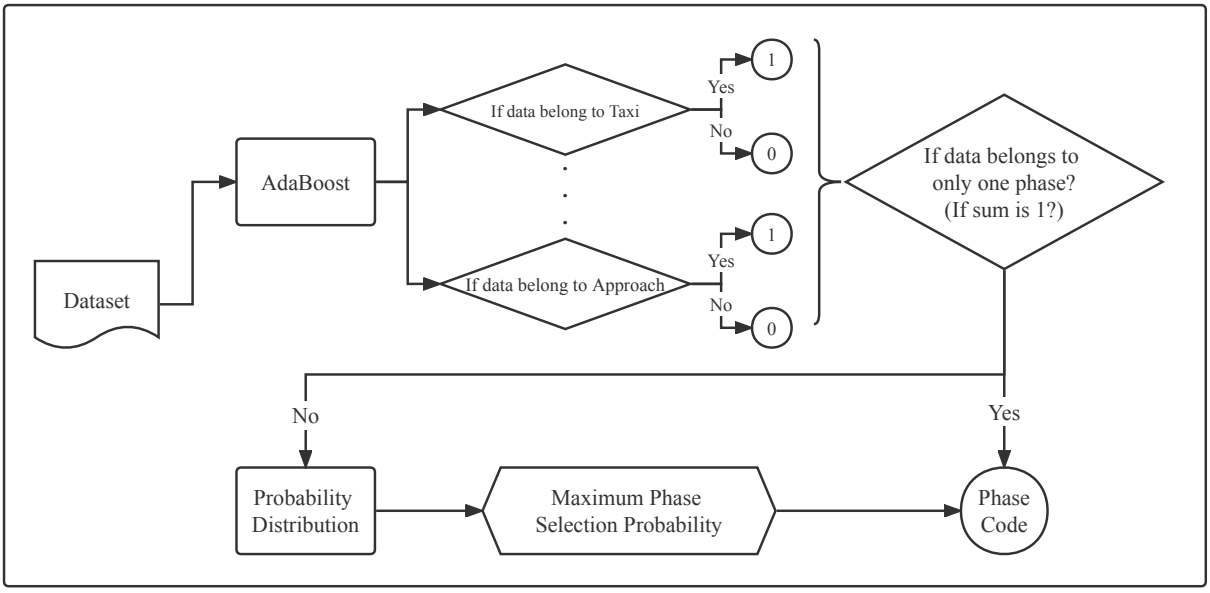


Figure 5. This is the flowchart of the hybrid methods. The flight phase duration is then summarized according to the phase code.

Table 2. Precision, Recall, and F-1 Score Report for Phases.

Hybrid Method				**AdaBoost (Multi-class)**			
Phase	Precision	Recall	F1-score	Phase	Precision	Recall	F1-score
Taxi	0.99840	0.99989	0.99914	Taxi	0.99856	0.99982	0.99919
Takeoff	0.99977	0.99962	0.99969	Takeoff	0.96936	1.00000	0.98444
Climb	0.99540	0.97027	0.98267	Climb	0.99999	0.96426	0.98180
Cruise	0.98990	0.99848	0.99417	Cruise	0.98974	1.00000	0.99484
Descent	0.97651	0.98900	0.98272	Descent	0.97034	0.96982	0.97008
Approach	0.93693	0.86893	0.90165	Approach	0.83976	0.83854	0.83915
Adaboost (Binary Combination)				**Probability Distribution**			
Phase	Precision	Recall	F1-score	Phase	Precision	Recall	F1-score
Taxi	0.99856	0.99987	0.99921	Taxi	0.88623	0.99858	0.93906
Takeoff	0.99985	0.99746	0.99865	Takeoff	0.98197	0.49431	0.65760
Climb	0.99539	0.96811	0.98156	Climb	0.98677	0.96807	0.97733
Cruise	0.98998	0.99101	0.99049	Cruise	0.98989	1.00000	0.99492
Descent	0.98221	0.98591	0.98406	Descent	0.98209	0.97668	0.97937
Approach	0.94947	0.84865	0.89623	Approach	0.87266	0.90032	0.88628

Table 3. Model Performance Comparison.

