Finding Precursors to Anomalous Drop in Airspeed During a Flight's Takeoff

Vijay Manikandan Janakiraman
USRA/ NASA Ames Research Center
Moffett Field, CA 94035, USA
vjanakir@mail.nasa.gov

Bryan Matthews
USRA/ NASA Ames Research Center
Moffett Field, CA 94035, USA
bryan.l.matthews@nasa.gov

Nikunj Oza
NASA Ames Research Center
Moffett Field, CA 94035, USA
nikunj.c.oza@nasa.gov

ABSTRACT

Aerodynamic stall based loss of control in flight is a major cause of fatal flight accidents. In a typical takeoff, a flight's airspeed continues to increase as it gains altitude. However, in some cases, the airspeed may drop immediately after takeoff and when left uncorrected, the flight gets close to a stall condition which is extremely risky. The takeoff is a high workload period for the flight crew involving frequent monitoring, control and communication with the ground control tower. Although there exists secondary safety systems and specialized recovery maneuvers, current technology is reactive; often based on simple threshold detection and does not provide the crew with sufficient lead time. Further, with increasing complexity of automation, the crew may not be aware of the true states of the automation to take corrective actions in time. At NASA, we aim to develop decision support tools by mining historic flight data to proactively identify and manage high risk situations encountered in flight. In this paper, we present our work on finding precursors to the anomalous drop-in-airspeed (ADA) event using the ADOPT (Automatic Discovery of Precursors in Time series) algorithm [12]. ADOPT works by converting the precursor discovery problem into a search for sub-optimal decision making in the time series data, which is modeled using reinforcement learning. We give insights about the flight data, feature selection, ADOPT modeling and results on precursor discovery. Some improvements to ADOPT algorithm are implemented that reduces its computational complexity and enables forecasting of the adverse event. Using ADOPT analysis, we have identified some interesting precursor patterns that were validated to be operationally significant by subject matter experts. The performance of ADOPT is evaluated by using the precursor scores as features to predict the drop in airspeed events.

CCS CONCEPTS

•Information systems →Data mining; •Applied computing →Aerospace; •Computing methodologies →*Machine learning approaches;*

KEYWORDS

Aircraft Stall, Aviation Data Mining, Precursor, Reinforcement Learning, Anomaly Detection, Prognostics, PHM

ACM Reference format:
Vijay Manikandan Janakiraman, Bryan Matthews, and Nikunj Oza. 2017. Finding Precursors to Anomalous Drop in Airspeed During a Flight's Takeoff. In *Proceedings of KDD '17, August 13-17, 2017, Halifax, NS, Canada, ,* 10 pages.
DOI: 10.1145/3097983.3098097

1 INTRODUCTION

Loss of control (LOC) in flight is one of the significant causes of fatal aircraft accidents [2, 3, 16]. Loss of control occurs when the controllability of the flight is lost because it deviates from the safe operating limits. Some of the human induced causes of LOC include flight mismanagement by the crew, erroneous decision making, erroneous performance calculations and loss of situational awareness particularly during high workload situations such as takeoff and landing. Aerodynamic stall is a major loss of control event that results in fatal consequences [16]. Some of the recent accidents where pilots failed to recognize stall conditions include the crash of Continental Connection flight 3407 and Air France flight 447 (see [2, 3] for details). To address stall based safety concerns, the Federal Aviation Administration (FAA) issued rule changes to the Code of Federal Regulations (see [4] for details) mandating air carriers in the United States to train the pilots in recognizing, avoiding, and properly recovering from stalls by 2019. In this work, we aim to recognize and avoid stall risks early by discovering precursors that increase the risk of an aerodynamic stall and use the operational insights towards designing early warning systems and decision support tools for stall monitoring.

A stall may be described as follows [6]. The aircraft generates lift by forcing the air to flow around its wings. If the angle between the wing and the air flow (angle of attack) is increased aggressively, the airflow around the wing becomes turbulent and loses contact with the wing surface. The wings at this point is unable to generate sufficient lift and is said to be "stalled". A stall risk depends upon many variables including critical angle of attack, flight's airspeed, altitude rate, pitch rate, roll rate, aircraft weight, flap settings, etc., which are in turn controlled by the crew and/or the autopilot. Exogenous variables like weather also affect the risk of stall. Refer to Table 1 for the variables and their descriptions.

Although there exists some research on stall warning and recovery based on the aerodynamics and the physics of flying, such

DOI: 10.1145/3097983.3098097

models do not capture the risks due to human errors, environmental factors and errors in following operational procedures. Further, current systems are reactive (detect after occurrence) that gives less time for the crew to recover. Our goal is to develop proactive systems that can prevent the stall from occurring. Thus, we aim to detect the precursors to a stall instead of the stall itself so that the crew becomes aware of the evolving risk and take mitigation actions early in the process. We also take a data mining approach to include the system effects as well as the real-world effects arising from operational, human and environmental factors. Most airplanes are equipped with sensors to capture in high detail, the behavior of the flight subsystems, and actions of the pilot and automation. We aim to use this information to build data based decision support tools to complement the physics based models to provide real world insights on the causal factors of stall events. The insights could be used to inform pilot training programs to achieve better management of take off profiles and design early warning systems to alert the crew on potential loss of control.

A challenge in applying data mining techniques to detecting precursors to stall (and other extreme events in general) is that the number of positive examples (flights that stalled) is very low. A severe stall event is rare and if it is used directly as the target event, it would be difficult to achieve statistical significance using a small number of labeled data. However, it is fairly common for flights to approach a "stall risk" which is often mitigated. For instance, the risk of the airplane stalling during initial climb is high. Immediately after takeoff, the flight's airspeed typically increases as the flight gains altitude. During this time, if the airspeed drops intentionally or unknowingly, the safety margin to a stall is reduced and the flight goes into a state of high risk. If such "high-risk" states are considered as the target events, there could be two benefits - (1) substantial data becomes available for mining precursors, (2) stall risk can be detected earlier which increases the time available to the flight crew to make corrective actions much earlier than that provided by detecting the stall itself. Thus, in this work, we consider the the anomalous drop in airspeed (ADA) during initial climb which is a precursor to the stall and aim to find precursors to the ADA events using data mining.

We consider the flight operations quality assurance (FOQA) data from a commercial airline for this study. Mining FOQA data is challenging because of the following reasons. The data is high-dimensional, includes continuous and categorical variables and is of variable trajectory length. The variability within the data is high because (1) data from different aircraft weight classes that have significantly different flight characteristics, (2) data from different airports whose geography, weather, takeoff and landing patterns are significantly different, (3) different operational procedures that change slightly every few months which make comparing data from past history difficult. To keep the problem manageable, we restrict our scope to certain type of flights and applying feature selection based on Granger Causality along with domain knowledge. The reason for using Granger Causality is to present causally relevant variables to the precursor mining algorithm for it to detect actionable precursors to the ADA events. On the selected features, we apply a recently developed technique - the Automatic Discovery of Precursors in Time series (ADOPT) algorithm to discover precursors to the ADA events. ADOPT [12] has been developed at NASA Ames Research Center and has been shown to detect precursors to adverse events in FOQA type time series data. ADOPT works by constructing a model that separates the nominal times series data (no ADA event) from the adverse data (with ADA event) using reinforcement learning. The nominal data is also called as the "expert" data whose state transitions have evidence of the optimal behavior of human, automation and external factors combined, that mitigated the ADA risk. Precursors are then determined as state transitions that deviated from the expert's behavior.

