doi: 10.2514/1.I010582

Anomaly Detection in General Aviation Operations Using Energy Metrics and Flight Data Records

Tejas G. Puranik* and Dimitri N. Mavris†
Georgia Institute of Technology, Atlanta, Georgia, 30332-0150

Among operations in the General Aviation community, one of the most important objectives is to improve safety across all flight regimes. Flight data monitoring or Flight Operations Quality Assurance programs have percolated in the General Aviation sector with the aim of improving safety by analyzing and evaluating flight data. Energy-based metrics provide measurable indications of the energy state of the aircraft and can be viewed as an objective currency to evaluate various safety-critical conditions. The use of data mining techniques for safety analysis, incident examination, and fault detection is gaining traction in the aviation community. In this paper, a generic methodology is presented for identifying anomalous flight data records from General Aviation operations in the approach and landing phase. Energy-based metrics, identified in previous work, are used to generate feature vectors for each flight data record. Density-based clustering and one-class classification are then used together for anomaly detection using energy-based metrics. A demonstration of this methodology on a set of actual flight data records from routine operations as well as simulated flight data is presented highlighting its potential for retrospective safety analysis. Anomaly detection using energy metrics, specifically, is a novel application presented here.

I. Introduction

One of the most important objectives among operations in the General Aviation (GA) community is to improve safety across all flight regimes. In the past, accidents have been the primary triggers for identifying problems and developing mitigation strategies [1]. However, the industry is now moving towards a more pro-active approach to safety enhancement in which potential unsafe events are identified beforehand and mitigation strategies are implemented in order to prevent accidents and loss of life. According to the National Transportation Safety Board (NTSB) [2], the number of total accidents per million flight hours in GA is an order of magnitude higher than that of commercial operations. With air traffic expected to grow tremendously over the next decade and GA set to receive a significant impetus, improving safety is of paramount interest to the aviation industry. Energy state awareness and energy management are critical concepts in the characterization, detection, and prevention of safety-critical conditions. The Federal Aviation Administration (FAA) has recognized Loss Of Control (LOC) and Controlled Flight Into Terrain (CFIT) as the leading causes of fatal accidents in GA [3]. Previous studies have shown that improper or poor energy management and loss of energy state awareness (LESA) are among the top contributors to LOC and CFIT accidents [4]. Paradoxically, energy state awareness and energy management have been addressed almost exclusively in commercial aviation where the concepts are intrinsic in operational safety and have been the subject of much research.

Flight Data Monitoring (FDM) or Flight Operational Quality Assurance (FOQA) [5] programs, which are well-established in commercial operations have percolated into GA with the aim of improving operational safety. Typical FOQA programs involve a continuous cycle of data collection from on-board recorders, retrospective analysis of flight data records, identification of operational safety exceedances, design and implementation of corrective measures, and monitoring to assess their effectiveness. Current practice in FOQA is chiefly underpinned by a priori definition of safety events known as ‘exceedances’ [5]. This method performs well on known safety issues but is blind to safety-critical conditions that may be captured by flight data records but not included in the set of pre-defined events. Data mining techniques for safety analysis, incident examination, and fault detection have garnered increased interest in the aviation community in recent years. While formal techniques for flight data analysis are not new, applications of data mining for retrospective operational safety analysis are fairly sparse. Additionally, a large portion of the existing literature is

*Graduate Researcher, Aerospace Systems Design Laboratory, School of Aerospace Engineering, AIAA Student Member.
†S.P. Langley NIA Distinguished Regents Professor, Director of Aerospace Systems Design Laboratory, School of Aerospace Engineering, AIAA Fellow.

dedicated to commercial aviation despite the fact that GA operations have historically had considerably greater accident and incident rates [2].

Transport category airplanes have certain minimum requirements related to Digital Flight Data Recorder Systems (DFDRS) set by the Federal Aviation Administration (FAA) for U.S carriers. [6] Therefore, data mining techniques have lent themselves well to data obtained from FOQA programs in commercial operations due to the wealth of information available. In GA operations that are the subject of this work (small airplanes weighing less than 12,500 lbs and having seating configuration of 10 or less), [7] a large number of parameters may not be recorded due to limited capability of their data collection hardware. Therefore, using recorded flight data with quantitative aircraft performance models (such as those developed in Harrison et al. [8] and Min et al. [9]) are a key enabler for retrospective safety analysis in GA. These models are used for evaluating a number of energy based metrics which provide a measurable quantification of the energy state of the aircraft and can be viewed as an objective currency to evaluate various safety-critical conditions across a large fleet of aircraft with minimal amount of recorded parameters. In the previous work (Puranik et al. [10, 11]), the authors conducted a thorough literature review of existing energy based metrics used in aircraft performance studies and defined some completely new metrics using GA flight data. Methods of generating a data-driven reference profile of different metrics was demonstrated in Puranik et al. [12]. The feasibility of using these energy metrics, reference profiles, and performance models in conjunction was explored. The current work aims to build upon these concepts by implementing them in an anomaly detection framework on a large data set of GA flights. Anomaly detection using energy metrics, specifically, is a novel application presented here. Some of the challenges identified earlier include accounting for the heterogeneity of GA operations and fleet, limited flight data recording capability, and lack of validation for identified outliers [13].

Unlike other applications of data mining or anomaly detection, aviation data is typically not labeled. There is no knowledge a priori as to which flight records (if any) are actually anomalous. This makes validation of developed techniques harder. Majority of the flights from routine operations are expected to contain non-anomalous data. Also, there is no set definition for what an anomaly in this context refers to. Therefore, unsupervised or semi-supervised algorithms need to be used to identify anomalies. Identifying anomalies as outliers of a clustering algorithm is useful way of anomaly detection as it allows the possibility of multiple standard patterns. Alternatively, one-class learning algorithms can also be used for identifying outliers. There are a number of ways in literature by which the “anomalousness" (or anomaly score) of an outlier can be quantified. Some anomaly detection algorithms directly provide an anomaly score while others use an external metric such as Local Outlier Factor (LOF) [14] to provide the anomalousness. Campos et al. [15] have provided a review of some of these scores using different data sets to quantify their relative performance. In most cases, the performance of an anomaly score is dependent on the type of data.

Considering the above observations we articulate the needs motivating the work in this paper as follows:

1) Demonstrate the use of energy metrics (or a subset of energy metrics) to identify anomalous GA flight data records in the approach and landing phase
2) Demonstrate identification of known anomalous flights (obtained through a flight simulation model) using the developed methodology
3) Quantify the performance of the methodology when limited data parameter recording capability is available (as is the case in a lot of GA aircraft)

It should be noted that throughout this paper, the term anomalous or abnormal flights refer to those flights behaving significantly differently than others in terms of certain critical parameters and metrics. This does not necessarily mean that the flights are inherently unsafe. While some of these may be benign flights that simply followed different procedures, others could point to potential safety issues which are of interest to safety analysts. It is the identification of these emergent outliers that is the aim and scope of this work. The main purpose of identifying these anomalies is the retrospective safety analysis of flight data from heterogeneous GA operations. The methodology presented in this research can aid industry experts and GA operators in better understanding and identifying unsafe practices. It could be used to improve flight training and instruction for student pilots as well as deployed on a large database of flights such as the National General Aviation Flight Information Database (NGAFID) to understand trends and behaviors.

The following sections will build the methodology and discuss results with the aim of addressing the objectives mentioned above. The rest of the paper is organized as follows: Section II provides a review of existing data mining techniques applied in the aviation domain. Section III contains the outline and description of the key elements of the methodology used in this work. Section IV presents the implementation of the methodology and results obtained and Section V presents the observations and conclusions and outlines avenues of future work being pursued.

II. Review of Existing Work

Previous applications of data mining in the aviation safety domain have primarily treated it as an anomaly detection problem with data objects as multivariate time series. [16–20] In the broader data mining community, anomaly detection is loosely defined as the "*task of obtaining patterns in data that do not conform to a well defined notion of normal behavior*" [21]. The literature identifies two main types of anomalies in aviation data – *Flight Level Anomalies* in which the entire flight record or phase of flight considered are off-nominal and abnormal data patterns persist over a period of time, and *Instantaneous Anomalies* in which only an instant or small part (less than a minute) of the flight record is off-nominal. In general different techniques are used for identifying each type of anomaly. The focus of this paper is flight-level anomalies as instantaneous anomalies have been addressed in separate work [13].

Chandola et al. [21] have provided a comprehensive survey of anomaly detection that covers techniques applied across all domains. Liao et al. [22] have specifically surveyed clustering techniques pertaining to time series data and provided a taxonomy of techniques. Within the aviation domain, Gavrilovski et al. [23] have surveyed data mining techniques and provided a review of published work on its application to flight data. They have also identified challenges and opportunities for its application to fixed and rotary wing GA applications and have also noted salient distinctions between application of these methods to commercial aviation against rotorcraft and fixed wing GA operations. Anomaly detection techniques used in literature can be broadly classified into two categories - supervised learning and unsupervised learning.

Supervised learning methods such as Inductive Monitoring System (IMS) [24] rely on a training set consisting of typical system behaviors which is compared with real-time data to detect anomalies. Each point is monitored standalone and therefore, the temporal aspect of anomalous sub-sequences is lost when identifying anomalies. SequenceMiner [25] has been shown to detect anomalies in discrete parameter sequences by learning from a model of normal switching. This technique detects flight-level anomalies but is limited to discrete data. Some of the methods rely on developing an approximate model using flight data and detecting those flight records which deviate greatly from this model as outliers. For example, Chu et al. [26] have proposed an approach for detecting anomalies from aircraft cruise flight data using a model trained using historical data of a fleet of aircraft. Anomalies are detected as outliers that exceed the scatter caused by turbulence and the modeling error. Gorinevsky et al. [20] have described an application of data mining technology called Distributed Fleet Monitoring (DFM) to Flight Operational Quality Assurance data. This application consists of fitting a large scale multi-level regression model to the data set and finding anomalies using these built models. The algorithm is able to identify anomalies within a flight record (instantaneous), abnormal flight-to-flight trends (flight-level anomalies) and abnormally performing aircraft. Hotelling T^2 statistics for residuals from the built models are calculated and used for monitoring and identifying anomalies. While this framework is capable of identifying instantaneous anomalies, it is limited to models fitted in the (aerodynamically) clean configuration. Also, most of the anomalies detected are in the determination of aerodynamic or propulsion parameter estimates or gross weight. Melnyk et al. [19] have treated each multivariate time series using a vector auto-regressive exogenous model. Dissimilarity between two flights is measured as the residuals obtained by using the model of one flight on the data of another. Outliers are identified using LOF which is a nearest neighbor based anomaly detection method. This method requires that a different model be built for each flight record being analyzed. This method also requires that pilot inputs be recorded which may not necessarily be the case for GA data since the required instrumentation is usually not present.