Index	**Method**	**Accuracy**[1]	**SMAPE**
1	Hybrid Method	98.83% (99.04%)	1.04% (0.77%)
2	AdaBoost (Multi-class)	98.40% (98.47%)	1.19% (1.15%)
3	AdaBoost (Binary Combination)	98.30% (98.87%)	1.21% (0.79%)
4	Probability Distribution	96.48% (96.68%)	2.03% (1.93%)
5	TICC + Fuzzy + Boolean[2]	91.99%	5.90%
6	K-means + Fuzzy + Boolean	81.79%	21.10%
7	DBSCAN + Fuzzy + Boolean	80.30%	22.63%
8	K-means + Boolean	52.78%	46.58%

[1] Values in parentheses are the training set results.
[2] "Fuzzy" stands for fuzzy logic, and "Boolean" stands for Boolean logic.

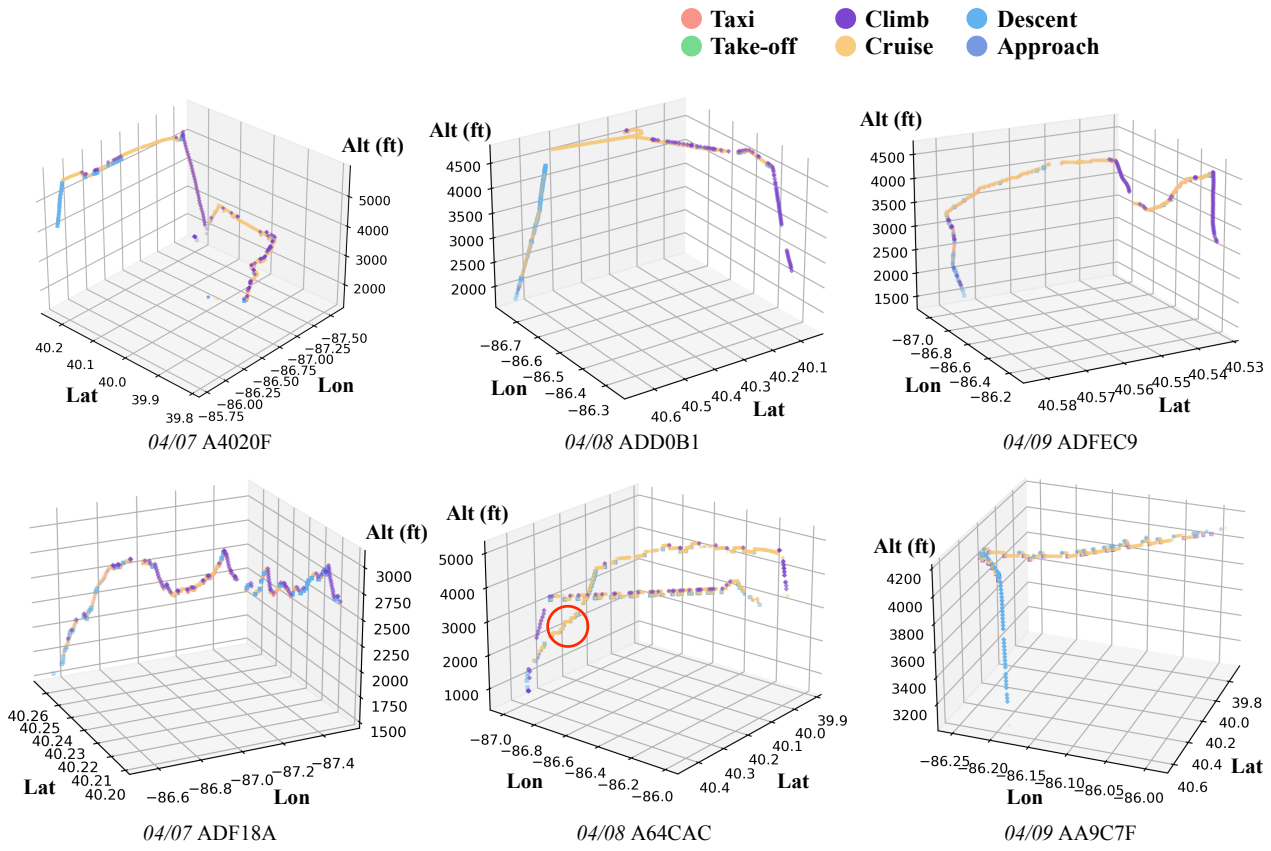


Figure 6. ADS-B flight trajectory example classified by hybrid method model. All six examples are noted by the operation date and hex identification code.