The contributions of the paper can be summarized as follows:

(1) The application of data mining to recognize takeoff stall risk is novel.
(2) We demonstrate a methodology to mine for precursors using the FOQA data by scoping out manageable subsets of data, performing a causality based feature selection and applying ADOPT algorithm to find precursors.
(3) We improve the original ADOPT algorithm by simplifying the reward definition which reduces its computational complexity as well as gives a probability interpretation to the precursors.
(4) We conduct experiments based on real flight operational data to find operationally significant insights about drop in airspeed events.

2 RELATED WORK

While there exists no literature that analyzes precursors to the anomalous drop in airspeed (ADA) events, the work done in this paper fits in the research landscape on stall warning and loss of control. Current airplanes are equipped with industry standard stall warning systems that are mandated by regulations. They are available in the form of a stick-shaker that shakes the pilot's control arm to warn the onset of a stall, audible warning systems that provide an audible signal based on sensors that sense angle of attack or wind flow on the airplane's wings [1]. Such systems often warn the pilots when the aircraft stalls or when it is close to a stall. Other advanced systems have active stall protection controls that will add power to prevent a stall. However there may be situations where this protection is inactive. While such systems neither give insights to the precursors nor have a long lead time, they are the only choice when it comes to practical stall detection. Thus, much of the focus is on detecting stall better and developing recovery maneuvers. Most of the published work focuses on using physics or signal processing to improve stall detection [10] and develops guidance and control algorithms [16] to recover from a stall. Physics based modeling and flight test simulations are typically used to gain insights on stall approach and recovery [5, 8]. Current techniques can also predict the flight's trajectory and estimate the safe operating envelope[19] to help pilots with better situational and state awareness in flight.

All published literature on stall and loss of control are based on physics based modeling. While system based physics are useful for analysis and validation, they do not capture operational factors such as airport demands, maintenance issues, geography of the airport, communication, external factors such as weather, and human factors such as fatigue and attention. By building models directly using operational data, we can hope to capture some of these effects. Anomaly detection is a common way to study abnormal flight situations and identify precursor patterns from data [9]. For example, flight deck human-automation interaction issues could be detected using intent inference [14]. The ADOPT algorithm used in this work is different from anomaly detection because it starts with a definition of an adverse event and finds only those deviations that are correlated to the adverse event. On the other hand, anomaly detection algorithms find deviations from a nominal behavior that is defined more generally, making predictions of specific risks challenging. Time series classification algorithms [7] may be able to find the correlated features specific to an adverse event. However for precursor discovery, it is important to prune some of the correlations that occur later in time to reduce false alarms, which is the main novelty of ADOPT [12]. Precursors to high risk situations for flight landing and high energy approaches have been studied [12] using the ADOPT. However, the application to study drop in airspeed in flights is novel. Compared to rule mining [13] based methods where the number of candidate rules grow exponentially in high dimensional data, the proposed method handles the curse of dimensionality better by using kernel machines such as SVM.

3 DATA DESCRIPTION

For this study, we use the FOQA data provided by a de-identified commercial airline[1]. Figure 1 shows the airports in Europe where the de-identified airlines operate. The FOQA data consists of most of the sensory measurements on board the aircraft including flight speed, altitude, flight control surfaces, thrust, engine power, fuel consumption, pilot switches, pitch, roll, pressure, temperature among many others. The data is sampled at 1 Hz. Most of the data used in this work are recorded from flights that takeoff from these airports. To eliminate variability in aircraft characteristics, we consider only the airbus models A319 and A320 because they belong to the same weight category of interest in this work. We have not filtered the flights based on weather, wind patterns and airports to because any change in these variables will be compensated in flight so as to achieve the desired flight profile. Errors in such compensations may be responsible for the drop in airspeed and it is important to retain these features.

Although it is not possible to explain all 370 time series variables in the data, a subset of the variables that are relevant to this study are shown as distributions in Figure 2. It should be noted that the figure shows only distribution limits of the nominal data (the data without ADA events) to give the reader an understanding on how the variables nominally change during takeoff and initial climb (the first 100 seconds of flight after takeoff). It can be seen from Figure 2 that the airspeed is nominally around 150 knots[2] until about 1500 ft altitude after which the speed increases monotonically. In comparison with the nominal data, the flights that had the ADA events start around a similar nominal value but decreases as the aircraft climbs (See Figure 3). Coming back to Figure 2, the pitch is initially increased to gain altitude but after the flight reaches the acceleration altitude of 1500 ft, the pitch is reduced. This increases the airspeed at a higher rate because now the rate of altitude increase is lowered (altitude is traded-off for airspeed). It is also during this time, the autopilot and autothrottles are usually engaged. The reference primary flight display speed (PFD speed) is also seen to go up after reaching the acceleration altitude of 1500 ft.

[1] Owing to proprietary nature of the work, we do not disclose the airline name, data and code to the public. However, the approach reported in this paper is detailed enough to be applied to similar applications.

[2] 1 knot = 1.15078 mph

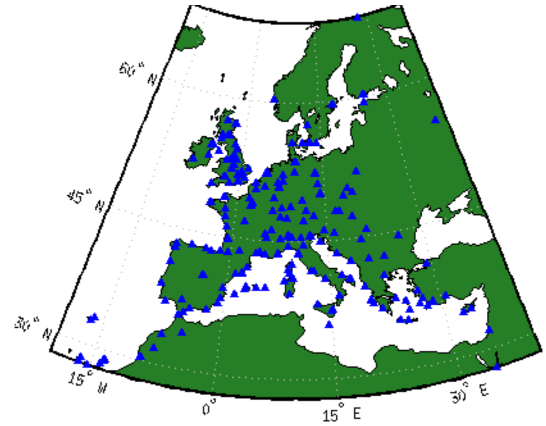

Figure 1: Airport locations in Europe where the de-identified airline company operates.