On the other hand, unsupervised learning methods do not rely on a training set and try to obtain anomalous records from a large data set using techniques such as clustering. Bay and Schwabacher [27] have described a method called Orca which uses Euclidean distance of a point to its k-nearest neighbors to gauge the anomalousness of each point. However, as with IMS, this method treats each point independently, therefore it loses the temporal aspect of abnormal flights. Das et. al. [16] have developed Multiple Kernel Anomaly Detection (MKAD) which applies a one-class support vector machine for anomaly detection. MKAD uses the normalized Longest Common Sub-sequence (nLCS) kernel which is useful for discrete data, but it results in loss of some finer features for continuous data (when it is discretized). Matthews et al. [17] have discussed and summarized the aviation knowledge discovery pipeline using various algorithms. Li et al. [18] have implemented ClusterAD, which uses cluster-based anomaly detection on pre-processed flight data parameters to identify abnormal operations. One of the potential issues in this method can be that it isolates each sample in each signal as a unique feature, when in fact, the change over that time may be an important factor. Another potential issue is the fact that anomalies identified are purely mathematical artifacts and may need subject matter expertise to validate them. Despite these potential limitations, MKAD and ClusterAD represent the state-of-the-art for anomaly detection in the aviation domain and have shown to be very effective at uncovering a host of anomalies that are missed by traditional approaches. There appears to be consensus in the literature that performance of anomaly detection techniques can be improved by including domain-specific knowledge (in this case, the energy metrics as features) in the data-mining task as this will increase the chances of identifying truly anomalous conditions [17, 19, 23].

III. Methodology

The methodology followed in this work to identify flight-level anomalies is outlined in Figure 1.

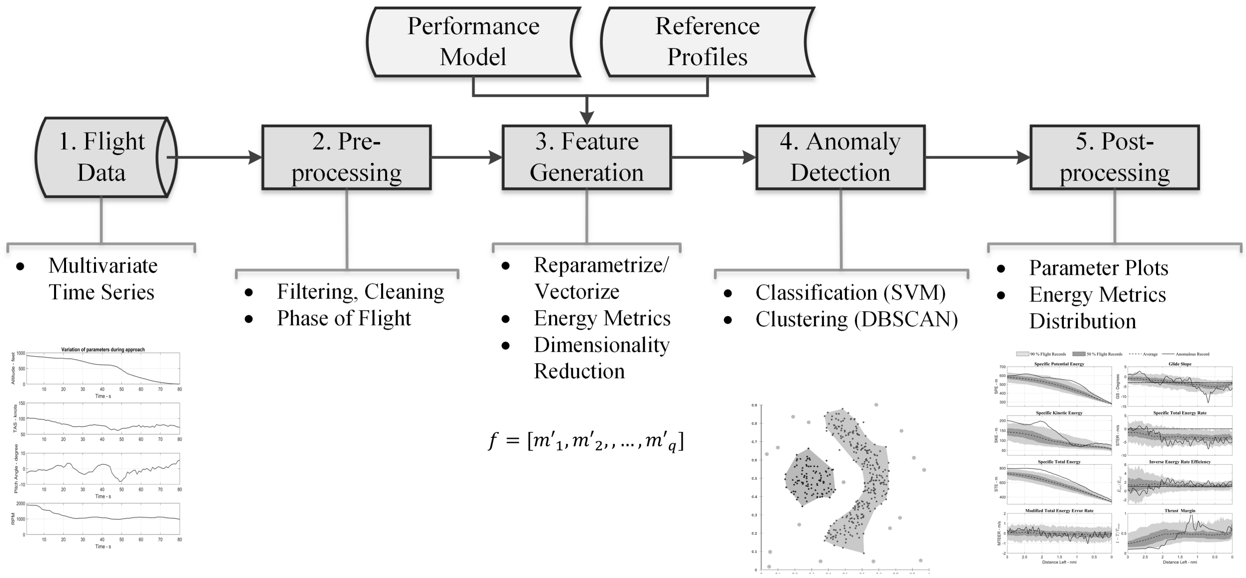

Fig. 1 Methodology followed to identify anomalous records using energy metrics

For identifying flight level anomalies, well-defined phases of flight need to be considered as they can be easily compared across different flights. Due to the heterogeneity in operations in GA, the phases of flight are not very easy to define. Goblet et al. [28] have provided definitions for common phases of flight in GA operations and their identification based on recorded parameters. In GA operations maximum accidents occur during the approach and landing phase [3]. This phase of flight can also be defined in a relatively straightforward manner according to the definitions provided by Goblet et al. Therefore, this paper focuses on the approach and landing phase in GA operations. The various steps in this methodology correspond to steps in a general anomaly detection framework. There are two external inputs to the framework that are not usually observed in the literature surveyed – *performance models* and *reference profiles*. This is because most approaches in literature are pure data-mining based and therefore only utilize the recorded data. The details of analyses performed in each step are provided in the following subsections. The entire framework is implemented in MATLAB.

A. Flight Data

The flight data obtained from DFDR is a multivariate time series, whose lengths typically vary between records due to varying duration of flight. The number of parameters varies from around twenty (in GA operations) to thousands (commercial operations). These parameters are recorded at a specific frequency (e.g., once per one second interval). In the current implementation, approximately one thousand flight data records collected from routine training flights on a Cessna 172S aircraft are utilized. The flights are conducted at the same airport but on multiple runways. The experience level of the pilots is variable but all are flown by student pilots. In addition, a simulated anomalous flight flown by a private pilot is also included with the actual flight data records (details provided in results – Section IV). The aircraft are equipped with a Garmin G1000 [29] glass cockpit for recording the flight data. Each data record contains approximately thirty continuous parameters related to the state, attitude, basic engine information, environmental conditions, Global Positioning System (GPS) information, and others. Discrete parameters such as pilot's control inputs or flap position are not explicitly recorded in GA DFDR and presents a some limitations. Specifically, lack of logged data corresponding to the pilot's control inputs implies that the capability of any anomaly detection methodology in *attributing* identified outliers to causes is curtailed. Therefore, this (and any other methodology with similar limitations in logged data) would strictly be looking at the effects of different actions (pilot actions, weather etc.) on the system in terms of its recorded parameters and identifying anomalies based on these effects. It also restricts the usage of methods that build a model from the recorded data (Melnyk et al. [19]) and require a set of inputs (*unavailable here*) and outputs (*available here*). On the other hand, due to presence of only continuous data in the set, feature generation and pre-processing are streamlined and made uniform across all parameters and metrics and special techniques requiring handling of heterogeneous data are not required.

B. Pre-processing

In all techniques surveyed from the literature, the raw data obtained from flights is pre-processed to obtain data suitable for anomaly detection. Noise that may be present in the original data is smoothed using a moving average filter to remove spikes caused due to noisy sensor readings. Similarly, records with incomplete or corrupted data are removed from the data set. Following basic cleaning and filtering of the data set, the approach and landing phases of flight are extracted from the flight data. The definitions given in Goblet et al. [28] are used to extract these phases from the flight data records. Once the data is pre-processed it is then used for generation of feature vectors for each flight record.

C. Feature Generation

In anomaly detection algorithms for flight level anomalies, a flight data record or phase of flight is characterized using *feature vectors*. These feature vectors are typically generated using the data obtained from the flight record and account for the temporal aspect of data. Anomaly detection algorithms then compare feature vectors of different flights to each other to obtain similarities between them. There are two important requirements on these feature vectors – Firstly, the feature vector generated must have the same dimensions for all flights in order to be comparable. Secondly, the corresponding elements of the vectors of different flights should be comparable. These two requirements give rise to multiple ways of mapping the multivariate time series to a single vector.

Das et al. [16] use the technique of Symbolic Aggregate Approximation (SAX) [30] to convert a continuous time series of parameters during approach and landing into discrete symbols by averaging the data over a number of seconds into a single symbol. An advantage of this type of approach is that approach profiles of varying duration can be incorporated as the total number of symbols is fixed. However, approximating continuous time series by discrete symbols usually results in loss of information due to down sampling. Also, this approach works well when the time series associated with the phase of flight is of a large duration (which is not the case for GA approach and landing). Another approach used by Li et al. [18] is to anchor the flight data record at certain significant points (example – application of take-off power in take-off phase and touchdown on runway in approach and landing). From the anchor point, the multivariate time series in takeoff are sampled at a fixed temporal interval whereas in approach and landing, they are re-sampled based on the distance remaining to touchdown. The reason for re-sampling in approach and landing based on distance rather than time is that many of the procedures in this phase are distance-specific. This approach is advantageous because each feature vector thus obtained is of the same dimension without down-sampling or loss of information. A potential disadvantage of this approach is that this requires the ground-track distance of the approach and landing phase to be approximately similar in each case, which may not necessarily be true for GA operations. It should be noted that in GA flight data records, the exact touchdown point of the aircraft on the runway is not recorded even using high-end recorders such as the G1000. Therefore, the anchor point in the approach and landing phase is not directly available but needs to be obtained through data analysis. Previous work done on flight data analysis and extracting information from flight records (Puranik et al. [12]) addressed this issue by using altitude difference between successive points in the flight data along with combinations of other parameters to obtain the approximate touchdown point on the runway. Therefore, in this paper, the resampled flight data record based on distance remaining is used to generate features. The approach followed in this paper for feature generation is shown in Figure 2.