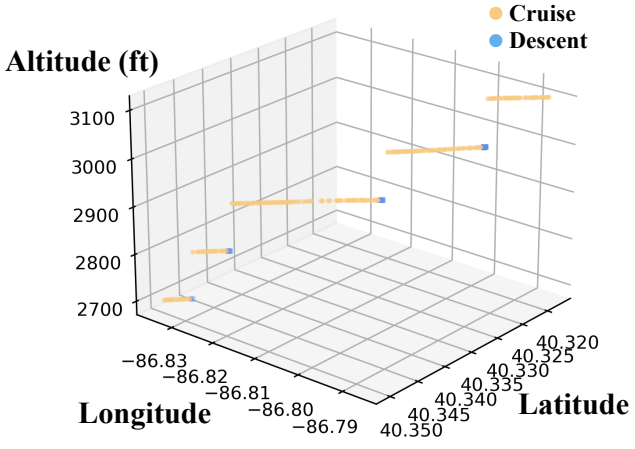


Figure 7. Partial enlarged ADS-B flight trajectory from aircraft coded A64CAC on April 8, 2020.

trained model, thereby resulting in inaccurate predictions. Including other auxiliary information such as TAS or power settings or other diverse surveillance data sources may alleviate issues due to missing ADS-B data, allowing the achievement of more precise phase identification.

Conclusion and Future Work

Flight phase identification is an important yet challenging task whose objective is to classify flight data into different flight phases. The output estimates can provide summary information for aircraft operations and contribute to environmental impact evaluation as input. Efficiently obtaining operational statistics is beneficial for nontowered general aviation airports as well as for community action groups seeking to reduce the perceived environmental impact due to noise and exhaust. This research addresses the need for more accurate emissions estimates by developing a flight phase identification framework. In this study, two methodologies are proposed and achieve higher accuracy compared with other traditional methodologies. The hybrid methods give even better results than the individual cases. Additionally, the experiment presents results when the phase identification framework is applied to an empirical ADS-B dataset.

The methodology proposed in this article still has room for improvement in terms of reliability and practicability. The problem of noisy or missing ADS-B data at certain phases may be resolved by collecting multiple data sources, such as from the Garmin 1000 (G1000) system. These additional

data would complement unknown information to provide more robust training data. Furthermore, the current synthetic flight data generator cannot simulate the various behavior of different aircraft types. The problem may be tackled using deep learning techniques. Specifically, construct a generative adversarial network (GAN) (*41*) to self-learn the pattern of diverse aircraft types, generating data closer to those in an empirical dataset. It is also necessary to conduct a long-term field case study to verify the method's feasibility. Targeted improvements can be made to address the deficiencies raised by practitioners. Future work will also consist of developing a real-time system that can monitor a dynamic operational metric. Other possible applications utilizing this summarized information, such as pollution statistics, shall be pursued. Their development will allow managers and regulatory agencies to better understand the environmental impacts and provide improved support for decisions regarding airport operations.

Acknowledgements

This work was funded by a grant from the Purdue Polytechnic Institute.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: Qilei Zhang and John H. Mott; data collection: Qilei Zhang; analysis and interpretation of results: Qilei Zhang; draft manuscript preparation: Qilei Zhang and John H. Mott. All authors reviewed the results and approved the final version of the manuscript.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was sponsored by a grant from the Purdue Polytechnic Institute.