3.1 Adverse Event Definition

In this work we are interested in the anomalous drop in airspeed (ADA) immediately after lift-off[3] with a severity of at least 20 knots in speed reduction. This severity level detects adverse airspeed drop events that are operationally significant as well as gives sufficient data hits to train our models. The rationale behind defining the adverse event purely based on a single variable - the airspeed, is as follows. The airspeed is an important part of the energy state of the aircraft and is managed both strategically as well as tactically. The flight crew determines the desired takeoff profile when the aircraft is on the ground prior to takeoff. The desired profile is a calculation that takes into account the current wind conditions, the gross weight of the airplane, and using the performance characteristics of the airplane. Once the desired profile is calculated, the aircraft is flown using the reference speeds that are calculated based on this profile both by the human pilot and the autopilot. The dynamic variations in wind, flight behavior, operating procedures and air-traffic controller commands are tactically compensated to achieve the optimal flight profile. Any error in strategic calculations or in tactical control and external factors will gain visibility via the airspeed. Thus, we define the adverse event based on a threshold on airspeed and look back in the flight's history to determine if any erroneous action/external disturbance occurred and how it propagated. While the definition of the adverse event is based on one variable, the precursor analysis considers multiple time series variables in combination.

4 PRECURSOR DISCOVERY USING ADOPT

4.1 Precursor Mining Problem

Consider a multivariate time series with d variables as shown in Figure 4. If the adverse event E_a occurs at time $L + 1$, the data

[3] lift-off is a part of takeoff when the wheels lose contact with the ground.

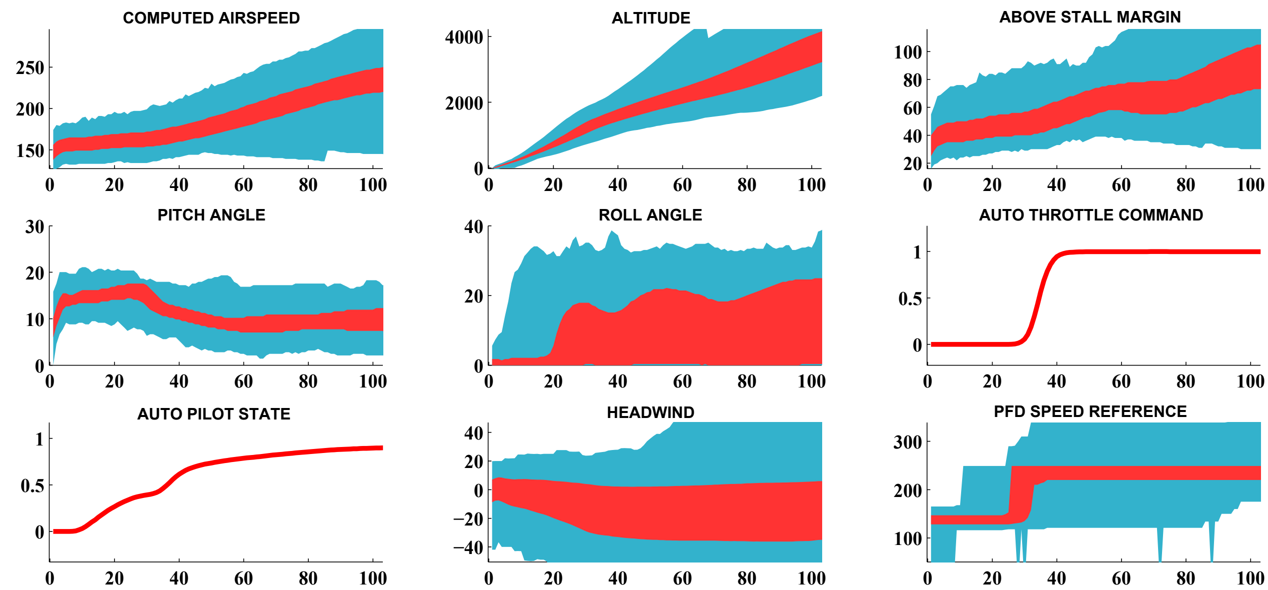

Figure 2: Figure showing the nominal distributions of selected time series variables. The outer distribution in blue captures the 10 - 90 percentile while the inner red patch shows the 25-75 percentile range. For the binary variables, the proportion of the nominal data being ON (equal to 1) is shown at every time step. The x-axis for each subplot is time in seconds after the flight lifts off from the ground.

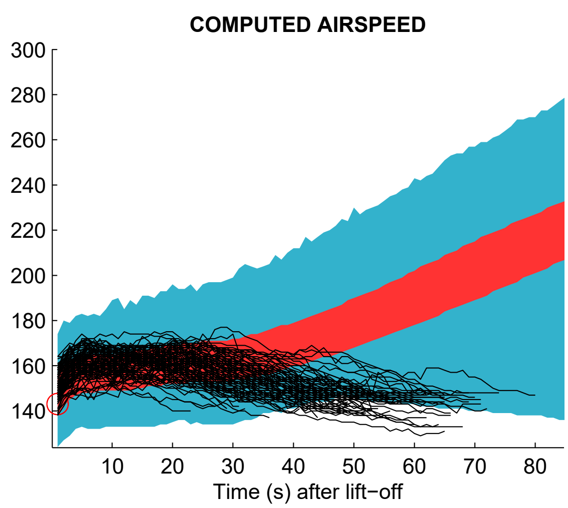

Figure 3: Figure showing nominal distribution of airspeed with some adverse flights overlaid (shown in black). The adverse flights are shown until the occurrence of ADA event. Refer to the caption from Figure 2 for interpretation.

prior to the adverse event (corresponding to time $1, 2, .., L$) can be considered the search data where precursors may be present. Let $\overline{\mathcal{N}} = \{X_i\}_{i=1}^{\overline{N}}$ be a database containing $\overline{N}$ such time series records. Similarly let $\mathcal{N} = \{X_i\}_{i=1}^{N}$ be a database containing N time series records that are nominal; i.e., data where E_a does not occur. A data record X_i can be represented in matrix form as follows

$$X_i = \begin{bmatrix} x^1(1) & x^1(2) & . & x^1(L_i) \\ x^2(1) & x^2(2) & . & x^2(L_i) \\ . & . & . & . \\ x^d(1) & x^d(2) & . & x^d(L_i) \end{bmatrix}$$

where L_i is the length of the multivariate time series X_i. Note that each X_i may have a different length. Let the event at time k be defined as

$$\mathbf{x(k)} = [x_1(k), x_2(k), .., x_d(k)]\,. \tag{1}$$

The time series record X_i can now be represented in terms of events as $X_i = [\mathbf{x(1)}, \mathbf{x(2)}, .., \mathbf{x(L_i)}]$. With this setup, ADOPT defines a precursor as follows

Definition 4.1. Given a sequence of events $X = [\mathbf{x(1)}, \mathbf{x(2)}, .., \mathbf{x(L)}]$, an action is any state transition $a_k : \mathbf{x(k)} \longrightarrow \mathbf{x(k+1)}$ where $1 \le k \le L$, then a_k is a precursor to E_a if

$$V(a_k) - V(a_k^*) > \delta, \tag{2}$$

where a_k^* is the expert action at time k and $V(.)$ is the expert's value function [12].

A value function in this context, is a metric that evaluates a given event (a_k in this case) for its long term consequence with respect to the adverse event. It will be shown in Section 4.2.2 that the value function is equivalent to the conditional probability $P(E_a|a_k)$; i.e., a high value of a_k translates to a high probability of adverse event occurring in the future. Thus, the expert's value $V(a^*)$ is always less than or equal to $V(a)$. We assume that the nominal data is generated by an expert who manages the state transitions so that the adverse event is prevented. Consequently, the expert's action

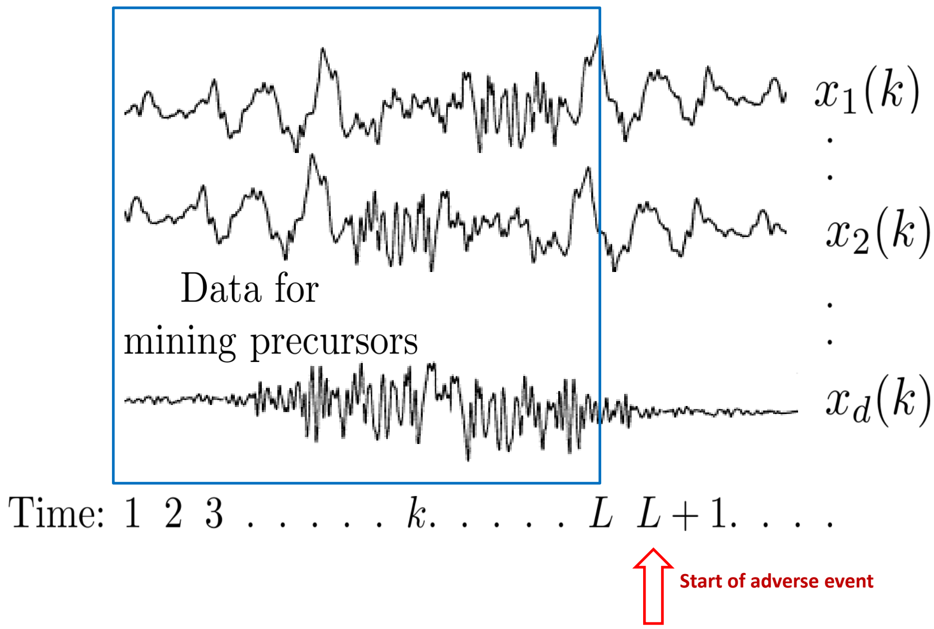

Figure 4: Schematic showing the data setup for precursor discovery in ADOPT. Each data record is a time series of d dimensions.

a_k^* is the best action that can be taken at time k to prevent the adverse event.

4.2 ADOPT Algorithm

ADOPT (abbr. for Automatic Discovery of Precursors in Time series data) is a recently developed algorithm [12] that detects precursors based on definition 4.1. The various steps involved in ADOPT [12] include data preparation, modeling the expert's reward function, modeling the expert's value function, identifying expert's actions and evaluating precursor events. ADOPT approaches the precursor discovery problem from a decision making perspective. The time series data is considered as a Markov decision process where a hidden agent makes a decision at every time instant and the quality of decisions decide the outcome of the adverse event; i.e., if the agent makes optimal decisions, the adverse event is avoided whereas if the agent makes poor decisions, the adverse event occurs as a consequence. In this work, the agent is a unified entity that includes the actions of flight crew, external air-traffic controller, flight automation and external disturbance such as weather. ADOPT infers the level of optimality in decision making using historical data and ranks the poor decisions as candidate precursors.

4.2.1 Background on MDP. An MDP is a tuple ($\mathcal{X}$, $\mathcal{A}$, $P^a_{x,x'}$, γ, R) where $\mathcal{X} \subseteq \mathbb{R}^d$ is a continuous state space with d state variables, $\mathcal{A}$ is an action space (could be a combination of continuous or categorical), $P^a_{\mathbf{x},\mathbf{x}'}$ (or $P_{\mathbf{x}\mathbf{x}'}$ if actions are unknown) are the state transition probabilities corresponding to an action a at state $\mathbf{x}(k) = \mathbf{x}$ transitioning to $\mathbf{x}(k+1) = \mathbf{x}'$, $\gamma \in [0, 1)$ represents the discount factor and $R : \mathcal{X} \to \mathbb{R}$ is the underlying reward function. A policy π can be defined as any map $\pi : \mathcal{X} \mapsto \mathcal{A}$ specifying an action a at every state $\mathbf{x}$. The value function of a policy π evaluated at a state $\mathbf{x_0}$ is be given by

$$V^{\pi}(\mathbf{x_0}) = E[R(\mathbf{x_0}) + \gamma R(\mathbf{x_1}) + .. + \gamma^L R(\mathbf{x_L})|\pi] \quad (3)$$

where the expectation is over the distribution of state sequences starting from $\mathbf{x_0}$ and following the policy π [15, 17]. The expert's policy π_E is given by

$$\pi_E(\mathbf{x}) \geq \pi_i(\mathbf{x}) \iff V^{\pi_E}(\mathbf{x}) \geq V^{\pi_i}(\mathbf{x}) \quad \forall \mathbf{x} \quad \forall \pi_i. \quad (4)$$

Using the expert's value function V^{π_E}, the optimal action can be determined using Bellman's optimality as

$$\pi_E(\mathbf{x_k}) = arg \max_{a_k} Q^{\pi_E}(x_k, a_k), \quad (5)$$

where $Q^{\pi_E}(x, a)$ is the state-action value function. If one has no access to the actions, the state transitions may be abstracted as actions and Bellman's optimality can be applied to get the optimal state transition [12] as follows

$$\pi_E(\mathbf{x_k}) = arg \max_{\{feasible\ \mathbf{x_{k+1}}\}} V^{\pi_E}(\mathbf{x_{k+1}}) \quad (6)$$

4.2.2 Probability Interpretation of the Value. The original algorithm presented in [12] are based on modeling the reward function first and then using it to model the value function. In such a setup, the value function indirectly gives an estimate of the probability of the adverse event and in fact it is inversely proportional to the probability. For example, a high value indicates a low probability of the adverse event and so on. A direct interpretability is missing in the original formulation of ADOPT. To overcome this, we define the reward in a simpler way as follows. If a time series record is given by $X = [\mathbf{x}(1), \mathbf{x}(2), .., \mathbf{x}(\mathrm{L})]$, the reward may be defined based on the knowledge of the adverse event in the data as

$$R(\mathbf{x}) = \begin{cases} 0, & \text{if} \quad \mathbf{x} = \mathbf{x}(\mathbf{L}) \quad \text{and} \quad X \in \mathcal{N} \\ 1, & \text{if} \quad \mathbf{x} = \mathbf{x}(\mathbf{L}) \quad \text{and} \quad X \in \overline{\mathcal{N}} \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

If the reward function is defined in this way, it can be proved [18] that $V(\mathbf{x}) = P(E_a|\mathbf{x})$; i.e., the value function of a state $\mathbf{x}$ directly gives the conditional probability of the adverse event given the occurrence of that state. In ADOPT, the precursor definition from equation (2) can now be expressed as

$$P(E_a|a_k) - P(E_a|a_k^*) > \delta, \tag{8}$$

This means that ADOPT's precursors can be interpreted as follows. A precursor is an action that has a high probability of the adverse event compared to the equivalent expert's action at that time instant. The new definition of reward function (equation (7)) not only gives a better interpretation (in terms of probability) for the precursors but also eliminates the reward estimation step which reduces the overall computational complexity of ADOPT.

4.2.3 Estimating Expert's Value Model. In ADOPT, a model of the expert's value function is estimated using the knowledge of the reward function and the time series data. The time series data may be considered as Monte Carlo samples and the value function can be estimated using the Monte Carlo (MC) reinforcement learning method [17]. This fits well in our setup as the MC method requires only the policy demonstrations (time series data in our case) without any prior knowledge of the environment's dynamics [17] such as a full transition probability matrix which may be very hard to estimate for the FOQA data. As shown in [12], for every trajectory, the first visit Monte Carlo return $V(s)$ for each state s is determined using the reward model $R(s)$. In contrast to the previous version of ADOPT [12], the above modified definition of a reward makes the $V(s_i)$ to be binary. Thus, instead of a regression model as originally proposed in [12], here we use a classification model such as SVM to obtain a model $(\hat{V}^{\pi_E}(\mathbf{x};\theta^*))$ of the expert's value function.

4.2.4 Algorithm. The pseudocode of the modified ADOPT algorithm is shown in Algorithm 1. The algorithm begins by taking data corresponding to nominal $\mathcal{N}$ and adverse $\overline{\mathcal{N}}$ time series as trajectory demonstrations of the expert and non-expert respectively. The expert's reward model $R(\mathbf{x})$ is defined as in equation (7). Using the reward model and the trajectory data, the expert's value model $\hat{V}^{\pi_E}(\mathbf{x})$ is estimated using reinforcement learning. A detailed description of these steps is given in [12]. The value model along with a test trajectory X_T is given as input to the precursor discovery module in Step 2. In this work, the action space $\mathcal{A}$ is unknown. So actions are abstracted from the state transitions. For each action a_k in X_T, the value $V^{\pi_E}(a_k)$ is determined using the expert's value model. Using the data history, the set of feasible states $\mathcal{R}(\mathbf{x_k}) = \{feasible\ \mathbf{x_{k+1}}|\mathbf{x_k}\}$ is determined. This represents the set of all possible states the agent can reach by taking any possible action at time k including a_k and the optimal action a_k^*. The values corresponding to all such actions $V^{\pi_E}(\mathcal{R}(\mathbf{x_k}))$ using the expert's value model is determined. Using Bellman's optimality, the action corresponding to the minimum value $V^{\pi_E}(a_k^*)$ is identified. Note that the expert's value model gives the probability of the adverse event and so the expert tries to minimize it. A precursor score PS is determined as the difference $V^{\pi_E}(a_k) - V^{\pi_E}(a_k^*)$. By learning a threshold model δ_P from data (described in Section 5.3), the precursors are detected.

Algorithm 1: ADOPT Pseudocode

```
Data: Time series data from nominal database 𝒩 and adverse
      database 𝒩̄.
Result: Precursors in a given time series X_a.
Step 1. Estimation of Expert's Value function
Input  : Expert's demonstration X_i, i = 1, 2, ..., N + N̄,
         Reward R(x).
S ← [];
V ← [];
for i ← 1 to N + N̄ do
    for k ← 1 to L_i do
        S ← [S,   x_k];
        V ← [V,   {Σ_{p=k}^{L_i} γ^{p-1} R(x_p)}];
    end
end
θ* ← arg min_θ (1/|V|) Σ_{i=1}^{|V|} ||V_i − V̂^{π_E}(S_i; θ)||² + (μ/2)||θ||²;
Output: Expert's value model V̂^{π_E}(x) = f_V(x; θ*).
Step 2. Discovery of Precursors
Input  : Expert's value model V̂^{π_E}(x), test trajectory X_T.
for k ← 1 to L_T do
    V^{π_E}(a_k) ← V̂^{π_E}(x_k → x_{k+1});
    ℛ(x_k) ← {feasible x_{k+1}|x_k} calculated from data;
    V^{π_E}(a_k^*) ← min_{x∈ℛ(x_k)} {V^{π_E}(x)} (Bellman's optimality);
    PS_k ← V^{π_E}(a_k) − V^{π_E}(a_k^*);
end
𝒫_{E_a} ← {PS_k | PS_k > δ_P};
Output: PS_k and 𝒫_{E_a}.
```

5 EXPERIMENTS

We consider the FOQA data from a de-identified airlines company and obtained two sets of flights - the ones that had ADA events as defined in Section 3.1 and the ones that did not have the ADA events (the nominal). Let these two datasets be identified as $\overline{\mathcal{N}}$ and $\mathcal{N}$ respectively. The FOQA data has more than 370 time series variables in each flight record and it is not trivial to decide which ones to choose for precursor analysis. We do feature selection based on Granger Causality, to get a candidate set of sensory variables from which we use domain knowledge to get a smaller and refined set for further mining using ADOPT.

5.1 Feature Selection

One of the goals of this work is to recommend actionable insights to the flight crew and air-traffic controllers. Thus, we begin by selecting those time series variables that have a strong causal relationship with respect to the flight's airspeed so that the detected precursors may give insights into the true causal factors. Granger Causality [11] is a simple and scalable technique to infer causality in time series data. Granger Causality is based on statistical hypothesis testing which tests if there is a significant improvement in using a given variable to predict a target variable. The target variable in our case is the flight's airspeed. We check all the other time series variables to see if they causally influence the airspeed. The causal strength is determined using the magnitude of the F-score obtained in Granger causality analysis.

A maximum lag of 10 was used for the analysis. About 150 variables out of the 370 were eliminated using Granger causality method. From the rest, the top 10% of the variables that causally effects airspeed were short listed. It has to be noted that Granger causality method in its basic form only finds linear causality and considers variables independently. Thus, we verified the short-listed variables with a team of domain experts who eliminated some variables that were unrelated to the ADA events, and some that were consequential to the ADA events. For example, the flight's flap setting appeared as a causal variable but was removed because flaps are set based on the airspeed during initial climb and may not be a precursor to variations in airspeed. Other irrelevant variables such as 'database validity', 'Traffic Collision Avoidance System sensitivity' etc. that had no effect on stall were also removed from the list. There were some relevant variables that the Granger analysis missed which were added manually to the list. For example, important variables that define the aircraft state such as altitude, wind speed, were not listed in the top 10% of the causal variables. In addition, some hand-engineered features were added including adding a binary variable for PFD selected speed switching to a value of 1 when the speed is increased, binary variables to represent the current altitude mode of the flight, a binary variable that represents turning ON and OFF either autopilot. The list of variables that we considered for ADOPT analysis is shown in Table 1.

5.2 ADOPT Modeling

Before training the models for ADOPT, we normalize the data to have zero mean and unit standard deviation across all flights. The data is split into training and testing sets with testing data constituting about 10% of the data. The training set had about 36000 flights while the test data has about 4000 flights. The class proportions in the data are balanced with equal number of nominal and adverse flights by randomly sub-sampling the nominal flights which were in excess. The nominal flights were cutoff at 109 seconds (the length of the longest adverse time series) after lift-off, while the adverse flights had variable lengths starting from lift-off to when it loses about 20 knots of airspeed. The adverse trajectories vary in length between 10 and 109 and nominally around 50. The SVM model for ADOPT's value function was learned with Gaussian kernel parameters $\gamma = 10, C = 50$, achieving an accuracy of 87% on the unseen data. The values of γ and C are selected based on cross-validation.

5.3 Finding Optimal Threshold

In the previous version of ADOPT [12], the threshold δ_P on the precursor score was set based on the 90^{th} percentile value of a given flight, so that the top precursors in that flight are reported. A downside of this approach is that it does not generalize the predictive ability of the detected precursors to other flights. In this paper, we have improved this aspect by learning the threshold directly from the two classes of data for better generalization. After estimating the expert's value model, the precursor scores (PS) for all trajectories in the nominal and adverse data are calculated. Using the precursor scores as features and the adverse event knowledge (nominal or adverse) as labels, we learn a model that separates the precursor scores of the two classes of data. As this data is one dimensional, a simple threshold learning was used and the optimal threshold function is determined. The learned threshold is shown in Figure 5, where the threshold is shown in dotted black curve, the precursor scores of the adverse and nominal flights are shown in red and green respectively. The learned threshold is tested using the unseen test data which is also shown in the figure.

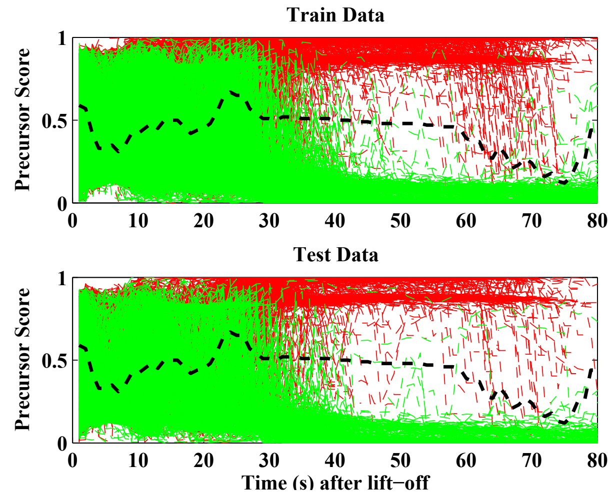

Figure 5: Precursor threshold learned from data as a function of time (s) after takeoff.

6 RESULTS

6.1 ADOPT Results

It can be seen from Figure 5 that ADOPT's precursor scores give an insight into how the risk of the drop in airspeed (ADA) is developing in the flight. For the nominal flights (colored green), although the precursor scores were high until about 30 seconds after takeoff, it goes down soon after possibly because of taking optimal actions which prevented the airspeed from dropping. On the other hand, the precursor scores of the adverse flights, remained mostly high (above the learned threshold) because of sub-optimal decision being made. ADOPT's underlying model discriminates the two classes of flights reasonably well with a overall accuracy of 87%.

6.2 Predictive Skill

Although it is not possible to quantitatively show ADOPT's performance in detecting precursors because precursor labels are unknown[4], we can use the precursor score from ADOPT as features to demonstrate its predictive power towards predicting the ADA

[4]Analyzing few minutes of a flight takes in the order of hours of analyst time mainly because of the high dimensional nature of the data. Thus it is not feasible to get sufficient precursor labels for a quantitative evaluation.

Table 1: Table showing the list of time series variables selected using Granger causality and using domain knowledge

Variable Name	Description	Type
Computed airspeed	Speed of aircraft relative to the surrounding air	Continuous
Auto throttle command	Command that allows a pilot to control the power setting of the aircraft's engines	Binary
Above stall margin	The speed margin that is available to reach an aerodynamic stall	Continuous
Mach number	Mach number	Continuous
Pitch angle	Angle between the longitudinal axis of the aircraft and the horizontal axis	Continuous
Angle of attack	Angle between the longitudinal axis of the aircraft and the motion trajectory of the plane	Continuous
Flight display selected speed	Reference speed (automation calculated) based on the aircraft's speed profile, wind, weight, flap settings.	Continuous
Immediate climb mode	Flight mode that commands an immediate climb	Binary
Pitch rate	Rate of change of pitch angle	Continuous
Longitudinal acceleration	Acceleration in the direction of the flight's longitudinal axis.	Continuous
Ground speed	Speed of aircraft relative to the ground	Continuous
Expedited climb mode	Flight mode that commands an expedited climb	Binary
Vertical speed mode	Maintains a specified vertical speed	Binary
Flap setting	Flap levels of the aircraft	Categorical
Altitude	Altitude of the aircraft with respect to mean sea level	Continuous
Roll angle	Angle of rotation around the longitudinal axis of the aircraft	Continuous
Headwind	Wind speed along the longitudinal axis of the flight (-ve is headwind, + is tailwind)	Continuous
Autopilot state	State of the autopilot. The value is 1 if ON.	Binary
PFD switch command	The value becomes 1 if the primary flight display reference speed is increased. The value remains 0 if unaltered.	Binary
Altitude mode M	Value is 1 when the flight is in the M^{th} mode of takeoff, 0 elsewhere.	Binary

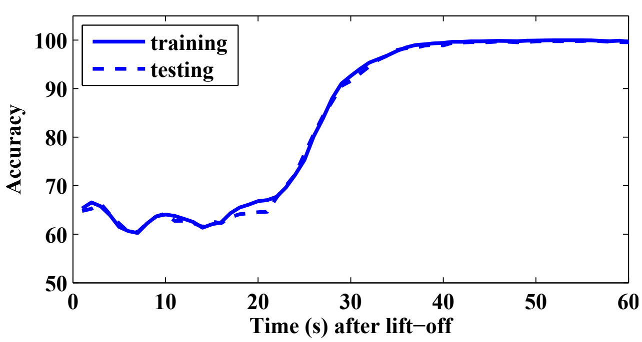

Figure 6: Time-wise accuracy of classifying a flight as nominal vs adverse using ADOPT's precursor score as inputs.

events. We can use the learned threshold from ADOPT at each time step to classify if the flight is nominal or adverse, and quantitatively show performance. For instance, in Figure 5, it can be seen that after 40 seconds into the climb, the threshold can separate the two classes of data well. This means, the precursor score at 40 seconds has a high predictive capability. The corresponding accuracy of prediction after 40 seconds into the flight is close to 100% (see Figure 6). As we look back to the past (the first 30 seconds of the flight), the data is not easy to separate, which is evident from the accuracy of classification (about 60%, see Figure 6). For the other time steps, refer to Figure 6 for the model's performance.

6.3 Flight Analysis

This section gives an example[5] of an adverse flight being analyzed using ADOPT. The flight takes off during normal winds and does not have a large roll during initial climb. However, around 30 seconds (see Figure 7), where most nominal flights had their PFD speed increased above 200 knots, the test flight's PFD was less than 150 knots. The PFD speed is a reference based on which the pitch commands are derived, which cues the human pilot (the autopilot is OFF in this case). ADOPT also ranks the top variables in the time series that contributed to the precursor score. This is important to understand and interpret the results of the model, particularly when the data is high dimensional. ADOPT breaks down the influence of each variable independently and ranks the top precursors at a given time instant. For the flight under study, the top precursors at various points during the flight are shown in Figure 8. It can be seen that at time T=27s, the PFD speed switch command was listed as the top precursor. At other time steps such as T=5s and T=42s, the pitch angle seems to be the dominating precursor and from the distribution plot in Figure 7, it is clear why this is the case. Although at T=1, there is no significant headwind (tailwind =5 knots), having a tailwind is sensitive to increasing the probability of the airspeed drop. This may be the reason the headwind is listed as a top precursor at T=1. The precursors discovered for this flight and for others were qualitatively validated by domain experts who did an independent analysis of the flights for precursors.

[5] We only show one example due to space constraints. The overall performance of ADOPT may be seen from the results in Figure 6 that is based on data from about 40000 flights.

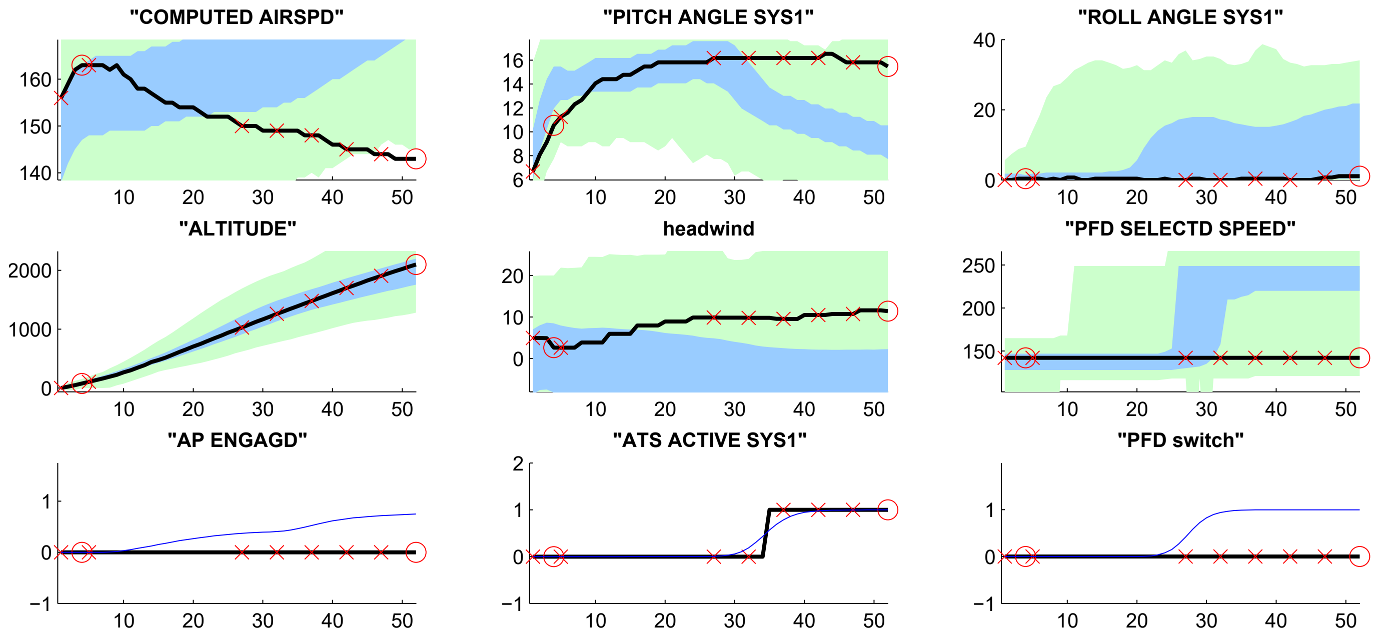

Figure 7: Flight time series variables plotted (thick black curve) against time (seconds) after lift-off. The distribution of nominal values are shown in blue (25-75 percentile) and green (10-90 percentile) as references. The X marker in red shows the precursors identified by ADOPT while the two O markers in red show the start of the speed drop and where the drop reaches 20 knots.

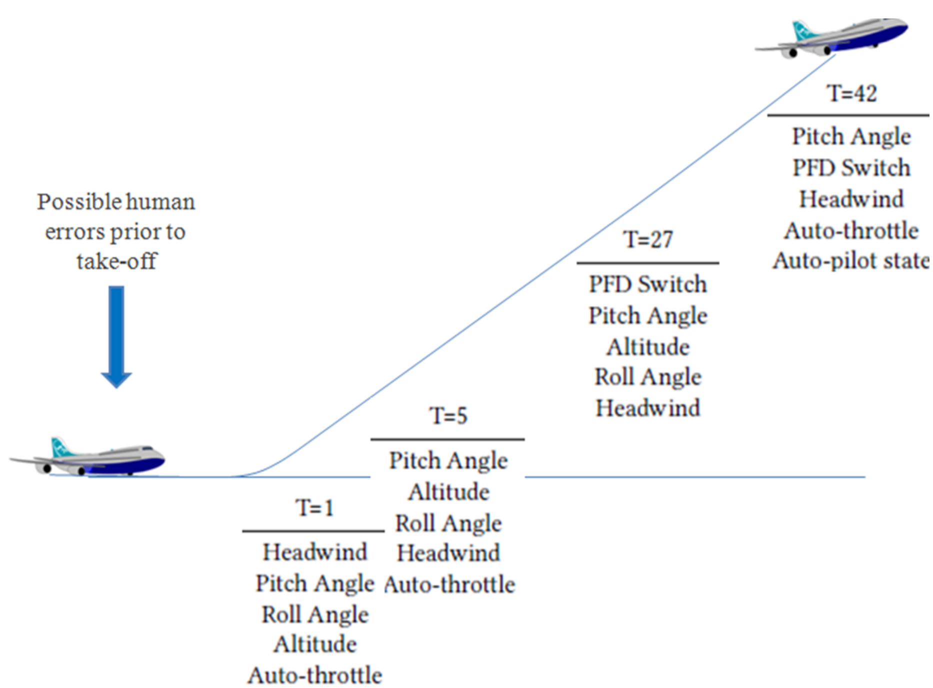

Figure 8: Precursors identified by ADOPT at different points in the flight.

6.4 Discussion

A major challenge in this work is that stall events are rare and it is not possible to obtain sufficient data to mine. We addressed this problem by considering a precursor to stall which is the anomalous drop in airspeed (ADA) events for which we were able to get sufficient data examples. Another challenge is that the precursor discovery problem is unsupervised. We only have labels for the consequence of the precursors. We used ADOPT that was specifically designed to solve this problem. ADOPT identified precursor events that are early predictors of the adverse event. It is impossible to obtain ground truth on precursors because (1) it is tedious for high dimensional data, (2) there could be precursors that are unknown to domain experts and so any ground truth labeling is prone to errors. Thus the only way to validate ADOPT's precursors is to evaluate their ability to discriminate the nominal and the adverse time series, which was performed under the section 6.2. Further, such precursors may be present in nominal flights as well and ADOPT was able to detect them along with showing a complete picture of how the probability becomes low as better actions were taken. Such actions may be used to learn about aircraft safety resilience starting from a high risk state.

Feature selection turned out to be extremely important in this work. It not only reduced the dimension and simplified model complexity, but also helped identify meaningful precursors. For example, it is well known that angle of attack is a major indicator to a stall. However, detecting a precursor in angle of attack may not provide a long lead time. So we removed some of the immediate indicators of stall such as angle of attack, airspeed, stall margin so that ADOPT tries to discriminate the data purely based on control commands from the pilot and autopilot. From our experience, this reduced the noisy detection as well as found meaningful precursors. ADOPT's precursors are identified as high dimensional vectors which sometimes makes it difficult to interpret. For example, if we considered 50 variables, a precursor detected by ADOPT would be a 50 dimensional vector. For a better interpretation, the sensitivity of each variable in the vector is analyzed and a ranking algorithm is used to rank the top-k variables (as shown in Figure 8). Although some of the combination effects are lost, this helped domain experts interpret the precursors easily.

The ADOPT algorithm was applied to find precursors to the ADA events and guide the discussions with subject matter experts on narrowing down to the operational factors behind the events. For some of the flights we analyzed, we were able to trace back the root causes of drop in airspeed to errors that possibly happened when the flight was on the ground prior to takeoff including error in setting the PFD speed reference, error in performance calculations prior to takeoff, overcompensation because of an error in taking-off at high headwinds. Such insights are important for improving pilot training programs, designing better automation systems and operating procedures.

7 CONCLUSION

ADOPT algorithm was applied to to find precursors to the drop in airspeed (ADA) events during takeoff. The reward function in ADOPT was simplified resulting in better interpretability and simpler modeling. A Granger causality based feature selection was performed prior to ADOPT that helped discover actionable and operationally significant precursors. ADOPT's precursors achieved about 87% accuracy towards its skill in predicting the drop in airspeed events. Future work will focus on developing decision support tools to assist pilots and controllers to manage the risk factors during a flight's takeoff.

8 ACKNOWLEDGMENTS

This research is supported by the NASA Airspace Operation and Safety Program. The authors would also like to thank the SMEs: Michael S. Feary, Mike Stewart and John Kaneshige for their insightful comments and perspective on the identified precursors.

REFERENCES

[1] 2016. Stall warning systems, SKYbrary. (2016). http://www.skybrary.aero/index.php/Stall_Warning_Systems
[2] February 2009. National Transportation Safety Board Accident Report. Loss of Control on Approach Colgan Air, Inc. Operating as Continental Connection Flight 3407 Bombardier DHC-8-400. In *Report Number NTSB/AAR-10/01, PB2010-910401*. N200WQ Clarence Center, NY, USA.
[3] July 2012. The Bureau of d'Enquetes et d'Analyses (BEA) pour la securite de l'avation civile. Final Report: On the accident on 1 June 2009 to the Airbus A330-203 registered F-GZCP operated by Air France flight 447 Rio de Janeiro-Paris.
[4] November 12 2013. Federal Aviation Administration. Qualification, Service, and Use of Crewmembers and Aircraft Dispatchers; Final Rule, 14 CFR Part 121. Federal Register, http://www.gpo.gov/fdsys/pkg/FR-2013-11-12/pdf/2013-26845.pdf. 78 (218).
[5] Gavin K Ananda and Michael S Selig. 2016. Stall/Post-Stall Modeling of the Longitudinal Characteristics of a General Aviation Aircraft. In *AIAA Atmospheric Flight Mechanics Conference*. 3541.
[6] J.D. Anderson. 2005. *Introduction to Flight*. McGraw-Hill Higher Education.
[7] Anthony Bagnall, Jason Lines, Aaron Bostrom, James Large, and Eamonn Keogh. The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances. *Data Mining and Knowledge Discovery* (????), 1–55.
[8] Kevin Cunningham, John V. Foster, Gautam H. Shah, Eric C. Stewart, Robert A. Rivers, James E. Wilborn, and William Gato. 2004. Simulation Study of a Commercial Transport Airplane During Stall and Post-Stall Flight, In SAE Technical Paper.
[9] Santanu Das, Bryan L Matthews, Ashok N Srivastava, and Nikunj C Oza. 2010. Multiple kernel learning for heterogeneous anomaly detection: algorithm and aviation safety case study. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 47–56.
[10] Thomas S. Donnelly. 1972. An Improved Stall Warning System for General Aviation Aircraft, In SAE Technical Paper.
[11] Clive WJ Granger. 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society* (1969), 424–438.
[12] Vijay Manikandan Janakiraman, Bryan Matthews, and Nikunj Oza. 2016. Discovery of Precursors to Adverse Events using Time Series Data. In *Proceedings of the 2016 SIAM International Conference on Data Mining*. Miami, FL, USA.
[13] Srivatsan Laxman and P.S. Sastry. 2006. A survey of temporal data mining. *Sadhana* 31, 2 (2006), 173–198. https://doi.org/10.1007/BF02719780
[14] Sangjin Lee, Inseok Hwang, and Kenneth Leiden. 2014. Flight deck human-automation issue detection via intent inference. In *Proceedings of the International Conference on Human-Computer Interaction in Aerospace*. ACM, 24.
[15] Andrew Y. Ng and Stuart Russell. 2000. Algorithms for Inverse Reinforcement Learning. In *in Proc. 17th International Conf. on Machine Learning*. Morgan Kaufmann, 663–670.
[16] Vahram Stepanyan, Kalmanje S Krishnakumar, John Kaneshige, and Diana M Acosta. 2016. Stall Recovery Guidance Algorithms Based on Constrained Control Approaches. In *AIAA Guidance, Navigation, and Control Conference*. 0878.
[17] R.S. Sutton and A.G. Barto. 1998. *Reinforcement Learning: An Introduction*. MIT Press.
[18] Xin Xu. 2010. Sequential anomaly detection based on temporal-difference learning: Principles, models and case studies. *Applied Soft Computing* 10, 3 (2010), 859 – 867.
[19] Steven D Young, Maarten Uijt De Haag, Taumi Daniels, Emory Evans, Kimberlee H Shish, Stefan Schuet, Timothy Etherington, and Daniel Kiggins. 2016. Evaluating Technologies for Improved Airplane State Awareness and Prediction. In *AIAA Infotech@ Aerospace*. 2043.