The multivariate time series is first reparametrized for the approach and landing phase at various distance-based intervals up to three nautical miles from the runway threshold. The reason for choosing three nautical miles is that the data indicates significant variability beyond this marker due to heterogeneity in mission profiles (more details and justification on this choice are presented in Puranik et al. [12]). Following reparameterization, the energy metrics at each point are evaluated. The complete list of energy metrics is provided in the appendix (Table 2) and a more detailed description of these can be found in Puranik et al. [11]. As required for the metrics, the values of recorded parameters along with aircraft performance models and reference profiles are used. Due to the limited number of parameters recorded in GA operations compared to commercial operations, obtaining additional information can enhance the effectiveness of the safety analysis task. Predictions provided by aircraft performance models can add this value to the analysis. Reference profiles, on the other hand are nominal variations of energy metrics over certain phases of flight. These profiles are used

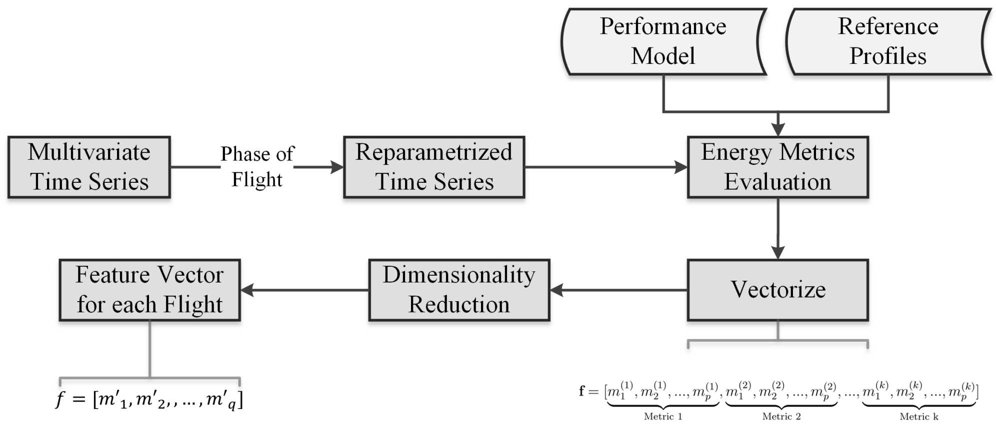

Fig. 2 Process of feature vector generation from flight data

to calculate deviation of certain energy metrics from nominal in those phases of flight. In some cases, these profiles are readily available. For example: the altitude profile that aircraft try to follow during approach and landing consists of a nominal (3°) glide slope. This corresponds to a certain variation of the specific potential energy metric profile. On the other hand, when they are not as easily available, they can be calculated using a data-driven approach by averaging over a large number of flight records to get nominal profiles which can be used as a reference. This approach has been elaborated in prior work (Puranik et al. [12]). Thus, each energy metric is evaluated at every distance-based interval along the approach and landing and the feature vector for each flight record can be represented by a single vector by concatenating the contribution from each metric as follows:

$$\mathbf{f} = \left[\underbrace{m_1^1, m_2^1, ..., m_p^1}_{\text{Metric 1}}, \underbrace{m_1^2, m_2^2, ..., m_p^2}_{\text{Metric 2}}, ..., \underbrace{m_1^k, m_2^k, ..., m_p^k}_{\text{Metric k}} \right] \tag{1}$$

where m_j^i is the value of the i^{th} metric at the j^{th} distance-based location or time-stamp in the approach and landing. Thus, the j^{th} element of each vector is now comparable to every other vector. Even though all the feature vectors have the same length, different metrics have different magnitudes. Therefore, each metric is normalized such that it has a zero mean and unit variance. This is achieved using z-score normalization. Let $m_{j,1}^i, m_{j,2}^i, ...m_{j,n}^i$ be the values of a particular metric i for all n flights at a particular distance j from the runway. The z-score normalized values are given by:

$$m_{j,1}^i = \frac{m_{j,1}^i - \overline{m_j^i}}{std m_{j,1...n}^i} \tag{2}$$

where std denotes the standard deviation. The feature vectors obtained can contain hundreds of dimensions depending on the resolution of the parameterization. However, while identifying outliers and clusters, the variability is typically embedded in a smaller number of dimensions [18]. Principal Component Analysis (PCA) is a linear transformation that is used to transform data into a new orthogonal coordinate system [31]. The coordinates in the new system are ranked in descending order of the amount of embedded information (variance) they contain. Dimensionality reduction is achieved by only retaining the first few components that explain majority of the variance (in this method the number is chosen such that 99% of the variance is captured). If the reduced dimensional vector contains p elements, each flight record can now be represented as:

$$\mathbf{f'} = \left[m_1', m_2', ..., m_p'\right] \tag{3}$$

The dimensionality reduction comes at the cost of losing the physical meaning of the components of the vectors thus obtained. Dimensionality reduction techniques such as PCA become intractable when the size of the data set becomes larger [18, 32]. The number of energy metrics is typically smaller than the number of parameters recorded (even in GA flight records), however, this step may be skipped if the size of the data set becomes intractable. It is noted that, while PCA or other dimensionality reduction techniques may speed up computational time, there should be minimal to no loss of information regarding which flights are identified as anomalous. To that end, a comparison is made of the set of anomalous flights obtained with and without PCA for the same feature vector choice. In each case of feature vector, the exact same set of anomalous flights is obtained.

D. Anomaly Detection

The next step in the framework is to use the feature vectors thus generated in an anomaly detection algorithm. A number of anomaly detection algorithms have been proposed in the literature for different types of problems. Various considerations need to be taken into account when selecting the most appropriate anomaly detection techniques. Since the data obtained from routine operations is not labeled, there is no prior knowledge of anomalies in the flight data (if there exist any at all). Therefore, only semi-supervised or unsupervised learning techniques may be utilized. However, even with these techniques, the number of anomalies is expected to be a small fraction of all the flight data records. This information should be utilized in tuning the parameters of the algorithm. In the literature, classification (in the form of one-class learning) [16, 17, 33] and clustering [17, 18, 25, 27] have been utilized extensively for anomaly detection in flight data records. In the current work, both clustering and classification have been used together to identify anomalous flight data records.

In *clustering-based* approaches, a clustering algorithm is applied on the feature vectors using a function to compare the dissimilarity between feature vectors. One of the most common functions used in clustering literature is the euclidean distance. Let the feature vectors corresponding to two flight records be $f_1 = m_1, m_2, ..., m_p$ and $f_2 = n_1, n_2, ..., n_p$. Then the euclidean distance between f_1 and f_2 is given as:

$$Df_1, f_2 = \sqrt{\sum_{i=1}^{p} m_i - n_i{}^2} \tag{4}$$

Other measures of distance or dissimilarity may also be used. In typical GA operations, the number of clusters present in the data is difficult to predict a priori. Therefore, algorithms which automatically identify the number of clusters (such as density-based clustering) are preferred over those which require specification of the number of clusters (such as k-means, k-medoids etc.) In the current work, the popular density-based clustering algorithm DBSCAN is used [34]. Given a set of points (flight data records), DBSCAN groups together instances that are closely packed together while marking points in low-density regions as outliers. A cluster forms when there are at least a minimum number of points (hereafter called $MinPts$) within a user specified threshold (hereafter called ϵ) of a given point. Clusters grow when additional points satisfy the density criterion specified by the algorithm until all the points have been allotted to a cluster or labeled as outliers.

There are two parameters that need to be supplied to DBSCAN - ϵ and $MinPts$. ϵ depends heavily on the similarity function used, normalization of data, and other factors. In many cases, rather than providing ϵ, its value is varied from the minimum distance observed among all flights in the data set to the maximum distance observed. Instead of ϵ, the user provides the proportion of flights that will be marked as outliers. This number has a direct correlation with the value of ϵ but is more intuitive to the user of the methodology. $MinPts$ on the other hand, depends on the homogeneity of operations and how similar flights are to each other in terms of the features chosen. Typically, $MinPts$ has a less significant effect on the algorithm than ϵ if it is within a nominal range of values [18, 34].

The main advantage of DBSCAN is that it has the ability to automatically determine the number of clusters and also detect outliers (anomalies) based on a user specified threshold. One of the drawbacks of DBSCAN is that its performance can suffer if there are multiple clusters with varying densities. Another drawback is that the algorithm needs to calculate the pairwise distance between all flights which can be computationally intensive. Also, in order to get a specific proportion of flights as outliers, the algorithm needs to try out different values of ϵ. Despite these limitations, the algorithm's most important quality of automatically determining the number of clusters will be utilized in this work. The algorithm provided by Ester et al. [34] is implemented in MATLAB for use in this work.

For *classification-based* approaches, a one-class classification model is trained assuming the available data set as nominal. One of the most powerful algorithms that is widely used is one-class Support Vector Machines (SVM) [35]. SVMs compare feature vectors using functions known as *kernels*. Kernel functions map pairs of feature vectors to the similarity between those vectors, with a value of 1 indicating maximum similarity and 0 indicating no similarity. Das et al. [16] used normalized Longest Common Subsequence (nLCS) kernel on discrete data and SAX-discretized data (which converts a continuous time series of parameters into discrete symbols by averaging the data over a number of seconds into a single symbol). The limitations of this approach were highlighted previously in the feature generation step. Therefore, in this work, Radial Basis Function (RBF) kernel - one of the most popular kernel functions is used. Let the feature vectors corresponding to two flight records be $f_1 = m_1, m_2, ..., m_p$ and $f_2 = n_1, n_2, ..., n_p$. Then the

RBF kernel is given by:

$$K f_1, f_2 = \exp\left(-\frac{||\boldsymbol{f_1} - \boldsymbol{f_2}||}{2\sigma^2}\right) \tag{5}$$

σ is a parameter that defines the scale of the classifier. The one-class SVM then uses this kernel function to find a decision boundary (by solving an optimization problem) to detect outliers. If the decision function predicts a negative label for a given test point z, then it is classified as an outlier. One of the main advantages of training a SVM model is that upon training, the prediction of anomalies is much quicker as it only has to use a fraction of the training points (called support vectors) in the prediction stage. The balance between model complexity and over-fitting can be managed by varying the proportion of points retained as support vectors. Details of the optimization problem and SVM implementation can be found in Schölkopf et al [35]. In the current work, the one-class SVM implementation in MATLAB is utilized and tuned for the GA flight data [36].

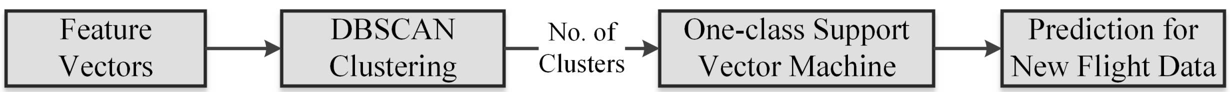

Fig. 3 Use of anomaly detection algorithms in the methodology

Figure 3 shows the algorithms used in the current work for identifying flight level anomalies. Compared to clustering methods, model-based methods like classification can be computationally more efficient in the detection stage by not requiring comparison of the new flight data record with each existing record in the database. In the present work, both DBSCAN clustering and one-class SVM are utilized. On an existing set of flight data records, DBSCAN is used to identify the approximate number of clusters in the existing data set. Each flight data record in the data set is then assumed to have a label of the cluster to which it belongs and this information is used to train a support vector machine model for each cluster identified. The trained SVM model(s) are then used to identify outliers from each class. The anomalousness of any existing or new flight data records can thus be directly obtained using the SVM score. Using both algorithms in this manner provides a consensus and enables combining the strengths of both algorithms while trying to overcome their individual weaknesses.