ORCID iDs

Qilei Zhang https://orcid.org/0000-0003-1364-9661
John H. Mott https://orcid.org/0000-0002-2087-3971

References

1. FAA. General Aviation Airports: A National Asset. *US Department of Transportation*.
2. Li, T. and A. A. Trani. A Least-Square Model to Estimate Historical Percentages of Itinerant General Aviation Operations by Aircraft Types and Flight Rules at an Airport. *Journal of Advanced Transportation*, Vol. 2017, 2017, p. 9575676. doi:10.1155/2017/9575676. URL https://doi.org/10.1155/2017/9575676.
3. Muia, M. J. *Counting aircraft operations at non-towered airports*, Vol. 4. Transportation Research Board, 2007.
4. Grothaus, J. H., T. J. Helms, S. Germolus, D. Beaver, K. Carlson, and T. Callister. Guidebook for Managing Small Airports. *Transportation Research Board*, 2009, pp. 20–23.
5. Mott, J. H. Measurement of Airport Operations Using a Low-Cost Transponder Data System. *Journal of Air Transportation*, Vol. 26, No. 4, 2018, pp. 147–156.
6. Johnson, M. E. and Y. Gu. Estimating airport operations at general aviation airports using the FAA NPIAS airport categories. *International Journal of Aviation, Aeronautics, and Aerospace*, Vol. 4, No. 1, 2017, p. 3.
7. FAA. ADS-B – Frequently Asked Questions, 2020. URL https://www.faa.gov/nextgen/programs/adsb/faq/. Accessed: 05-May-2021.
8. Carr, E., M. Lee, K. Marin, C. Holder, M. Hoyer, M. Pedde, R. Cook, and J. Touma. Development and evaluation of an air quality modeling approach to assess near-field impacts of lead emissions from piston-engine aircraft operating on leaded aviation gasoline. *Atmospheric Environment*, Vol. 45, No. 32, 2011, pp. 5795–5804.
9. EPA. Advance Notice of Proposed Rulemaking and Related Materials on Lead Emissions From Piston-Engine Aircraft Using Leaded Aviation Gasoline. *Federal Register*, Vol. 75, No. 81, 2010, p. 22442. URL www.epa.gov/regulations-emissions-vehicles-and-engines/advance-notice-proposed-rulemaking-and-related-materials/.
10. Lee, C., T. Thrasher, et al. Aviation Environmental Design Tool (AEDT) Version 3b Technical Manual. *Federal Aviation Administration*.
11. Muia, M. J. and M. E. Johnson. *Evaluating Methods for Counting Aircraft Operations at Non-Towered Airports*. The National Academies Press, Washington, DC, 2015. doi:10.17226/22182. URL www.nap.edu/catalog/22182/evaluating-methods-for-counting-aircraft-operations-at-non-towered-airports.
12. ICAO. *Airport Air Quality Manual—Doc 9889*, chap. 3. 2nd ed. Secretary General of International Civl Aviation Organization, 2020, pp. 3.1–3.6. URL www.icao.int/publications/Documents/9889_cons_en.pdf.
13. Zhang, Q., J. H. Mott, M. E. Johnson, and J. A. Springer. Development of a Reliable Method for General Aviation Flight Phase Identification. *IEEE Transactions on Intelligent Transportation Systems*, 2021, pp. 1–10. doi:10.1109/TITS.2021.3106774.
14. Berman, M., W. Daniels, J. Brian, A. Hill, L. Kirk, S. Martin, J. Seger, and A. Wiita. FAA Capstone Program: Phase II Baseline Report (Southeast Alaska). *Institute of Social and Economic Research*.
15. Churchill, A. M. and M. Bloem. Clustering aircraft trajectories on the airport surface. In *Proceedings of the 13th USA/Europe*

Air Traffic Management Research and Development Seminar, Chicago, IL, USA. 2019, pp. 10–13.

