ANALYSIS OF GENERAL AVIATION FIXED-WING AIRCRAFT ACCIDENTS INVOLVING INFLIGHT LOSS OF CONTROL USING A STATE-BASED APPROACH

Neelakshi MAJUMDAR [1], Karen MARAIS [1,*], Arjun RAO [2]

¹School of Aeronautics and Astronautics, Purdue University, West Lafayette, USA
²Collins Aerospace, 400 Collins RD, MS 124-319 Cedar Rapids, USA

Abstract. Inflight loss of control (LOC-I) is a significant cause of General Aviation (GA) fixed-wing aircraft accidents. The United States National Transportation Safety Board's database provides a rich source of accident data, but conventional analyses of the database yield limited insights to LOC-I. We investigate the causes of 5,726 LOC-I fixed-wing GA aircraft accidents in the United States in 1999–2008 and 2009–2017 using a state-based modeling approach. The multi-year analysis helps discern changes in causation trends over the last two decades. Our analysis highlights LOC-I causes such as pilot actions and mechanical issues that were not discernible in previous research efforts. The logic rules in the state-based approach help infer missing information from the National Transportation Safety Board (NTSB) accident reports. We inferred that 4.84% (1999–2008) and 7.46% (2009–2017) of LOC-I accidents involved a preflight hazardous aircraft condition. We also inferred that 20.11% (1999–2008) and 19.59% (2009–2017) of LOC-I accidents happened because the aircraft hit an object or terrain. By removing redundant coding and identifying when codes are missing, the state-based approach potentially provides a more consistent way of coding accidents compared to the current coding system.

Keywords: General Aviation safety, General Aviation accidents, loss of control, accident modeling, NTSB database.

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.3846/aviation.2021.15837

Introduction

Fixed-wing General Aviation (GA) accidents comprise approximately 64% of all aviation accidents in the United States (U.S.) every year (NTSB, 2019a). Most fixed-wing GA accidents result from inflight loss of control (LOC-I), controlled flight into terrain (CFIT), continued visual flight rules flight into instrumental meteorological conditions ('continued VFR into IMC'), engine failures, and fuel exhaustion/contamination (cf. AOPA, 2018; GAJSC, 2016). In particular, inflight loss of control (LOC-I) continues to be a significant cause of GA fixed-wing aircraft accidents each year. Loss of control is "a hazardous condition that involves an unintended departure of an aircraft from controlled flight regime" (FAA, 2019). Nearly 50% of fixed-wing GA accidents in the last two decades in the United States (U.S.) are attributed to LOC-I (NTSB, 2019a). In 2017, 21% of fixed-wing GA accidents overall involved LOC-I; for fatal accidents, this percentage increases to 57%.

There is a clear need to better understand the reasons for LOC-I accidents. One approach to improving our understanding is by analyzing historical accident reports. In the U.S., the National Transportation Safety Board [NTSB] investigates all civil aviation accidents. After concluding their investigation, the NTSB publishes a final report, which includes a prose section with summary analysis of the accident, a discussion of the probable cause and findings, and "factual information" on the flight history, personnel, aircraft, meteorological conditions, medical and pathological information, and any tests and research the investigators conducted (NTSB, 2019b). Each accident is also coded using a set of codes for occurrences, findings, and phases of flight to facilitate trend analysis (NTSB, 1998). Rao et al. (2016) provides a detailed discussion of this system. The NTSB coding system is based on an event-based model, where one event leads to another, but not all aspects of accidents are events. For example, an impaired pilot is better understood as a continuing condition, or a state (Rao & Marais, 2020). The pilot's impaired

condition makes subsequent errors more likely, and therefore does not fit well as "only" an initiating event. Additionally, multiple codes in the NTSB database have similar meanings. For example, the subject codes 24518: Altitude and 24519: Proper altitude both indicate that the pilot did not maintain the correct altitude. Such redundancy in codes can lead to inaccurate counts in accident causes. Finally, the NTSB database does not present all findings as codes, for example, in the pre-2008 coding system, there are no codes to capture improper aircraft heading.

Unfortunately, the prose content for GA accidents tends to be short. In 2017 (the most recent year to have completed factual reports), the average length for the 105 accidents that had LOC-I codes was just 449 words. The occurrence chains (number of occurrence codes) for these 105 accidents are also short (mean chain length = 3.36, SD = 1.42), albeit somewhat longer than that for all GA fixed-wing aircraft accidents (mean chain length = 2.48, SD = 1.60). 145 accidents had only a single recorded occurrence. The longest chain was 8 (for one LOC-I accident; NTSB ID: GAA17CA303). 80% of these reports included a code related to crashing into terrain/water. Thus, the potentially wide range of accident stories is reduced to a small set of short stories, most of which are some variation of "the pilot lost control and crashed into the ground/water". These problems are compounded by the lack of information about the cause for LOC-I. For instance, the most frequently used cause for fixed-wing LOC-I accidents is aircraft control not maintained – in other words, the pilot lost control because they did not maintain control (Houston et al., 2012; Franza & Fanjoy, 2012). So, we cannot easily determine why LOC-I happens, what most often causes it, or whether there have been any changes in its causes.

Several researchers have used NTSB codes to identify GA accident causes. Boyd (2015) found that failure to follow single engine procedures following loss of an engine was the highest factor in fatal twin-engine piston aircraft GA accidents under visual weather conditions. Fultz and Ashley's (2016) found that 60% of weather-related fatal accidents occurred in IMC. Bazargan and Guzhva (2007) found that hazardous weather and light conditions such as IMC and dark night conditions increased the likelihood that accidents would be fatal. Goldman et al. (2002) found that of maintenance errors, installation errors such as using the wrong parts were most likely to cause injury or fatality. Aguiar et al. (2017) found that GA accidents in mountainous terrain and high elevation environments most commonly involved CFIT and wind gusts/shear. Other analyses used NTSB accident narratives. Boyd and Stolzer (2016) identified accident-precipitating factors and found that not following the checklist/flight manual contributed the most to fatal or serious turbine-powered GA accidents. Ballard et al. (2013) considered three major risk factors for fatalities, post-crash fires, crashes after flight in IMC, and off-airport crashes (in other words, away from emergency services), and found that fatalities were most likely to occur in accidents occurring after flight in IMC contributed the most to fatal air tour accidents. Wiegmann et al. (2005)

used the Human Factors Analysis and Classification System (HFACS) to identify unsafe operator acts. 80% of the GA accidents were associated with at least one skill-based error such as handling. While these studies uncovered part of what causes GA accidents (e.g., flight into IMC is often involved in fatal accidents), they were not able to explain how, for example, IMC leads to fatal accidents.

Studies using NTSB data to understand LOC-I accidents face similar challenges. Previous work attempted to build chains of events in accidents using occurrence codes in the NTSB database. Rao and Marais (2015) found that 13.8% of 5051 GA rotorcraft fatal accidents had LOC-I as the first occurrence. Houston et al. (2012) found that 75% of the 147 instructional LOC-I accident reports cited LOC-I as the first occurrence – thus we cannot determine what led to the LOC-I. Other studies investigated the impact of aircraft characteristics on accidents. Franza and Fanjoy (2012) analyzed correlations between contributing factors to accidents from 2002–2012 in Cirrus SR20 and Piper PA28-161 aircraft. They also found that pilots' failure to maintain directional control contributed to 50% of fatal accidents in both aircraft models. Ud-Din and Yoon (2018) found that poor health and impairment due to medication, followed by poor manual control and inadequate pilot adherence to flight procedures were the most significant events for LOC during maneuvering.

One way to improve understanding is by modeling accidents. Several researchers have used Bayesian networks to identify causal factors and assess risk (Ancel & Shih, 2012; Ancel et al., 2015; Ayra et al., 2019; Xiao et al., 2020; Uğurlu et al., 2020). Ancel et al. (2015) developed an object-oriented Bayesian network (OOBN), based on HFACS, to model Part 121 and 135 LOC-I accidents. They identified organizational deficiencies as underlying flight-related and maintenance crew-related airline accidents. Bayesian networks are useful to visually represent a summary analysis of accidents. But they require detailed information that is often not available for GA accidents. The probability calculation for each node in a Bayesian network requires expert judgment and information from sources such as operators and aviation agencies. Further, some accident sequences have cyclic relationships, e.g., an aircraft stall may cause an LOC-I and vice versa. Since Bayesian networks are directed acyclic graphs, they cannot capture such cyclic relationships between aircraft states.

We investigate whether additional insight into fixed-wing LOC-I can be garnered from the NTSB database by taking a different modelling approach. We extend the state-based approach developed for rotorcraft accidents by Rao and Marais (2020) to fixed-wing aircraft accidents. Section 1 presents the state-based approach and develops a vocabulary of 108 state and 226 trigger definitions, along with a set of grammar rules for connecting states and triggers. We show how these rules can be used to logically infer some missing states and triggers. In Section 2, we identify and analyze GA LOC-I accidents using the approach and identify an additional 1,214 LOC-I accidents that were not directly identifiable using LOC-I

NTSB codes. Using the state-based approach, we found additional causes and different cause rankings for LOC-I accidents than those resulting from conventional analyses, such as those described earlier.

1. A state-based approach for fixed-wing aircraft accidents

Rao and Marais (2020) developed a state-based approach for modelling helicopter accidents by representing accidents as a sequence of states and triggers, rather than the event-centric current coding system. The state and trigger definitions are based on codes in the NTSB database that have been used for rotorcraft accidents (occurrence codes, finding codes, modifier codes, and phase of flight codes). In this section, we extend the state-based model to fixed-wing aircraft accidents.

1.1. State-based model introduction

The state-based model consists of two core concepts: accidents are modelled as a series of states and triggers; and, states and triggers (the dictionary) are ordered and linked by rules (the grammar), as shown in Figure 1.

The system comprises the aircraft and pilot(s) operating the aircraft. A state is a segment of time wherein a system exhibits a particular behavior. The nodes in Figure 1 represent states of a notional system where the first state represents the default or start state of the system and the last (end) state represents the system's behavior in the final segment of time in the accident. A system can be in only one state at any given point of time. There are two types of states: nominal and hazardous. A nominal state is a state of a system that is generally accepted as sufficiently safe by the applicable stakeholders. "Sufficiently safe" depends on the particular context and stakeholders. For example, safe states are those where the aircraft is operating in good weather with all systems functioning and with a competent and fit-to-fly pilot. A system is in a nominal state only when both the pilot and aircraft are in nominal states. A nominal state cannot lead directly to an accident state – it must be directly preceded by a hazardous state.

A hazardous state is an off-nominal state that may lead to an accident or an incident. For example, a pilot's poor physiological condition is a "pilot hazardous state", and loss of engine power is an "aircraft hazardous state".
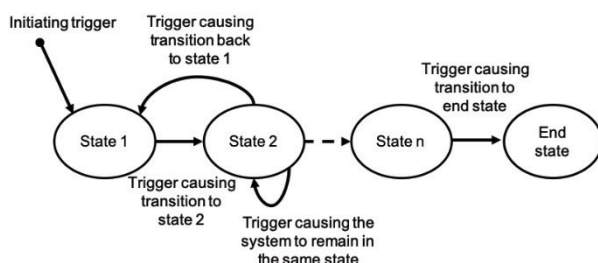

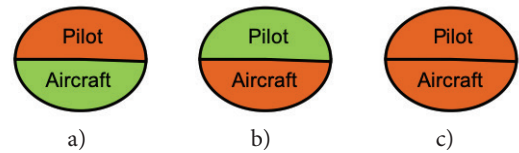Figure 1. State-based representation of a notional system

Figure 2. Illustration of three possible scenarios where a system is in hazardous state. Examples of these states are: (a) pilot's poor physiological condition; (b) loss of engine power; (c) pilot's poor physiological condition during a loss of engine power (adapted from Rao & Marais, 2020)

A system is in a hazardous state if either the pilot(s), the aircraft, or both the pilot(s) and aircraft are in hazardous states, as shown in Figure 2. We categorize hazardous states based on when they occur in an accident sequence. A preflight hazardous state is a hazardous state that exists before a flight starts, for example, preflight mechanical issue. An intermediary hazardous state occurs between a preflight state and an end state (in this case, an accident), for example, inflight loss of control. Each flight terminates in an end state, which can be nominal (e.g., safe landing), an incident (e.g., bounced landing), or an accident (e.g., midair collision).

A trigger is an event that occurs at a precise instant of time, causing either the aircraft, pilot(s), or both the aircraft and pilot(s) to transition between states or remain in the same state. For example, failure of an engine can cause a system to transition from a nominal state to a hazardous state. The links connecting to each state in Figure 1 represent triggers to each state. The *initiating trigger*, points to the default or start state of the system.

1.2. Fixed-wing aircraft dictionary of hazardous states, triggers, and additional information

The existing rotorcraft data dictionary has 84 state definitions and 182 trigger definitions. Here, we extend the data dictionary to fixed-wing aircraft accidents.

1.2.1. Fixed-wing state definitions

Building on the existing rotorcraft dictionary and using the NTSB database, we amended states and created new states, resulting in a set of 108 states that are applicable to fixed-wing aircraft, as shown in Table 1.

Fixed-wing aircraft differ from rotorcraft in four ways relevant to accident modelling: (1) Maneuvering. Rotorcraft and fixed-wing aircraft have different maneuvering capabilities due to different flight mechanics. For example, rotorcraft, unlike fixed-wing aircraft, can perform maneuvers such as hovering, and can autorotate in the event of losing engine power. So, these rotorcraft states are not applicable to fixed-wing aircraft. (2) Control surfaces. Fixed-wing aircraft, unlike rotorcraft, have ailerons, a rudder, and an elevator for aerodynamic stability. For example, fixed wing aircraft have flaps, unlike rotorcraft. Therefore, we created a new state for improper flaps extended speed (V_{FE}). (3) Takeoff and landing characteristics. Advanced

rotorcraft with wheels that can perform running takeoffs, hover taxi, and air taxi, are relatively rare in civil aviation, and therefore rotorcraft accidents associated with these maneuvers are also rare (there were no such accidents in the 34 years covered in Rao and Marais' 2020 analysis). We therefore created new fixed-wing states such as improper takeoff, improper taxi speed, water loop/swerve, and aircraft hydroplaning. (4) Airspeed factors. Fixed-wing aircraft have additional airspeeds to rotorcraft. For example, fixed-wing aircraft have five different airspeeds that convey takeoff or rotation speed: lift-off speed (V_{LOF}), takeoff safety speed (V_2), minimum takeoff speed (V_{2MIN}), rotation speed (V_R), and maximum speed from which the airplane can stop within the accelerate-stop distance (V_1). We define a new improper takeoff/rotation speed state, as

shown in Table 2. Table 2 also shows the Boolean logic for the improper takeoff or rotation speed state, which serves as input to our translation code. Similarly, we created 10 additional airspeed states for fixed-wing aircraft such as improper landing gear operating/ extended speed (V_{LO} and V_{LE}) and improper flaps extended speed (V_{FE}).

Finally, we added several states that may also apply to rotorcraft but did not appear in any of the rotorcraft accidents in the database. For example, Rao and Marais (2020) defined two LOC states for rotorcraft: inflight loss of control (LOC-I) and on-ground loss of control (LOC-G). Because the database does not always specify whether the LOC was inflight or on the ground, we created an unknown phase LOC state (LOC-U). Table 3 shows the definition and coding for the LOC-I state.

Table 1. Summary of the states for rotorcraft and fixed-wing aircraft

	Rotorcraft States	Fixed-wing States	State Example
States applicable only to rotorcraft	13	N/A	Improper autorotation only occurs in rotorcraft.
Rotorcraft states applicable to both rotorcraft and fixed-wing aircraft	54		Both rotorcraft and fixed-wing aircraft can experience hard landings.
Rotorcraft states re-coded for fixed-wing aircraft	17		Fixed-wing aircraft have additional airspeed types, compared to rotorcraft such as minimum takeoff speed (V2MIN).
New states defined for fixed-wing aircraft only	N/A	37	Rotorcraft do not have flaps, unlike fixed-wing aircraft, so we created improper flaps extended speed (VFE) state for fixed-wing aircraft.
Total states	84	108	

Table 2. Improper takeoff or rotation speed state definition for fixed-wing aircraft

NTSB Codes (Pre-2008)	Description
24507 AND (3101 OR 3104 OR 3107 OR 3108 OR 3109 OR 3112 OR 3115 OR 3120 OR 3122 OR 3127 OR 3129 OR 3138)	Airspeed, lift off speed (VLOF) AND ("Below" OR "Delayed" OR "Exceeded" OR "Excessive" OR "Improper" OR "Inattentive" OR "Inadequate" OR "Misjudged" OR "Not attained" OR "Not maintained" OR "Not obtained" OR "Reduced")
24568 AND (3000 OR 3011 OR 3107 OR 3127 OR 3129)	Airspeed, maximum speed from which the airplane can stop within the accelerate-stop distance (V1) AND ("Above" OR "Not obtained/maintained" OR "Exceeded" OR "Not obtained" OR "Not maintained")
24569 AND (3122 OR 3127)	Airspeed, takeoff safety speed (V2) AND ("Not attained" OR "Not maintained")
24570 AND (3011 OR 3115 OR 3122)	Airspeed, minimum takeoff safety speed (V2min) AND ("Not obtained/maintained" OR "Inadequate" OR "Not attained")

Table 3. Inflight loss of control (LOC-I) state definition for fixed-wing aircraft

NTSB Codes (Pre-2008)	Description
250	Loss of control – in flight
110	Altitude deviation, uncontrolled
553	Descent – uncontrolled phase
24524 AND (3140)	Descent AND ("Uncontrolled")
24525 AND (3140)	Proper descent rate AND ("Uncontrolled")
NTSB Codes (Post-2008)	Description
240	Loss of control in flight
650	Uncontrolled Descent
01062022	Pitch control
01062023	Lateral/bank control

1.2.2. Fixed-wing trigger definitions

Using the NTSB codes used for fixed-wing aircraft accidents, and combining codes that convey the same meaning, we defined 226 triggers (Table 4). Similar to hazardous states, we accounted for the differences between helicopters and fixed-wing aircraft when augmenting and creating new triggers for fixed-wing aircraft. For example, based on different speed characteristics of fixed-wing aircraft, we re-coded the rotorcraft trigger *improper aborted landing/takeoff* for fixed-wing aircraft by adding a subject code *24503 Abort above V_1* with its modifiers. V_1 is the takeoff decision speed, beyond which a flight can continue to take off even in case of an engine failure.

1.2.3. Additional information

Rao (2016) used *information codes* to translate NTSB codes that provide additional information about the prevailing conditions during an accident, but do not translate to states or triggers. Here, we amend this definition by adding a fourth category, pre-existing condition (PEC), and redefining the information codes to exclude PECs.

Pre-existing Condition (PEC): A condition in the aircraft's environment that remains true or applicable throughout a flight and is neither a state nor a trigger is defined as a pre-existing condition. We define three pre-existing conditions: unsuitable airport facilities, unsuitable runway, and unsuitable physical environment. For

example, unsuitable runway PEC gives information about a runway condition but does not describe a state or a trigger in an accident (Table 5).

Information code: Detail about a system that is neither a state, a trigger, *nor a pre-existing condition*, is defined as an information code. Information codes describe terrain/object(s) that an aircraft collided with and phases of flight in accidents. For example, the code *03022020: tree* indicates that an aircraft collided with a tree during the accident.

In the pre-2008 system, the NTSB uses subject codes 19200: terrain condition or 20200: object with modifiers to describe the type of terrain or objects. In the post-2008 system, the NTSB uses different finding codes to describe the type of terrain or objects with modifiers such as 91: contributed to outcome. Additionally, the NTSB uses a separate set of codes to describe phases of flight with each occurrence in accidents. Therefore, we defined three information code categories: information about objects, information about terrain, and information about phases of flight.

1.3. Illustrative example

We demonstrate the working of fixed-wing state and trigger definitions and the grammar rules using an accident (NTSB ID: ERA13FA059) that happened in November 2012 in Owls Head, Maine involving a Cessna 172N.

Table 4. Breakdown of rotorcraft and fixed-wing aircraft triggers

	Rotorcraft Triggers	Fixed-wing Triggers	Trigger Example
Triggers applicable only to rotorcraft	43	N/A	Rotor system failure can occur only in rotorcraft.
Rotorcraft triggers applicable to both rotorcraft and fixed-wing aircraft	76		Both fixed-wing and rotorcraft can experience improper engine shutdown.
Rotorcraft triggers re-coded for fixed-wing aircraft	63		Both rotorcraft and fixed-wing aircraft can experience improper aborted landing/takeoff, but there are additional NTSB codes that apply for fixed-wing aircraft.
New triggers defined for fixed-wing aircraft	N/A	87	Rotorcraft have rotors and not propellers, so they do not have a propeller control failure trigger.
Total triggers	182	226	

Table 5. Unsuitable runway pre-existing condition

NTSB Codes (Pre-2008)	Description
19201	Runway/landing area condition
NTSB Codes (Post-2008)	**Description**
03023000	Runway/land/takeoff/taxi surface – (general)
03023010	Runway/land/takeoff/taxi surface – Wet
03023015	Snow/slush/ice covered
03023020	Soft
03023025	Glassy
03023030	Choppy

During the departure roll, the aircraft collided with a ground vehicle that was crossing the runway, breaking the right elevator. The pilot continued taking off, stalled the aircraft, and went into a low-altitude spin before hitting the ground. The first two columns of Table 6 show the resulting NTSB codes for the accident report.

We model the accident in five steps:

1. **Identify states and triggers from the accident data:** We map the finding codes and occurrence codes from the database with corresponding states and triggers as shown in Table 6. Figure 3 shows the states and triggers. Since there are no codes indicating that the pilot was impaired or the aircraft was functioning improperly, we indicate their state as nominal.

2. **Identify preflight, intermediary, and end states:** Next, we identify the preflight, intermediary and end states, as shown in the last column of Table 6.

3. **Sequence hazardous states:** We apply the grammar rules to sequence hazardous states. The sequencing rules are based on flight physics and the sequence that the NTSB used to report accidents. See Rao and Marais (2020) for a detailed discussion of grammar rules. Figure 3 shows the accident model after applying the sequencing rules.

4. **Link states and triggers:** Using the grammar rules, we link triggers to the sequenced states, as shown in Figure 3. Three states do not have entering triggers, because the accident report does not mention any applicable trigger related codes.

5. **Infer triggers and states based on grammar rules:** The NTSB codes for an accident may not be sufficient to identify all states and triggers in that accident, as shown by the three missing triggers in Figure 3. We use the state-and-trigger sequencing rules to infer some of the missing information.

Consider for example the trigger recovery action not possible after loss of control. We infer this trigger whenever an end state succeeds a loss of control state in an accident, and the accident does not include any codes related to an improper remedial action or a lack of action triggers (Loss of control state AND (end state) AND NOT ("Improper remedial action" trigger OR "Lack of action" trigger).

2. Analysis of inflight loss of control accidents

This section compares conventional and state-based statistical analyses of LOC-I accidents involving fixed-wing aircraft operating under 14CFR Part 91 that occurred from 1999–2017 and are recorded in the NTSB database. Because the NTSB coding system changed in 2008, we consider two different time frames: 1999–2008 and 2009–2017.

Table 6. NTSB Codes and Corresponding States or Triggers (NTSB ID: ERA13FA059)

| Finding code/ Occurrence code	Modifier/Person/Phase Code	Resulting State/Trigger	Preflight/intermediary/ end state
2041015: Incorrect action performance	44: Pilot	Improper action performance trigger	N/A
1062022: Pitch control	20: Not attained/maintained	Inflight loss of control state	Intermediary
2041030: Lack of action	48: Airport personnel	Lack of action trigger	N/A
3036000: Light condition	91: Contributed to outcome	Prevailing weather and light state	Preflight
320: Runway incursion	300: Takeoff	Runway incursion state	Intermediary
490: Collision during takeoff/ landing	300: Takeoff	Collision during takeoff/landing state	Intermediary
240: Loss of control in-flight	350: Initial Climb	Inflight loss of control state	Intermediary
470: Collision with terrain/ object (non-CFIT)	650: Uncontrolled descent	Inflight collision with terrain/object state	End
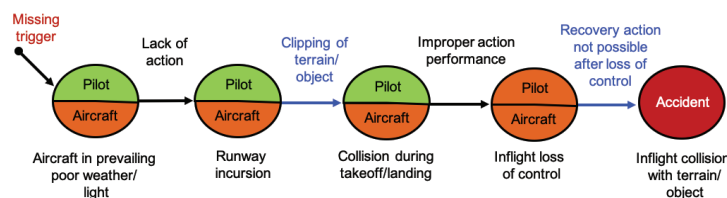
Figure 3. State-based representation of the accident.

Note: Using the grammar rules, we infer triggers to the sequenced states that have missing triggers. The links (and text) in blue are the inferred triggers. The text in red shows a missing trigger that cannot be inferred using the grammar rules.

2.1. Conventional analysis of LOC-I accidents

Conventional analyses such as those discussed in the introduction, analyse the relative frequencies with which NTSB codes are cited in accident reports. Here we do such an analysis for LOC-I accidents. In the pre-2008 coding system, LOC-I accidents are indicated by *250: Loss of control – in flight*; in the post-2008 system, by *240: Loss of control – in flight*. The NTSB uses subject code and modifier combinations (and finding codes in the post-2008 system) to provide greater detail about the level of contribution each finding had to the outcome. Each combination is designated as a cause or contributing factor. We identified the subject code and modifier combinations that the NTSB designated as causes (denoted as $Cause_j$) and calculated the presence for each subject code and modifier combination as the number of times it was used at least once in an accident, normalized by the total number of accidents (cf. Sorenson & Marais, 2016).

$$presence\left(Cause_j \mid Accident\right) =$$
$$\frac{\sum_{i=1}^{n_{accidents}} TRUE\left(Cause_j \geq 1 \mid Accident_i\right)}{Total\ Accidents}. \quad (1)$$

We use the same method to calculate the presence of contributing factors in accidents. In this section, we discuss the most frequent (top) causes and contributing factors in LOC-I accidents. Table 7 shows the top ten causes for LOC-I accidents in 1999–2008. The highest cause of LOC-I is that the pilot did not maintain aircraft control (24566–3127: Aircraft control − not maintained), thus providing no indication of what caused the loss of control. Failure to maintain airspeed (24506–3127: Airspeed − not maintained) has the second highest presence, providing at least some suggestion that airspeed is an important factor in LOC-I accidents. This finding is corroborated by the other top-ten causes related to airspeed in our analysis.

Table 8 shows the top ten causes for LOC-I accidents in 2009–2017. Three of the top four causes of LOC-I, aircraft control − pilot, directional control − not attained/maintained, and performance/control parameters − not at-

tained/maintained, are variations of "loss of control". Only the second highest cause, airspeed − not attained/maintained provides some indication of what happened during the LOC-I accident. Some of the top causes in 2009–2017 were different from the 1999–2008 findings. In 2009–2017, we found more causes related to aircraft performance (performance/control parameters − not attained/ maintained) and pilot actions (decision making/ judgment − pilot, incorrect action performance − pilot, and angle of attack − not attained/maintained). This difference arises from two main reasons: (1) the NTSB codes and their descriptions changed in 2008 (the NTSB built an entirely new coding system); and (2) the NTSB started using some code descriptions more extensively as causes in post-2008 LOC-I accidents compared to the pre-2008 accidents. For example, the NTSB used aircraft performance related finding as a cause more in the post-2008 system (1062000-20: Performance/control parameters − not attained/maintained) than in the pre-2008 system (17300: Aircraft performance).

The pre-2008 coding system has four separate subject codes to indicate decision making or judgment (24000: Planning/decision, 24010: Inflight planning/decision, 24031: Improper decision, and 60000: Judgment) as compared to only one finding code in the post-2008 system (02041520: Decision making/judgment). Out of these four subject code and modifier combinations in the pre-2008 system, 24010-3109: Improper inflight planning/decision appears the most frequently (2.15%) in 1999–2008 (cited as 12th most frequent cause in LOC-I accidents). The NTSB used decision making/judgment more frequently as a cause in the post-2008 LOC-I accidents than in the pre-2008 accidents (presence of 11.29%). Unlike in the post-2008 system, there are no codes in the pre-2008 system that indicate incorrect action by pilot or angle of attack and therefore these are some new causes that we identified in 2009–2017. Some of the top causes from 1999–2008 such as stall − inadvertent and stall/spin − inadvertent were not identifiable in 2009–2017 because the post-2008 coding system does not use any stall- or spin-related finding codes as causes. However, the post-2008 system uses 241: Aerodynamic stall/spin as an occurrence code in 16.79% of LOC-I accidents.

Table 7. Top ten causes for LOC-I accidents in 1999–2008

Subject code with modifier	Description	Presence, %
24566–3127	Aircraft control − Not maintained	27.43
24506–3127	Airspeed − Not maintained	21.97
24551–3113	Stall − Inadvertent	7.68
24511–3127	Airspeed, stall (Vs) − Not maintained	5.78
25000–3001	Reason for occurrence undetermined − No Modifier	5.60
24539–3127	Directional control − Not maintained	5.20
24026–3115	Compensation for wind conditions − Inadequate	4.22
24552–3113	Stall/spin − Inadvertent	3.93
33400–0	Spatial disorientation − No Modifier	3.49
24506–3115	Airspeed − Inadequate	2.84

Table 9 shows the top contributing factors to LOC-I accidents in 1999–2008. From 1999–2008, weather and light conditions, collision with objects, and low altitude were the top contributing factors in LOC-I accidents. Adverse weather conditions appear in 33.25% of LOC-I accidents, out of which wind gust (6.37%) and crosswind (4.69%) are the most dangerous weather conditions for LOC-I. Further, 6.11% of LOC-I accidents involved collision with an object, out of which collision with tree(s) was the most frequent.

Table 10 shows the top ten contributing factors in LOC-I accidents from 2009–2017. In contrast to the pre-2008 results, terrain, object, or weather-related codes do not appear in the top ten factors. This difference in top causes is because the NTSB has tended to designate weather and object related finding code-modifier combinations as causes rather than factors in the post-2008 system. In the post-2008 coding system, the NTSB used pilot-related codes more as factors which give more information about pilot's decision making, actions, physical wellness, and experience.

Table 8. Top ten causes for LOC-I accidents in 2009–2017

Finding code	Description	Presence, %
206304044	Aircraft control – Pilot	59.39
106201020	Airspeed – Not attained/maintained	21.10
106200020	Performance/control parameters – Not attained/ maintained	18.43
106202020	Directional control – Not attained/maintained	11.68
204152044	Decision making/judgment – Pilot	11.29
500000000	Not determined – No Modifier	7.03
202202544	Spatial disorientation – Pilot	6.41
106204220	Angle of attack – Not attained/maintained	5.90
204101544	Incorrect action performance – Pilot	5.62
206304046	Aircraft control – Student pilot	4.59

Table 9. Top ten contributing factors for LOC-I accidents in 1999–2008

Subject code with modifier	Description	Presence, %
20000–2207	Weather condition – Gusts	6.37
20000–2203	Weather condition – Crosswind	4.69
20200–2517	Object – Tree(s)	3.20
24518–3004	Altitude – Low	3.02
20000–2214	Weather condition – Low ceiling	2.91
20000–2212	Weather condition – High density altitude	2.58
20100–2305	Light condition – Dark night	2.55
20000–2204	Weather condition – Clouds	2.44
20000–2222	Weather condition – Tailwind	2.36
24551–3113	Stall – Inadvertent	1.89

Table 10. Top ten contributing factors for LOC-I accidents in 2009–2017

Finding Code	Description	Presence, %
204152044	Decision making/judgment – Pilot	4.08
203102544	Total experience with equipment – Pilot	1.87
206304044	Aircraft control – Pilot	1.08
206301544	Use of equipment/system – Pilot	0.96
201202544	Prescription medication – Pilot	0.91
500000000	Not determined – No Modifier	0.79
206151044	Preflight inspection – Pilot	0.74
201202044	Illicit drug – Pilot	0.68
204102540	Delayed action – Instructor/check pilot	0.68
201203044	OTC medication – Pilot	0.62

2.2. State-based analysis of LOC-I accidents

This section presents the top hazardous states and triggers using the state-based approach and compares our findings to the conventional analysis. We identified LOC-I accidents in 1999–2017 using the data dictionary to map the NTSB codes to the LOC-I state definition. Our conventional analysis identified 4,512 LOC-I accidents in 1999–2017. By mapping the NTSB codes to the LOC-I state, we found 1,214 additional LOC-I accidents, as shown by year in Figure 4.

We calculated the presence of hazardous states and triggers in the LOC-I accidents using Equation (1). Figure 5 compares the top hazardous states for LOC-I accidents in 1999–2008 and 2009–2017, ranked based on the top hazardous states for the recent years (2009–2017). Figure 5 shows that in 2009–2017, most (77.8%) of the LOC-I accidents typically ended with an inflight collision with terrain/water/object.
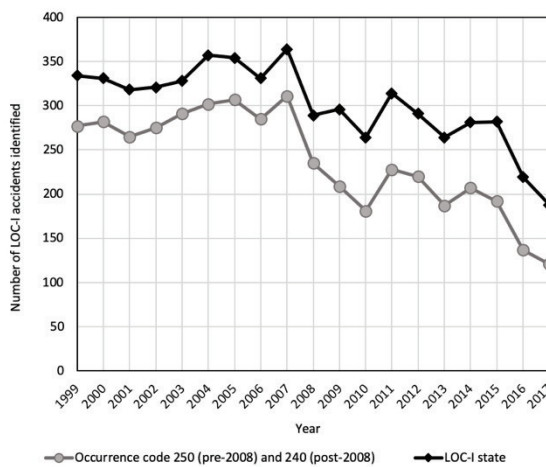
Figure 4. Comparison of number of LOC-I accidents found using the NTSB codes vs. state definitions
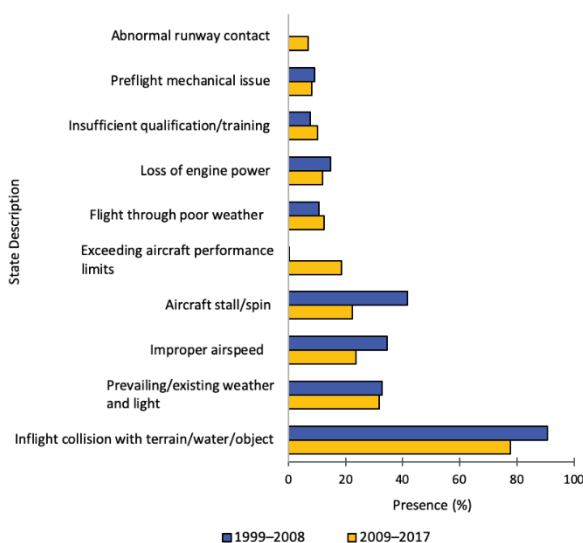

Figure 5. Comparison of presence of top hazardous states for LOC-I accidents in 1999–2008 and 2009–2017

The state-based analysis helped in identifying some new findings that could not be identified from a conventional analysis, such as abnormal runway contact and exceeding aircraft performance limits. As shown in Figure 5, we identified abnormal runway contact state only in 2009–2017 accidents because the NTSB's pre-2008 coding system does not have any codes to describe abnormal contact of the aircraft with the runway. Similarly, the pre-2008 system contains only one code that indicates exceeding aircraft performance limits (17300: Aircraft performance), as compared to three different codes in the post-2008 system. The NTSB did not use the pre-2008 subject code (17300) extensively in the 1999–2008 accidents and it therefore has a presence of only 0.28%. We identified new findings such as preflight mechanical issue and insufficient qualification/training as important causes for LOC-I, with a presence of 8.13% and 10.15% respectively in 2009–2017 accidents. Preflight mechanical issue involves scenarios such as improper weight and balance calculations by pilot and operating an aircraft with known deficiencies. Insufficient qualification/training includes lack of experience in a type of aircraft, night or instrument flying, inadequate flight training, and the pilot not being current in their certification.

Weather factors such as prevailing/existing weather and light conditions (31.72% in 2009–2017 and 32.88% in 1999–2008) and flight through poor weather (12.38% in 2009–2017 and 10.73% in 1999–2008) play a major role in LOC-I accidents. Although the number of prevailing weather-related codes increased from just two codes to 47 in the new NTSB coding system, the number of times that the NTSB cited codes describing "prevailing weather/light conditions" or "flying through a poor weather" in LOC-I accidents remains similar.

Aircraft stall/spin appears in only 22.3% of LOC-I accidents in 2009–2017 as compared to the high presence (41.72%) in pre-2008 LOC-I accidents because the pre-2008 coding system has two stall/spin related finding codes whereas the post-2008 coding system has only one stall related occurrence code, but no finding codes.

Further, pilot in a disoriented/lacking awareness state (not shown in Figure 5) was present in 6.42% of 2009–2017 accidents. This state involves situations when a pilot loses a reference point especially when flying through poor weather (for example, low visibility and instrumental meteorological conditions).

Figure 6 shows the top ten triggers in LOC-I accidents. Additional findings into LOC-I accidents such as improper inflight planning/decision-making, improper maintenance, improper preflight planning, and improper use of procedure or directives could not be identified from a conventional analysis of the NTSB database. In 2009–2017, improper inflight planning/decision-making has the highest presence (17.34%) in LOC-I accidents. This trigger involves scenarios such as reduced/improper judgement or decision-making by pilot, and not recognizing or comprehending risks. Improper maintenance and preflight planning put the flight in a hazardous state
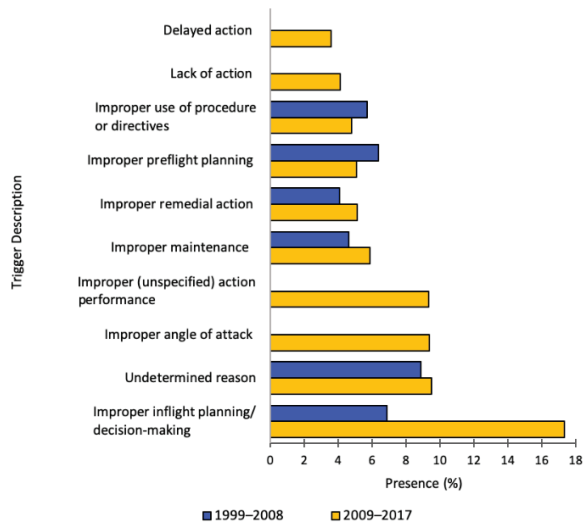
Figure 6. Comparison of presence of top triggers for LOC-I accidents in 1999–2008 and 2009–2017
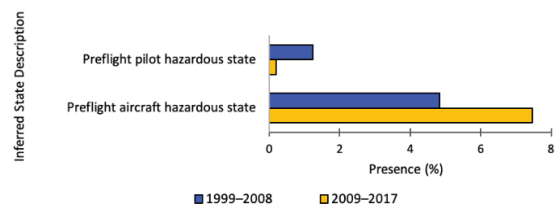

Figure 7. Presence of inferred states in LOC-I accidents in 1999–2008 and 2009–2017

Figure 8. Comparison of presence of top inferred triggers in LOC-I accidents in 1999–2008 and 2009–2017

(such as an unsafe to fly aircraft or severe weather conditions) even before it starts. Some of the top triggers such as delayed action, lack of action, improper action performance, and improper angle of attack were only cited in the NTSB's post-2008 coding system, since the pre-2008 coding system does not use any relevant codes to describe such pilot actions. The trigger Undetermined reason, although used extensively (8.87% in 1999–2008 and 9.50% in 2009–2017), does not provide any useful information about how LOC-I happened.

Next, we use the state-based approach to identify omissions in accident coding and infer the missing information using the grammar rules. Figure 7 shows the inferred hazardous states. 4.84% and 7.46% of LOC-I accidents in 1999–2008 and 2009–2017 respectively did not cite any codes related to aircraft preflight hazardous state (such as preflight mechanical issue state and preflight low engine fluids states). We inferred this state by using other trigger codes that implied that the aircraft was in a hazardous state before starting the flight. Similarly, we inferred preflight pilot hazardous state in 1.23% and 0.20% of LOC-I accidents in 1999–2008 and 2009–2017 respectively. Figure 8 shows the inferred hazardous triggers. 94.87% of the LOC-I accidents had no corresponding codes to describe how an LOC-I state leads to an accident. Using our logic rules, we inferred the trigger recovery action not possible from loss of control which represents the missing data in the accidents.

While the NTSB reports terminating occurrences (or end states in this paper) that immediately followed the LOC-I state, the accident codes do not indicate what triggers the aircraft to transition from the LOC-I state to the end state. In some cases, the NTSB codes translate to triggers that described how an LOC-I state transitioned to an accident such as improper remedial action (which is used in 5.14% of LOC-I accidents) and lack of action (presence of 4.13%).
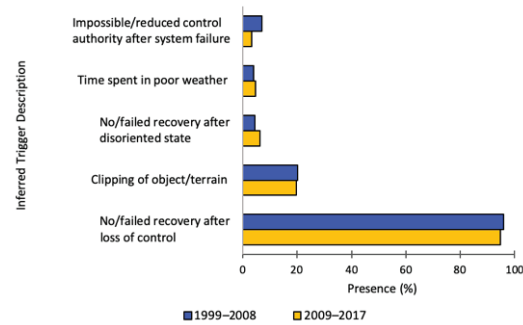
In 19.59% of LOC-I accidents, aircraft clipped (hit) terrain or object and continued the flight, suggesting that the aircraft did not crash and collided with the object or terrain, thus inferring clipping of object/terrain trigger. In 3.29% of LOC-I accidents, the NTSB database did not describe how an aircraft transitioned from system failure to LOC-I state. We infer impossible/reduced authority after system failure state trigger for such accidents, where no other related trigger codes were used. We inferred the trigger no/failed recovery from disoriented state (6.29% in 2009–2017) whenever a disoriented pilot directly transitioned into a loss of control state, with no related trigger information to describe the transition. We also inferred time spent in poor weather (4.79% in 2009–2017) as a trigger to pilot's disoriented state when the NTSB cites prevailing weather/light or a flight through poor weather as the immediate former state with no related trigger information.

Conclusions

We extended Rao and Marais' (2020) state-based approach for rotorcraft accidents to fixed-wing aircraft accidents by modifying the existing rotorcraft state and trigger definitions and adding a total of 130 new states, triggers, and additional information applicable to fixed-wing aircraft. We created a new category to store additional accident information called pre-existing condition (PEC) that describes an aircraft's environment that remains true throughout a flight. We developed a new set of grammar rules to sequence states and link triggers to states. These grammar rules help to logically infer some of the missing information and provide additional insights into accidents. We

investigated the usefulness of the state-based approach to model fixed-wing LOC-I accidents and revealed some new findings that were not discernible from the conventional analysis.

The state-based approach steps away from the chain of events accident modeling technique by viewing aviation accident as a set of hazardous states and triggers. Our approach also helps to provide a more correct count of the LOC-I accidents and their causes in the NTSB database by accounting for coding redundancies. By mapping the LOC-I state definition codes, we identified 1,214 additional LOC-I accidents that had not been labelled as such in the NTSB database.

The conventional analysis provides little information about LOC-I accident causation by using tautologies of LOC-I (such as directional control and aircraft control). These causes provide additional information about the type of LOC-I (directional or aircraft), but do not mention why loss of control happened. The state-based approach helped to provide a deeper statistical understanding of the LOC-I accidents in the NTSB database. We ranked the top hazardous states and triggers in 5,726 LOC-I accidents in two different timeframes 1999–2008 and 2009–2017 to understand the causal patterns in LOC-I accidents. In addition to the already known causes of LOC-I such as prevailing weather and light conditions and improper airspeed from the conventional analysis, our state-based analysis reveals that hazardous states such as exceeding aircraft performance limits, insufficient qualification/ training, and preflight mechanical issues are prevalent in LOC-I accidents. We also found that triggers such as improper inflight planning/decision-making, preflight planning, and improper use of procedures are some of the top causes for LOC-I.

The NTSB database sometimes omits important findings codes from accidents. This approach helps infer missing codes from reports and construct logical accident sequences (or stories). By using the grammar rules to model the LOC-I accidents, we inferred that aircraft clipping with object or terrain caused LOC-I in 19.9% of LOC-I accidents in 2009–2017, a finding that was not discernible from the conventional analysis. Additionally, we inferred that 4.84% and 7.46% of the accidents in 1999–2007 and 2009–2017 respectively had missing information about hazardous aircraft state before the start of the flight. These additional insights help to provide a better understanding of loss of control accidents. Further, considering these additional insights in loss of control prevention and recovery training techniques may help in reducing LOC-I accidents and incidents in the future.

31.8% of LOC-I accidents do not record any codes relevant to the trigger definitions and 9.6% of LOC-I accidents do not record any codes relevant to the preflight state definitions. In future work, we plan to expand the grammar rules to potentially infer specific triggers that cause LOC-I and the preflight states that lead to accidents. Text mining offers an additional source of state and trig-

ger information. In related work, we found that the narratives sometimes contain new and detailed information as compared to the NTSB codes. We found this information by "manually" reading each narrative, which is time consuming, tedious, and prone to subjectivity. An automated text mining approach can alleviate these issues and yield additional information. Augmenting our state-based model by including such a text mining approach can offer an additional source of state and trigger information to code and model accidents more efficiently. Having more data from the narratives would likely reveal new states and triggers. To address this aspect, we consider machine learning as a way to automatically identify potential hazardous states, triggers and new grammar rules and thus create a "self-developing coding system" based on the extracted information from the accident narratives. Further, by investigating potential associations between pre-existing conditions (PECs) and hazardous states, we can create additional rules to gain more insights from accident modeling. For example, the grammar rules may help to find the likelihood of the PEC, wet runway condition, to be associated with the landing to overrun state.

Funding

Author contributions

NM was responsible for the design and development of the data analysis and wrote the first draft of the article. KM provided oversight of the analysis and reviewed the paper drafts. AR helped develop the data analysis and reviewed the paper drafts.

Disclosure statement

We declare that we have no significant competing financial, professional, or personal interests that might have influenced the performance or presentation of the work described in this manuscript.

References

Aguiar, M., Stolzer, A., & Boyd, D. D. (2017). Rates and causes of accidents for general aviation aircraft operating in a mountainous and high elevation terrain environment. *Accident Analysis and Prevention*, *107*, 195–201. https://doi.org/10.1016/j.aap.2017.03.017

Aircraft Owners and Pilots Association. (2018). *27ᵗʰ Joseph T. Nall report: General aviation accidents in 2015*. AOPA. https://www.aopa.org/-/media/files/aopa/home/training-and-safety/

nall-report/27thnallreport2018.pdf?la=en&hash=C52F88B3 8FD95CB7C0A43F3B587A12E2692A8502

Ancel, E., & Shih, A. (2012, September). The analysis of the contribution of human factors to the in-flight loss of control accidents. In *12th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference and 14th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference* (p. 5548). Indianapolis, Indiana. https://doi.org/10.2514/6.2012-5548

Ancel, E., Shih, A. T., Jones, S. M., Reveley, M. S., Luxhøj, J. T., & Evans, J. K. (2015). Predictive safety analytics: Inferring aviation accident shaping factors and causation. *Journal of Risk Research*, *18*(4), 428–451. https://doi.org/10.1080/13669877.2014.896402

Ayra, E. S., Ríos Insua, D., & Cano, J. (2019). Bayesian network for managing runway overruns in aviation safety. *Journal of Aerospace Information Systems*, *16*(12), 546–558. https://doi.org/10.2514/1.I010726

Ballard, S.-B., Beaty, L. P., & Baker, S. P. (2013). US Commercial air tour crashes 2000–2011: Burden, fatal risk factors and FIA Score Validation. *Accident Analysis & Prevention, 57*, 49–54. https://doi.org/10.1016/j.aap.2013.03.028

Bazargan, M., & Guzhva, V. S. (2007). Factors contributing to fatalities in general aviation. *World Review of Intermodal Transportation Research, 1*(2), 170–182. https://doi.org/10.1504/WRITR.2007.013949

Boyd, D. (2015). Causes and risk factors for fatal accidents in non-commercial twin engine piston general aviation aircraft. *Accident Analysis and Prevention, 77*, 113–119. https://doi.org/10.1016/j.aap.2015.01.021

Boyd, D. D., & Stolzer, A. (2016). Accident-precipitating factors for crashes in turbine-powered general aviation aircraft. *Accident Analysis & Prevention*, *86*, 209–216. https://doi.org/10.1016/j.aap.2015.10.024

Federal Aviation Administration. (2019). *Fly Safe: Prevent loss of control accidents*. FAA. https://www.faa.gov/newsroom/fly-safe-prevent-loss-control-accidents-34?newsId=94566

Franza, A., & Fanjoy, R. (2012). Contributing factors in Piper PA28 and cirrus SR20 aircraft accidents. *Journal of Aviation Technology and Engineering, 1*(22), 90–96. https://doi.org/10.5703/1288284314662

Fultz, A. J., & Ashley, W. S. (2016). Fatal weather-related general aviation accidents in the United States. *Physical Geography, 37*(5), 291–312. https://doi.org/10.1080/02723646.2016.1211854

General Aviation Joint Steering Committee. (2016). *GAJSC Pareto*. GAJSC. https://www.gajsc.org/2016/01/ga-safety-performance-fy16/pareto/

Goldman, S. M., Fiedler, E. R., & King, R. E. (2002). *General aviation maintenance-related accidents: A review of ten years (1988–1997) of NTSB Data. DOT/FAA/AM-02/23*. Office of Aerospace Medicine, Washington. https://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/2000s/media/0223.pdf

Houston, S. J., Walton, R. O., & Conway, B. A. (2012). Analysis of General Aviation instructional loss of control accidents. *The Journal of Aviation/Aerospace Education and Research*, *22*(1), 35–49. https://doi.org/10.15394/jaaer.2012.1402

National Transportation Safety Board. (1998). *Aviation Coding Manual*. Washington DC, National Transportation Safety Board. NTSB. https://www.ntsb.gov/GILS/Documents/codman.pdf

National Transportation Safety Board. (2019a). *Aviation accident database & synopses*. NTSB. https://www.ntsb.gov/_layouts/ntsb.aviation/index.aspx

National Transportation Safety Board. (2019b). *Government Information Locator Service (GILS): Aviation accident database*. NTSB. https://www.ntsb.gov/GILS/Pages/AviationAccident.aspx

Rao, A. H. (2016). *A new approach to modeling aviation accidents* [Doctoral dissertation, Purdue University, USA]. https://docs.lib.purdue.edu/open_access_dissertations/993

Rao, A. H., & Marais, K. (2015). Identifying high-risk occurrence chains in helicopter operations from accident data. In *15th AIAA Aviation Technology, Integration, and Operations Conference* (p. 2848). American Institute of Aeronautics and Astronautics. https://doi.org/10.2514/6.2015-2848

Rao, A. H., Fala, N., & Marais, K. (2016). Analysis of helicopter maintenance risk from accident data. In *AIAA Infotech @ Aerospace* (p. 2135). American Institute of Aeronautics and Astronautics. https://doi.org/10.2514/6.2016-2135

Rao, A. H., & Marais, K. (2020). A state-based approach to modeling general aviation accidents. *Reliability Engineering and System Safety, 193,* 106670. https://doi.org/10.1016/j.ress.2019.106670

Sorenson, D. & Marais, K. (2016). Patterns of causation in accidents and other systems engineering failures. In *IEEE Systems Conference*. IEEE, Orlando, FL. https://doi.org/10.1109/SYSCON.2016.7490568

Ud-Din, S., & Yoon, Y. (2018). Analysis of loss of control parameters for aircraft maneuvering in general aviation. *Journal of Advanced Transportation, 2018,* 7865362. https://doi.org/10.1155/2018/7865362

Uğurlu, Ö., Yıldız, S., Loughney, S., Wang, J., Kuntchulia, S., & Sharabidze, I. (2020). Analyzing collision, grounding, and sinking accidents occurring in the Black Sea utilizing HFACS and Bayesian networks. *Risk Analysis, 40*(12), 2610–2638. https://doi.org/10.1111/risa.13568

Wiegmann, D., Faaborg, T., Boquet, A., Detwiler, C., Holcomb, K., & Shappell, S. (2005). *Human error and general aviation accidents: A comprehensive, fine-grained analysis using HFACS*. Office of Aerospace Medicine, Washington, DC. https://apps.dtic.mil/dtic/tr/fulltext/u2/a460866.pdf

Xiao, Q., Luo, F., & Li, Y. (2020). Risk assessment of seaplane operation safety using Bayesian network. *Symmetry*, *12*(6), 888. https://doi.org/10.3390/sym12060888