E. Post-processing

Once the anomaly detection algorithm has identified flight level anomalies, the post-processing step involves visualizing them and organizing according to their anomalousness. For visualization, flight parameter plots similar to those shown in Li et al. [18] can be used; but instead for the energy metrics. Since each energy metric provides information about a different dimension of the safety space, such plots of the energy metrics can provide a more holistic view of the safety state. Combinations of energy metrics going outside permissible bounds could be potentially more dangerous than the exceedance of individual metrics/parameters. The anomalousness of each flight data record can be visualized using anomaly scores such as Local Outlier Factor (LOF) [14] for clustering algorithms, whereas for classification, the support vector machine model predicts an anomaly score for each record which can be visualized for this purpose. In this paper, the anomaly score provided directly by the one-class SVM is utilized.

IV. Implementation and Results

All the steps of the methodology outlined earlier are demonstrated here on actual flight data records. First, the steps of pre-processing and feature generation are implemented to evaluate the energy metrics (formulas in Table 2 in the appendix) for all the flight records in the data set. In order to obtain reference energy profiles, the data driven approach outlined in previous work (Puranik et al. [12]) is utilized. Following generation of features, the parameters of the anomaly detection algorithms are fine-tuned for this data set. The first step in the anomaly detection step is tuning parameters and identification of number of clusters using DBSCAN clustering algorithm. As noted subsequently (Table 1), different subsets of energy metrics can be used in the clustering and anomaly detection phase of the methodology. Figure 4 shows the results obtained from the application of the clustering algorithm on the approach and landing phase of the current data set using all energy metrics (FV-4 of Table 1). In a similar manner, the other subsets of energy metrics are also investigated in the preliminary clustering to obtain their sensitivities. All the subsets considered in this work exhibited similar behavior to that of Figure 4. Therefore, rather than presenting each individually, the major trends are explained using Figure 4.

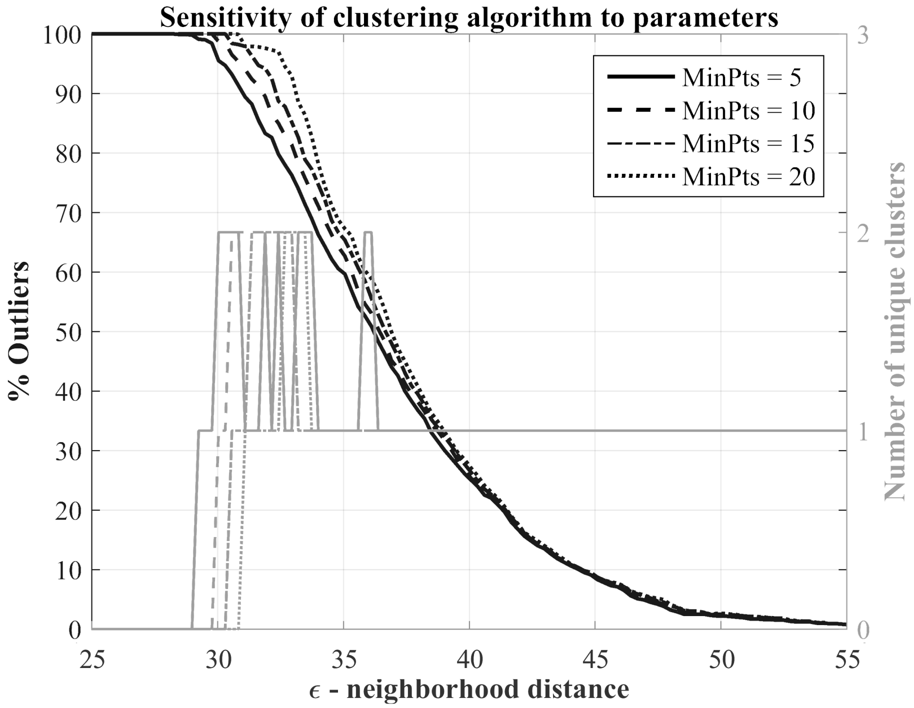

Fig. 4 Sensitivity of clustering algorithm to parameters

It is observed that, for all values of MinPts, as the value of ϵ increases from the minimum to the maximum value, the proportion of flights that are identified as outliers (black curves) steadily decreases. This is expected as more points will be included in the clusters when the cluster radius is increased. The curves for different values of MinPts eventually collapse almost into a single curve as ϵ is increased further. For the type of data dealt with in aviation safety, the proportion of anomalous or abnormal flights is expected to be very low (as seen from the low accident and incident rates per million flight hours). Therefore, at the values of ϵ that correspond to low outlier percentages, any value of MinPts from the set chosen gives almost equivalent results. This important observation leads to the conclusion that the value of MinPts can be set to a default of 5 for the purpose of this work.

The second set of curves (grey) from Figure 4 corresponds to the number of unique clusters at each setting of ϵ and MinPts. The trend observed is that as the value of ϵ increases, the number of clusters settles at one after initial oscillation. One reason for this oscillation is the nature of DBSCAN algorithm itself, in particular, the manner in which it starts forming clusters. At lower values of the neighborhood distance, individual points in dense neighborhoods start their own cluster by becoming 'core' points. As the distance increases, clusters consolidate until a new 'core' has enough points around itself to start its own cluster. This process of new cluster formation and consolidation causes the initial oscillations as seen in Figure 4. However, at the outlier significance levels of interest ($< 10\%$) there is only a single cluster present among the current set of flight data records. It is understood that this might change/update as more data is added to the set from other aircraft/airports etc. and therefore, the clustering algorithm will be periodically re-run to update this. Therefore, for the one-class support vector machine, all the flights will be given a default label of 1 prior to fitting the model. The one-class support vector machine is then implemented on this data set assuming all records nominally belong to a single cluster with an expected outlier percentage of 5% (this number can be changed based on the user preference). The algorithm automatically identifies anomalous flight data records using the trained model. The anomaly scores of each flight record thus obtained are plotted in Figure 5. The flights with scores greater than zero (depicted in gray color) are classified as normal while those with scores lower than zero (black color) are anomalies.

It is important for the methodology to be able to identify known anomalies as well as emergent artifacts that are not explicitly labeled as anomalies a priori. Therefore, the following two subsections present each of these types of anomalous flight data records identified by the methodology that have been annotated as 'simulated anomaly' and 'example anomaly in flight data' in Figure 5.

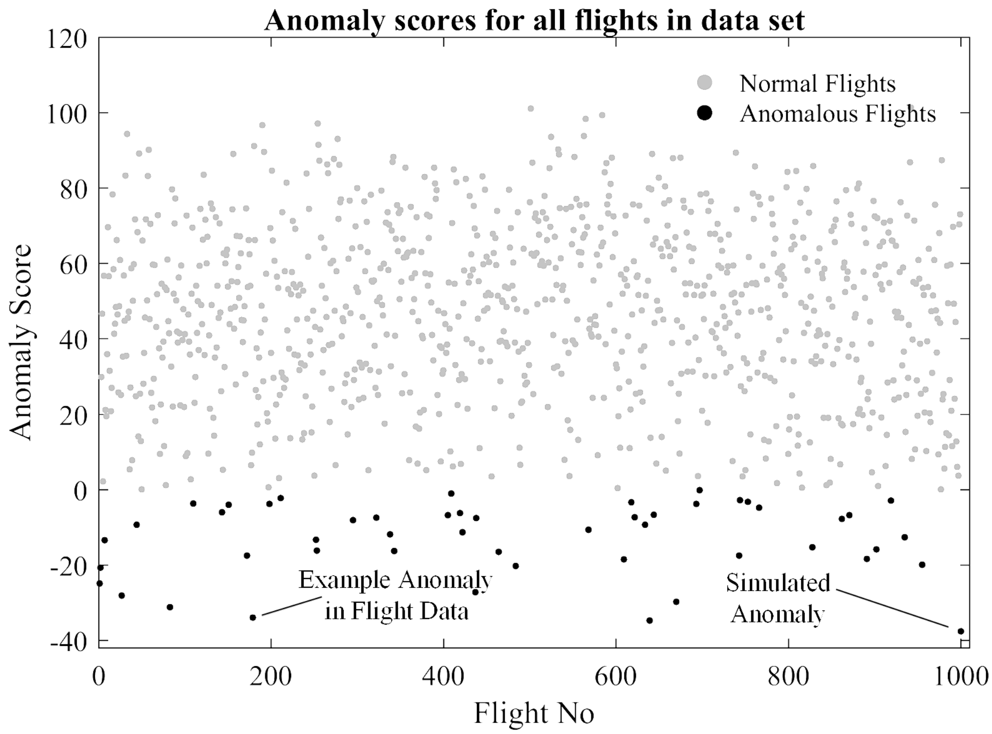

Fig. 5 Anomaly scores of all the flights in the data set

A. Artificial Anomaly
Within a set of flight data records from routine operations, flights that feature anomalous occurrences (if any) are not labeled as such. This makes it difficult to validate different methodologies for anomaly detection using such a set of routine flights alone. Therefore, in order to facilitate validation efforts, an artificial flight data record, generated using a flight simulation model, is included in the data set. The dynamics of the aircraft in this model are simulated using a MATLAB/Simulink model which is connected to FlightGear Flight Simulator to allow visual rendering of the motion of the aircraft. Further details of the model can be found in Chakraborty et al [37]. The simulated flight record can be set up to output a G1000-style data record with the same parameters as recorded in actual flights in the data set. Similar validation has previously been done by Chu et al. [26] at NASA using data from a flight simulator. The simulated approach and landing are flown by a private pilot instructed to deliberately fly an unstabilized 3 NM final approach with poor energy management. In the aviation safety community, this is classified as a high and fast approach and can lead to accidents like Runway Excursion (RE) or CFIT.

The simulated anomalous flight is included in the set of routine flights and the anomaly score of this record is also calculated (shown in Figure 5). A visualization of some important energy metrics during this approach and landing, along with variation of flight parameters is shown in Figures 6,7. In both figures, the dark gray bands represent 50^{th} percentile of all flight records the light gray bands represent the 95^{th} percentile of all flight records. The solid black line represents the anomalous flight data record that has been identified by the methodology. As is clearly evident, from the plots and the anomaly scores, this known anomalous approach and landing is correctly captured using this methodology and placed in the region containing flights with high anomaly scores from the data set. Therefore, this methodology is able to identify known anomalies which could potentially be unsafe that are introduced using a flight simulation model.

B. Anomaly in Flight Data
Along with the identification of known anomalies, one of the most useful attributes of anomaly detection techniques is the automatic identification of flight data records that were behaving abnormally or outside the bounds of nominal operations. Therefore, the data records with high anomaly scores (from Figure 5) are some of the flights that would be of immediate interest to analysts or instructors to identify various issues. Once the list of anomalous flight records is

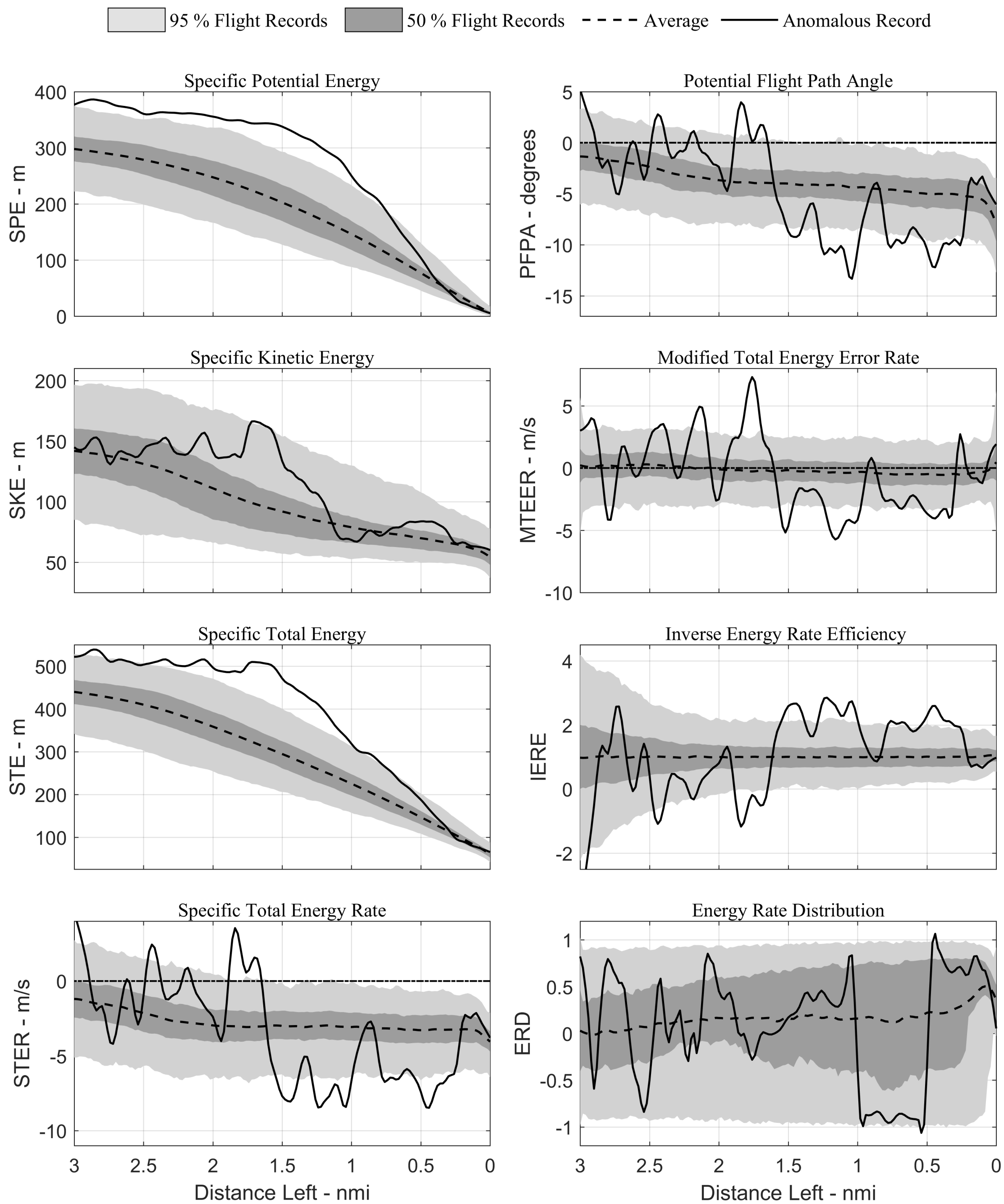

Fig. 6 Visualization of energy metrics for artificial anomaly

generated, one of the top anomalous flights from the data set having a high anomaly score is visualized in Figures 8, 9. Similar visualization and post-processing can be done for other anomalous flights identified.

The variation of some critical energy metrics and flight parameters during the approach and landing phase for the anomalous flight along with the bands for nominal variation within the data set can be seen in the figures. It is clearly evident that the flight data record is deviating from the norm on various occasions in the approach in terms of a

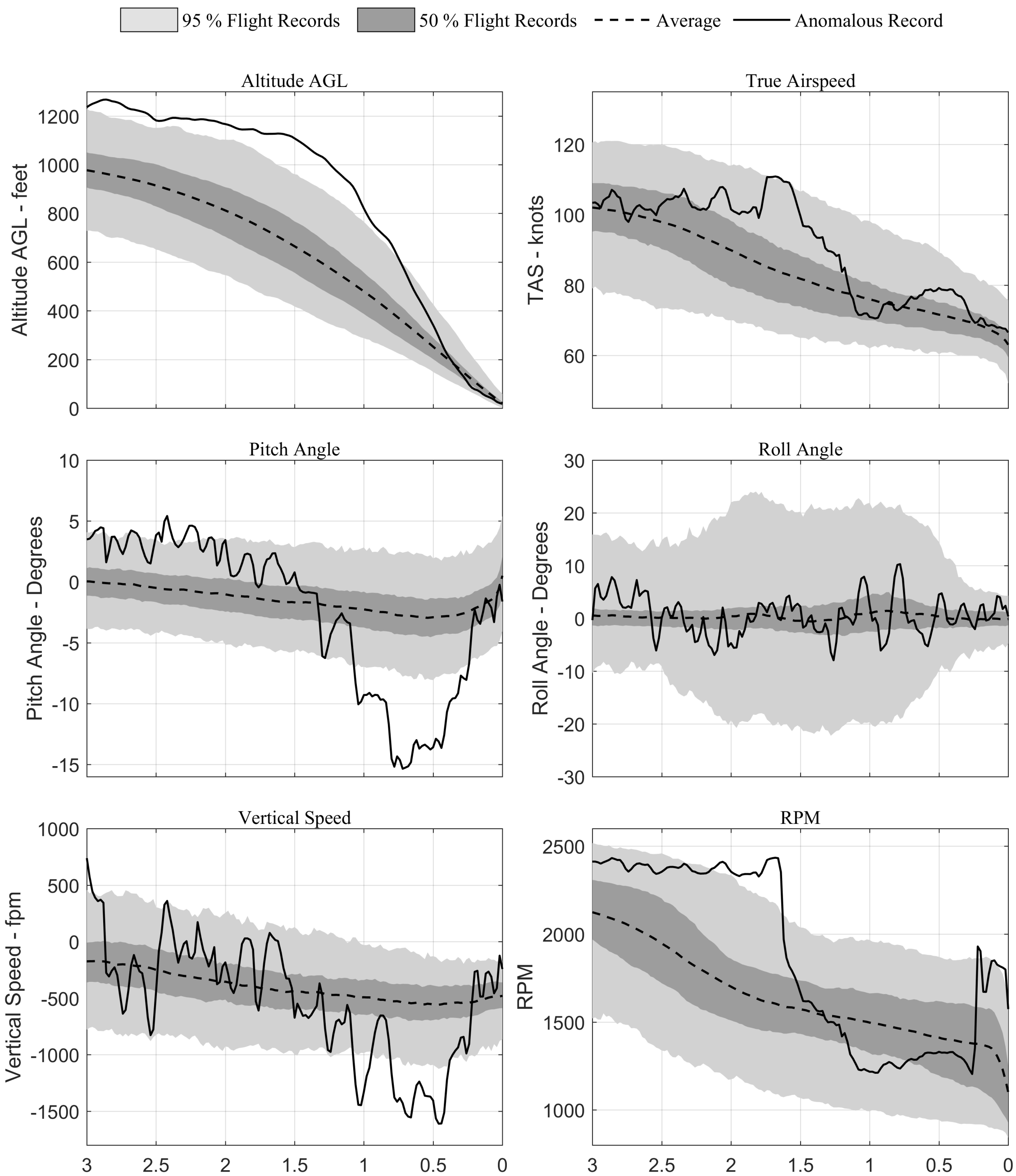

Fig. 7 Visualization of a subset of flight data parameters metrics for artificial anomaly

few key metrics (such as specific kinetic energy, thrust margin, and specific total energy). Based on the visualization of parameters and energy metrics, it could be classified as a low and fast approach and could be considered unsafe operations. The visualization of parameters shows that the Revolutions Per Minute (RPM), Vertical Speed, Altitude Above Ground Level (AGL), and Airspeed variation is outside of the bounds for majority of flight records and the approach has elements of an unstable approach as defined by Flight Safety Foundation [38].

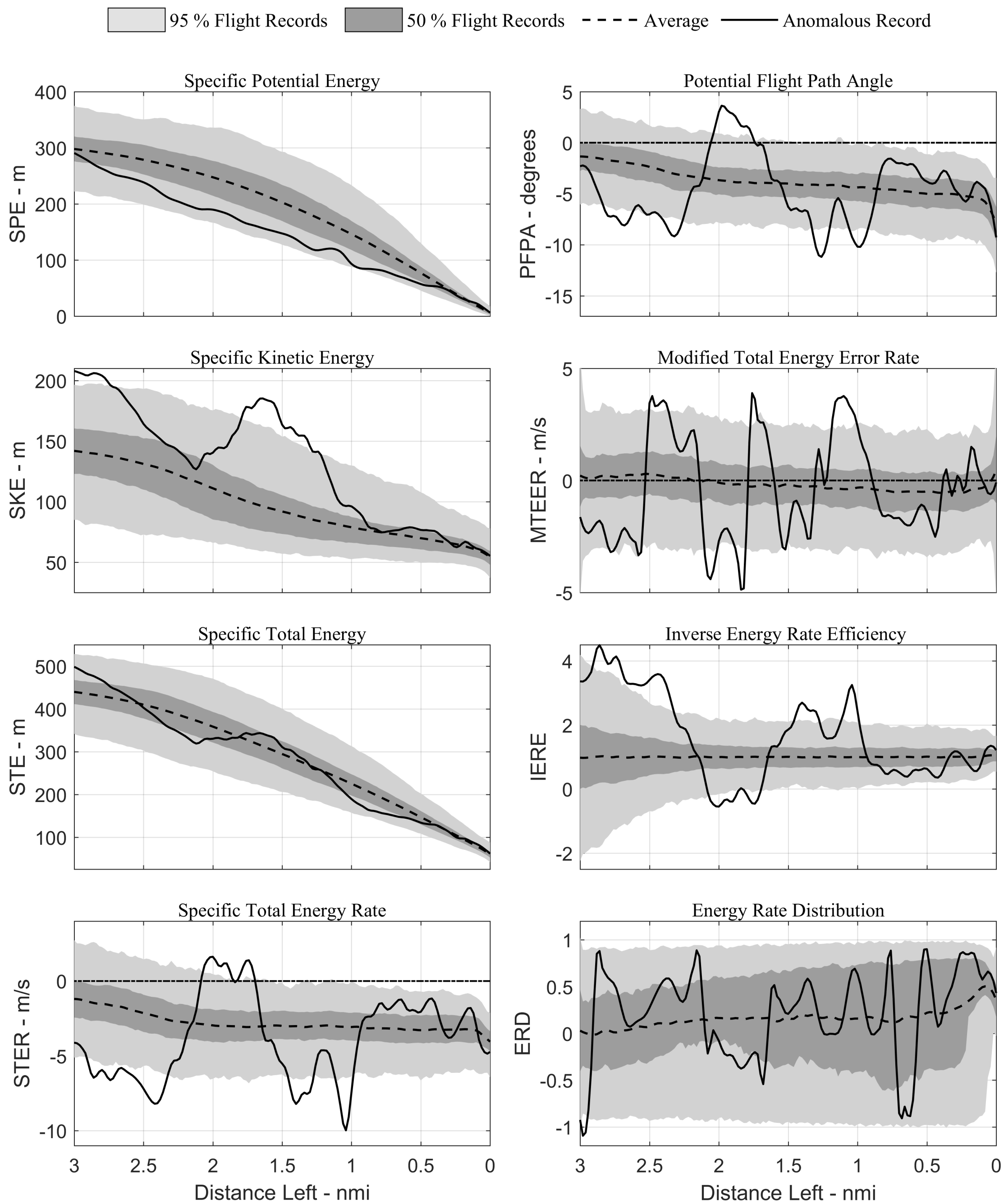

Fig. 8 Visualization of energy metrics for a flight data record with fast approach

Therefore, we see that this methodology is able to identify flight records that have high anomaly scores and upon visualization corroborate that the approach and landing deviates from nominally observed operations. In both the cases, it is observed that various key energy metrics deviate outside of the bounds observed during nominal operations. These could signal potential safety issues and therefore, may be selected for further inspection by the safety analyst.

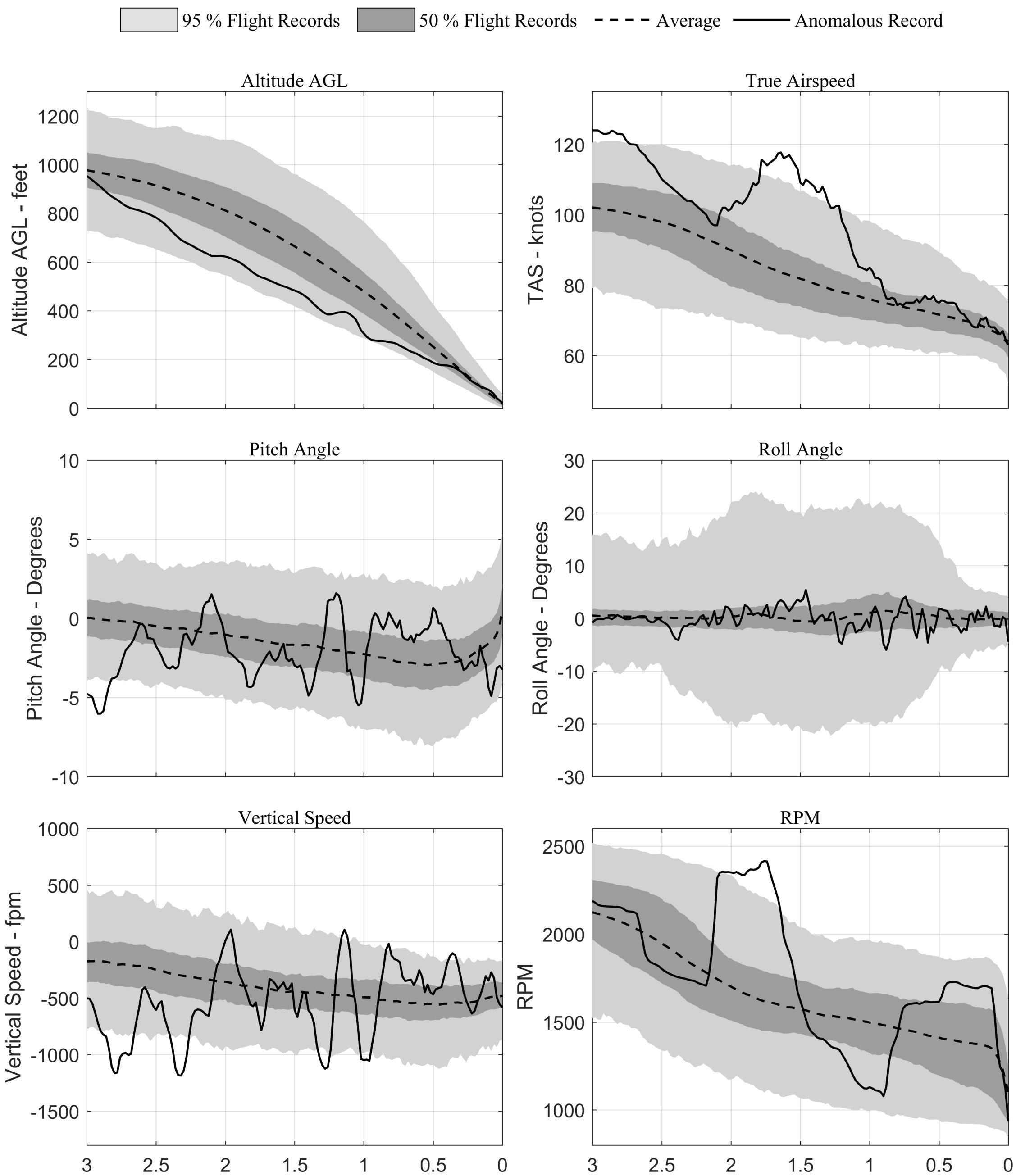

Fig. 9 Visualization of a subset of flight data parameters metrics for a flight data record with fast approach

C. Comparative Evaluation

It was noted earlier that in traditional FOQA analysis in commercial operations, unsafe behaviors are typically captured by safety events by different operators. These events typically consist of exceedance of one or more parameters beyond thresholds as defined by subject matter experts. Similar safety events can be defined for GA operations such as those by Higgins et al [39]. However, the definitions of the safety events assumes recording of a certain minimum set of parameters (most noteworthy is the flap position) which are not necessarily recorded on all GA aircraft. The Garmin G1000 (which is the higher end of instrumentation of GA aircraft) data available for this study did not have flap position recorded which makes it difficult to identify these events. Due to these reasons, a comparison to traditional exceedance detection events is not pursued. However, it is nevertheless of interest to examine how the methodology would perform in the presence of limited data recording capability which is a very realistic possibility in GA safety studies.

It is desired that data obtained from multiple types of recorders be utilized in the same unified anomaly detection environment. In order to facilitate this, the defined energy metrics are divided into three types of feature vectors based on less stringent data recorder requirements. This is achieved by omitting those additional parameters in the feature vector options that do not assume them available. The exact parameters that are used in each feature vector option can be seen in Table 3. The flight level anomaly detection methodology is then run using each feature vector (in addition to not using energy metrics at all which is FV-1) to compare resulting anomalous flights obtained. It should be noted that FV-1 (not using energy metrics but only raw parameters) is similar to the ClusterAD method developed by Li et al. [18] Therefore, this comparison will also demonstrate the differences between the energy metrics approach and the existing approach in literature.

Table 1 Feature vector options and description of contents

Name	Features	Description
FV-1	Raw Flight Parameters	Set of all flight parameters collected from flight data
FV-2	Energy Metrics Subset 1	Set of all energy metrics from Table 2 that can be obtained from basic flight data alone
FV-3	Energy Metrics Subset 2	Set of all energy metrics from Table 2 that can be obtained from basic flight data and reference energy profiles
FV-4	All Energy Metrics	Set of all energy metrics listed in Table 2

In each case, the top 5% anomalous flights are selected from the entire data set of 1000 flight records (50 anomalous flights in each case). The total amount of overlap in the anomalous flights is then obtained to identify which flight records get missed when limited information is available. Since there is no truth value to compare against, it is desired that all the top anomalous flights be the same or similar despite limitations in data recording capability. The overlap among the 50 anomalous flights identified by each feature vector option is shown in Figure 10.

The following observations can be made from the results in Figure 10. First, there is a fair amount of overlap among the anomalies identified by different feature vector options. The feature vectors only from energy metrics (FV-2, FV-3, FV-4) have 40 out of the 50 anomalous flights in common. Therefore, even with limited amount of information available for evaluating energy metrics, the methodology is able to retain 80% of the same anomalies. This implies that even if flight data records with limited amount of parameters are included in the data set, there is a good probability of identifying anomalous ones among these flights. Second, it is noted that there are only few flights that do not get recognized with energy metrics obtained from limited parameters and/or no performance models. These flight records have anomalies related to the specific metrics that cannot be evaluated without a richer data set or performance models. Thus, overall, there is consistency regarding which flights are considered anomalous using different energy metric subsets which is encouraging.

Using FV-1, which directly utilizes all the raw parameters from the flight data record, results in identification of $\approx 60\%$ of the same anomalies (29), but has 16 distinct anomalous flights than those obtained from any of the energy metrics feature vectors. The anomalies detected by FV-1 only, correspond to unusual variations of parameters that are not explicitly used in the energy metrics definitions (For example: Roll angle, Latitude, Longitude, Heading, Track, Fuel Quantity, Oil Temperature, Cylinder Head Temperature (CHT) etc.). The anomalies detected by FV-1 only can be of different types – flights that land at a different runway/airport (anomalous due to Latitude, Longitude, and Track etc.), flights that follow different procedures (to enter the traffic pattern for instance), flights that have abnormalities in other parameters not used in energy metrics (Oil Temperature, CHT etc.). While some of these types could have potential safety implications (Oil Temperature, CHT etc.), others such as those landing at different airport/runways do not necessarily need to be in the anomalous set. However, even for identifying those that are significant for safety

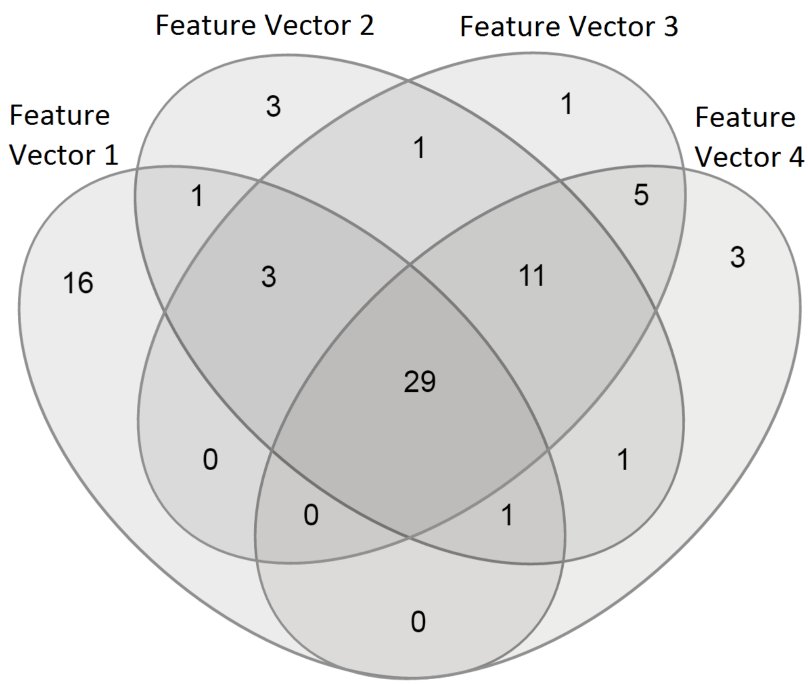

Fig. 10 Venn diagram showing overlap of anomalous flights as classified by the different feature vector options (Created using Herberle et al. [40])

analysis, the recording capability needs to be available for all the aircraft (which, as argued earlier, it is not). Therefore, using FV-1 only may be viewed as a complementary approach when the required data is available. It is, however, noted that the 29 common anomalies identified by each feature vector option were ranked at least within the top 35 anomalous flights of each individual set of 50 anomalies from different feature vectors. This points to the fact that the top anomalies (which would be of interest to safety analysts) are identified across the spectrum of feature vector options with the highest anomaly scores.

The main appeals of energy metrics approach (FV – 2, 3, 4) are - the requirement of a parsimonious set of parameters, (relative) independence of the size (weight) of the aircraft, applicability at multiple airport/runways. From the results shown in this paper and the previous discussions, it is evident that this approach can be easily used on a large variety of GA FDR data due to the limited amount of parameters required. This makes it ideal to be deployed on a large database such as the NGAFID where data sets from multiple types of GA FDR may be available. Additionally, unlike FV-1 approach, the results are insensitive to the runway used for approach and landing and are expected to be independent of airport as well (data in this study are for the same airport).

V. Conclusion

In this paper, the implementation of a generic anomaly detection methodology which utilizes energy-based metrics as features is demonstrated. It is shown that energy based metrics used in an anomaly detection framework consisting of a novel combination of clustering and one-class classification algorithms are able to quickly and efficiently identify anomalies in GA operations. Through specific examples, the implementation of this methodology and its potential in identifying anomalies as emergent artifacts from within a large data set is demonstrated. It was also verified that a simulated anomalous approach and landing does get identified as such using the algorithm, providing preliminary validation for the method in the absence of actual labeled anomalous flight data. Therefore, it is evident that this methodology is able to identify flight records that have high anomaly scores and upon visualization corroborate that the approach and landing are potentially unsafe, and warrant further inspection. Finally, the performance of the methodology in terms of anomalous flight identification was quantified when limited number of parameters are recorded or performance models are unavailable. It is shown that using the energy metrics approach results in a consistent identification of anomalies within the given data set and therefore has potential to be used across a heterogeneous GA fleet and operations. It is noted that GA operations include heterogeneity due to uncertain weather conditions and phenomenon such as cross winds, gusts, etc. The setup of the analysis ensures that the energy metrics are able to capture these phenomena implicitly because they are derived using flight parameters whose values are affected by weather phenomena. In particular, adverse/gusty/turbulent conditions may result in an approach/landing that is outside normal bounds and that may, therefore be tagged as anomalous. However, a capability to differentiate between purely weather-induced anomalies and those due to piloting procedures is at present lacking due to the unavailability of weather data in the FDR.

The implementation of the methodology and case studies highlight its potential in uncovering safety related events. The retrospective analysis of flights can be utilized by pilots and flight instructors to understand how their flights are performing compared to other pilots. This can also help uncover different behaviors during flight and potentially correct or improve flying skills. Better energy management by pilots using a data-driven approach will enable safer operations. De-identified aggregated data can also be used in this framework by safety analysts to understand how operations are actually being conducted and recommend changes or improvements.

Various ways of enhancing the applicability of this method are identified. The steps of the generic anomaly detection framework remain the same, but different alternatives or algorithms can be explored within each step. One of the main advantages of energy metrics is that they can be considered as an objective currency across different aircraft types. Therefore, demonstrating the use of flight data from different aircraft will enable the scope of this methodology to be widened. Methods of expanding this analysis to other phases of flight will be pursued to make it more generalizable. Similarly, as noted before, energy metrics have also been used to identify instantaneous anomalies [13]. In future work, a comparison between instantaneous and flight-level anomalies may be conducted to identify the correlation between these anomalies and to asses if successive instantaneous anomalies result in a flight-level anomaly. Finally, providing feedback to the algorithms using opinions of subject matter experts will increase the likelihood of identifying anomalies that are operationally significant from the perspective of continued safety.

Acknowledgments

This work was sponsored by the Federal Aviation Administration under Project No. 5 of the FAA Center of Excellence for General Aviation, PEGASAS (Partnership to Enhance General Aviation Safety Accessibility and Sustainability), via grant No. 12-C-GA-GIT-004 to the Georgia Institute of Technology. The authors are grateful to FAA project technical monitor Mr. Michael Vu and project partners. The information in this research does not constitute FAA Flight Standards or FAA Aircraft Certification policy. The authors would also like to thank Evan Harrison, Sanggyu Min, Dr. Hernando Jimenez, and Dr. Imon Chakraborty for their feedback and assistance.

A. Summary of Implemented Energy Metrics

Table 2 Summary of implemented energy metrics and data required for computation

		Can be estimated using		
Metric	**Formula**	**Flight Data**	**Flight Data + Ref. Profile**	**Flight Data + Perf. Model**
Specific Total Energy	$h + V^2 2g$	✓	✓	✓
Specific Potential Energy	h	✓	✓	✓
Specific Kinetic Energy	$V^2 2g$	✓	✓	✓
Specific Total Energy Rate	$h + VVg = T - DVW$	✓	✓	✓
Specific Potential Energy Rate	$h = V \sin \gamma$	✓	✓	✓
Specific Kinetic Energy Rate	VVg	✓	✓	✓
Potential Flight Path Angle	$\gamma + Vg$	✓	✓	✓
Energy Rate Distribution	$sign \frac{SKER}{SPER} \times \exp - \lvert \frac{SKER}{SPER} \rvert$	✓	✓	✓
Specific Total Energy Error	$h_{act} - h_{ref} + V_{act}^2 - V_{ref}^2 2g$	×	✓	×
Specific Potential Energy Error	$h_{act} - h_{ref}$	×	✓	×
Specific Kinetic Energy Error	$V_{act}^2 - V_{ref}^2 2g$	×	✓	×
Normalized Specific Energy Error	$STE_{act} - STE_{ref} STE_{tol}$	×	✓	×
Specific Total Energy Error Rate	$signSTEE \times \delta STEE \delta t$	×	✓	×
Inverse Energy Rate Efficiency	$V_{act} T - DV_{red} W \gamma_{ref} + V_{ref} g$	×	✓	×
Max. Potential Flight Path Angle	$T_{max} - DW$	×	×	✓
Min. Potential Flight Path Angle	$T_{idle} - DW$	×	×	✓
Thrust Margin	$1 - TT_{max}$	×	×	✓
Energy Rate Margin	$W\gamma_a + V_a g T_{max} - D$	×	×	✓
Energy Rate Demand	$W\gamma_c + V_c g T_{idle} - D$	×	×	✓

B. Summary of data required for various feature vectors

Table 3 Summary of parameters required for computation of various Feature Vectors (FV)

No	Parameter/Model	FV 1	FV 2	FV 3	FV 4
1	Altitude	✓	✓	✓	✓
2	True Airspeed	✓	✓	✓	✓
3	Vertical Speed	✓	✓	✓	✓
4	Ground Speed	✓	✓	✓	✓
5	Outside Air Temperature	✓			
6	Latitude	✓			
7	Longitude	✓			
8	Indicated Airspeed	✓			
9	Ground speed	✓			
10	Pitch	✓	✓		
11	Roll	✓			
12	Lateral Acceleration	✓			
13	Normal Acceleration	✓			
14	Heading	✓			
15	Track	✓			
16	Fuel Quantity (Left Tank)	✓			
17	Fuel Quantity (Right Tank)	✓			
18	Fuel Flow Rate	✓			
19	Oil Temperature	✓			
20	Oil Pressure	✓			
21	RPM	✓	✓		
22	Cylinder Head Temp.	✓			
23	Exhaust Gas Temp.	✓	✓		
24	Flight Path Angle	✓	✓	✓	✓
25	Reference Energy Profiles		✓	✓	
26	Aerodynamics Model		✓		
27	Propulsion Model		✓		

References

[1] Logan, T. J., "Error Prevention as Developed in Airlines," *International Journal of Radiation Oncology*Biology*Physics*, Vol. 71, No. 1, 2008, pp. S178–S181. doi:10.1016/j.ijrobp.2007.09.040.

[2] NTSB, "National Transportation Safety Board Website (Retrieved 10/2016)," , 2013. URL `http://www.ntsb.gov/investigations/data/Pages/AviationDataStats.aspx`, url: `http://www.ntsb.gov/investigations/data/Pages/AviationDataStats.aspx`.

[3] "Federal Aviaition Administration Fact Sheet - General Aviation Safety," , 2017. URL `https://www.faa.gov/news/fact_sheets/news_story.cfm?newsId=16774`, retrieved - 04/2017. url: `https://www.faa.gov/news/fact_sheets/news_story.cfm?newsId=16774`.

[4] Belcastro, C. M., and Foster, J. V., "Aircraft Loss-of-Control Accident Analysis," *AIAA Guidance, Navigation, and Control Conference*, 2010. Paper No. AIAA-2010-8004, doi:10.2514/6.2010-8004.

[5] "Advisory Circular, 120-82 - Flight Operational Quality Assurance," , April 2004. Url: `https://www.faa.gov/regulations_policies/advisory_circulars/index.cfm/go/document.information/documentID/23227`.

[6] "Federal Aviation Administration, 14 CFR §121.344 Digital Flight Data Recorders for Transport Category Airplanes, 2011," , ????. Url: `https://www.ecfr.gov/cgi-bin/text-idx?tpl=/ecfrbrowse/Title14/14cfr121_main_02.tpl`.

[7] "Federal Aviation Administration, 14 CFR §23.3 Airplane Categories," , ????. Url: `https://www.ecfr.gov/cgi-bin/text-idx?node=pt14.1.23`.

[8] Harrison, E., Min, S., Jimenez, H., and Mavris, D. N., "Implementataion and Validation of an Internal Combustion Engine and Propeller Model for General Aviation Aircraft Performance Studies," *AIAA Aviation*, American Institute of Aeronautics and Astronautics, 2015. doi:10.2514/6.2015-2850, URL `http://dx.doi.org/10.2514/6.2015-2850`, paper No. AIAA-2015-2850, doi:10.2514/6.2015-2850.

[9] Min, S., Harrison, E., Jimenez, H., and Mavris, D. N., "Development of Aerodynamic Modeling and Calibration Methods for General Aviation Aircraft Performance Analysis - a Survey and Comparison of Models," *AIAA Aviation*, American Institute of Aeronautics and Astronautics, 2015. doi:10.2514/6.2015-2853, URL `http://dx.doi.org/10.2514/6.2015-2853`, paper No. AIAA 2015-2853, doi:10.2514/6.2015-2853.

[10] Puranik, T., Harrison, E., Min, S., Jimenez, H., and Mavris, D., "Energy-Based Metrics for General Aviation Flight Data Record Analysis," *16th AIAA Aviation Technology, Integration, and Operations Conference*, 2016. Paper No. AIAA 2016-3915, doi:10.2514/6.2016-3915.

[11] Puranik, T., Jimenez, H., and Mavris, D., "Energy-Based Metrics for Safety Analysis of General Aviation Operations," *Journal of Aircraft*, 2017. doi:10.2514/1.C034196.

[12] Puranik, T., Harrison, E., Min, S., Jimenez, H., and Mavris, D., "General Aviation Approach and Landing Analysis Using Flight Data Records," *16th AIAA Aviation Technology, Integration, and Operations Conference*, 2016. Paper No. AIAA 2016-3913, doi:10.2514/6.2016-3913.

[13] Puranik, T. G., and Mavris, D. N., "Identifying Instantaneous Anomalies in General Aviation Operations," *17th AIAA Aviation Technology, Integration, and Operations Conference*, 2017. doi:10.2514/6.2017-3779, paper No: AIAA-2017-3779, doi:10.2514/6.2017-3779.

[14] Breunig, M. M., Kriegel, H.-P., Ng, R. T., and Sander, J., "LOF: Identifying Density-based Local Outliers," *ACM sigmod record*, Vol. 29, ACM, 2000, pp. 93–104. doi:10.1145/342009.335388.

[15] Campos, G. O., Zimek, A., Sander, J., Campello, R. J., Micenková, B., Schubert, E., Assent, I., and Houle, M. E., "On the evaluation of unsupervised outlier detection: measures, datasets, and an empirical study," *Data Mining and Knowledge Discovery*, 2015, pp. 1–37. doi:10.1007/s10618-015-0444-8.

[16] Das, S., Matthews, B. L., Srivastava, A. N., and Oza, N. C., "Multiple Kernel Learning for Heterogeneous Anomaly Detection: Algorithm and Aviation Safety Case Study," *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, 2010, pp. 47–56. doi:10.1145/1835804.1835813.

[17] Matthews, B., Das, S., Bhaduri, K., Das, K., Martin, R., and Oza, N., "Discovering Anomalous Aviation Safety Events Using Scalable Data Mining Algorithms," *Journal of Aerospace Information Systems*, Vol. 10, No. 10, 2013, pp. 467–475. doi:10.2514/1.I010080, URL `http://dx.doi.org/10.2514/1.I010080`, doi:10.2514/1.I010080.

[18] Li, L., Das, S., John Hansman, R., Palacios, R., and Srivastava, A. N., "Analysis of Flight Data Using Clustering Techniques for Detecting Abnormal Operations," *Journal of Aerospace Information Systems*, Vol. 12, No. 9, 2015, pp. 587–598. doi:10.2514/1.I010329, URL `http://dx.doi.org/10.2514/1.I010329`, doi:10.2514/1.I010329.

[19] Melnyk, I., Matthews, B., Valizadegan, H., Banerjee, A., and Oza, N., "Vector Autoregressive Model-Based Anomaly Detection in Aviation Systems," *Journal of Aerospace Information Systems*, 2016, pp. 161–173. doi:10.2514/1.I010394.

[20] Gorinevsky, D., Matthews, B., and Martin, R., "Aircraft anomaly detection using performance models trained on fleet data," *Intelligent Data Understanding (CIDU), 2012 Conference on*, IEEE, 2012, pp. 17–23. doi:10.1109/CIDU.2012.6382196.

[21] Chandola, V., Banerjee, A., and Kumar, V., "Anomaly detection: A survey," *ACM computing surveys (CSUR)*, Vol. 41, No. 3, 2009. doi:10.1145/1541880.1541882.

[22] Liao, T. W., "Clustering of time series data - a survey," *Pattern Recognition*, Vol. 38, No. 11, 2005, pp. 1857 – 1874. doi:http://dx.doi.org/10.1016/j.patcog.2005.01.025, URL `http://www.sciencedirect.com/science/article/pii/S0031320305001305`, doi:10.1016/j.patcog.2005.01.025.

[23] Gavrilovski, A., Jimenez, H., Mavris, D. N., Rao, A. H., Shin, S., Hwang, I., and Marais, K., "Challenges and Opportunities in Flight Data Mining: A Review of the State of the Art," *AIAA SciTech Forum*, American Institute of Aeronautics and Astronautics, 2016. doi:10.2514/6.2016-0923, URL `http://dx.doi.org/10.2514/6.2016-0923`, paper No: AIAA 2016-0923 doi:10.2514/6.2016-0923 .

[24] Iverson, D. L., "Inductive System Health Monitoring," Tech. rep., National Aeronautics and Space Administration, 2004. Url: `https://ntrs.nasa.gov/search.jsp?R=20040068062`.

[25] Budalakoti, S., Srivastava, A. N., and Otey, M. E., "Anomaly Detection and Diagnosis Algorithms for Discrete Symbol Sequences with Aapplications to Airline Safety," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, Vol. 39, No. 1, 2009, pp. 101–113. doi:10.1109/TSMCC.2008.2007248.

[26] Chu, E., Gorinevsky, D., and Boyd, S., "Detecting Aircraft Performance Anomalies from Cruise Flight Data," *AIAA Infotech@Aerospace, Atlanta, Georgia.*, 2010. Paper No: AIAA 2010-3307. doi:10.2514/6.2010-3307.

[27] Bay, S. D., and Schwabacher, M., "Mining Distance-based Outliers in Near Linear Time with Randomization and a Simple Pruning Rule," *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, 2003, pp. 29–38. doi:10.1145/956750.956758.

[28] Goblet, V., Fala, N., and Marais, K., "Identifying Phases of Flight in General Aviation Operations," *AIAA Aviation*, American Institute of Aeronautics and Astronautics, 2015. doi:10.2514/6.2015-2851, URL `http://dx.doi.org/10.2514/6.2015-2851`, paper No: AIAA 2015-2851. doi:10.2514/6.2015-2851.

[29] Garmin, *G1000 Integrated Flight Deck Pilot's Guide*, 2017. Retrieved 03/2017. url: `http://static.garmincdn.com/pumac/190-00498-07_0A_Web.pdf`.

[30] Lin, J., Keogh, E., Lonardi, S., and Chiu, B., "A symbolic representation of time series, with implications for streaming algorithms," *Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery*, ACM, 2003, pp. 2–11. doi:10.1145/882082.882086.

[31] Hotelling, H., "Analysis of a complex of statistical variables into principal components." *Journal of educational psychology*, Vol. 24, No. 6, 1933, pp. 417–441. doi:10.1037/h0071325.

[32] Amidan, B. G., and Ferryman, T. A., "APMS SVD Methodology and Implementation," Tech. rep., U.S. Department of Energy PNWD-3026, 2000. doi:10.2172/753847.

[33] Das, S., Matthews, B. L., and Lawrence, R., "Fleet level anomaly detection of aviation safety data," *Prognostics and Health Management (PHM), 2011 IEEE Conference on*, IEEE, 2011, pp. 1–10. doi:10.1109/ICPHM.2011.6024356.

[34] Ester, M., Kriegel, H.-P., Sander, J., and Xu, X., "A density-based algorithm for discovering clusters in large spatial databases with noise." *Kdd*, Vol. 96, 1996, pp. 226–231.

[35] Schölkopf, B., Platt, J. C., Shawe-Taylor, J., Smola, A. J., and Williamson, R. C., "Estimating The Support of a High-dimensional Distribution," *Neural computation*, Vol. 13, No. 7, 2001, pp. 1443–1471. doi:10.1162/089976601750264965.

[36] Mathworks, "Matlab Online Documentation," , 2017. URL `http://www.mathworks.com/help/matlab/`, url: `http://www.mathworks.com/help/matlab/`.

[37] Chakraborty, I., Lozano, B., and Mavris, D. N., “Pilot-Friendliness Considerations for Personal Air Vehicle Flight Control Systems,” *AIAA Aviation*, American Institute of Aeronautics and Astronautics, 2015. doi:10.2514/6.2015-2852, URL http://dx.doi.org/10.2514/6.2015-2852, paper No. AIAA 2015-2852, doi:10.2514/6.2015-2852.

[38] FSF, “Flight Safety Foundation ALAR Briefing Note 7.1 - Stabilized Approach,” Technical Report, August–November 2000. URL https://flightsafety.org/files/alar_bn4-2-energymgmt.pdf, url: https://flightsafety.org/wp-content/uploads/2016/09/alar_bn7-1stablizedappr.pdf.

[39] Higgins, J., Clachar, S., and Hennselek, K., “Flight Data Monitoring (FDM)/Flight Operational Quality Assurance (FOQA) for General Aviation,” Tech. rep., Federal Aviation Administration, December 2013.

[40] Heberle, H., Meirelles, G. V., da Silva, F. R., Telles, G. P., and Minghim, R., “InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams,” *BMC bioinformatics*, Vol. 16, No. 169, 2015. doi:10.1186/s12859-015-0611-3.