16. Reudink, M. Cellular range extension with narrow-beam antennas. In *Proceedings of 1997 Wireless Communications Conference*. IEEE, 1997, pp. 79–83.
17. Tian, F., X. Cheng, G. Meng, and Y. Xu. Research on flight phase division based on decision tree classifier. In *2017 2nd IEEE International Conference on Computational Intelligence and Applications (ICCIA)*. IEEE, 2017, pp. 372–375.
18. Sun, J., J. Ellerbroek, and J. Hoekstra. Flight extraction and phase identification for large automatic dependent surveillance–broadcast datasets. *Journal of Aerospace Information Systems*, Vol. 14, No. 10, 2017, pp. 566–572.
19. Goblet, V. P. Phase of flight identification in general aviation operations. *Open Access Theses*.
20. Alsabti, K., S. Ranka, and V. Singh. An efficient k-means clustering algorithm. *Electrical Engineering and Computer Science*.
21. Ester, M., H.-P. Kriegel, J. Sander, X. Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, Vol. 96. 1996, pp. 226–231.
22. Zhang, T., R. Ramakrishnan, and M. Livny. BIRCH: an efficient data clustering databases method for very large. In *ACM SIGMOD international conference on management of data*, Vol. 1. 1996, pp. 103–114.
23. Hallac, D., S. Vare, S. Boyd, and J. Leskovec. Toeplitz inverse covariance-based clustering of multivariate time series data. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2017, pp. 215–223.
24. Brown, F. M. *Boolean reasoning: the logic of Boolean equations*. Springer Science & Business Media, 2012.
25. Quinlan, J. R. Learning decision tree classifiers. *ACM Computing Surveys (CSUR)*, Vol. 28, No. 1, 1996, pp. 71–72.
26. Zadeh, L. A. Fuzzy logic. *Computer*, Vol. 21, No. 4, 1988, pp. 83–93.
27. Lan, M. and G. Yongsheng. Four-dimensional trajectory prediction method based on ADS-B data mining# br. *Journal of Civil Aviation University of China*, Vol. 37, No. 4, 2019, p. 1.
28. Gariel, M., A. N. Srivastava, and E. Feron. Trajectory clustering and an application to airspace monitoring. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 12, No. 4, 2011, pp. 1511–1524.
29. Marsh, R. and K. Ogaard. Mining heterogeneous ADS-B data sets for probabilistic models of pilot behavior. In *2010 IEEE International Conference on Data Mining Workshops*. IEEE, 2010, pp. 606–612.
30. Bian, L., Y. Liu, Y. Zhan, H. Tang, M. Bo, L. Zhang, K. Luan, X. Yin, and J. Ding. Hovering recognition by ADS-B data mining. In *2020 IEEE 2nd International Conference on Civil Aviation Safety and Information Technology (ICCASIT*. IEEE, 2020, pp. 121–127.
31. Lai, C.-P., Y.-J. Ren, and C. Lin. ADS-B based collision avoidance radar for unmanned aerial vehicles. In *2009 IEEE MTT-S International Microwave Symposium Digest*. IEEE, 2009, pp. 85–88.
32. Strohmeier, M., M. Schäfer, V. Lenders, and I. Martinovic. Realities and challenges of nextgen air traffic management: the case of ADS-B. *IEEE Communications Magazine*, Vol. 52, No. 5, 2014, pp. 111–118.
33. Mott, J. H. and N. A. Sambado. Evaluation of Acoustic Devices for Measuring Airport Operations Counts. *Transportation Research Record*, Vol. 2673, No. 1, 2019, pp. 17–25.
34. ICAO Commercial Aviation Safety Team. *Phase of Flight, Definitions and Usage Notes*. Tech. rep., International Civil Aviation Organization, Montreal, Canada, 2013. URL http://www.intlaviationstandards.org/Documents/PhaseofFlightDefinitions.pdf.
35. Sun, J., J. Ellerbroek, and J. Hoekstra. Large-scale flight phase identification from ads-b data using machine learning methods. In *7th International Conference on Research in Air Transportation*. 2016, pp. 1–8.
36. Freund, Y., R. E. Schapire, et al. Experiments with a new boosting algorithm. In *icml*, Vol. 96. Citeseer, 1996, pp. 148–156.
37. Hastie, T., S. Rosset, J. Zhu, and H. Zou. Multi-class adaboost. *Statistics and its Interface*, Vol. 2, No. 3, 2009, pp. 349–360.
38. Opitz, D. and R. Maclin. Popular ensemble methods: An empirical study. *Journal of artificial intelligence research*, Vol. 11, 1999, pp. 169–198.
39. Polikar, R. Ensemble based systems in decision making. *IEEE Circuits and systems magazine*, Vol. 6, No. 3, 2006, pp. 21–45.
40. Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, Vol. 12, 2011, pp. 2825–2830.
41. Goodfellow, I., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative Adversarial Nets. In *Advances in Neural Information Processing Systems*, Vol. 27 (Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Weinberger, eds.). Curran Associates, Inc., 2014. URL https://proceedings.neurips.cc/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf.