

# TOOLS AND APPROACHES FOR WORKING WITH AVIATION ACCIDENT DATA

Jana Price, Ph.D., and Deborah Bruce, Ph.D.  
National Transportation Safety Board  
Washington, DC

The National Transportation Safety Board investigates approximately 2,000 aviation accidents per year and is responsible for maintaining the official United States census of civil aviation accident data. This paper summarizes categories of aviation accidents; explains the phases of an NTSB accident investigation; describes the history and scope of the NTSB's accident database; and characterizes the structure and content of the database. Finally, it provides a brief summary of NTSB data available to researchers and identifies research products that have resulted from NTSB accident data.

The National Transportation Safety Board (NTSB) investigates all aviation accidents that occur in the United States involving civil aircraft and certain aircraft involved in public use operations. The NTSB also provides investigators to serve as U.S. Accredited Representatives as specified in international treaties for aviation accidents overseas involving U.S.-registered aircraft or involving aircraft or major aircraft components of U.S. manufacture. Investigations are conducted from NTSB Headquarters in Washington, D.C., or from one of ten regional offices in the United States.

## Categories of Aviation Accidents

An "aviation accident" is defined in 49 Code of Federal Regulations (CFR), Part 830.2 as, "an occurrence associated with the operation of an aircraft which takes place between the time any person boards the aircraft with the intention of flight and all such persons have disembarked, and in which any person suffers death or serious injury, or in which the aircraft receives substantial damage." The database also contains a select number of aviation "incidents," defined as, "occurrences other than accidents that are associated with the operation of an aircraft and that affect or could affect the safety of operations."

The NTSB is responsible for determining the probable cause of accidents involving U.S.-registered aircraft flying for both commercial and private purposes. Title 14 of the CFR divides commercial air carrier operations into three major groups, Part 121, Part 135, and Part 129. Briefly stated, Part 121 applies to major airlines and cargo carriers that fly large transport-category aircraft; Part 135 applies to commercial air carriers commonly referred to as commuter airlines (i.e.,

scheduled Part 135) and air taxis (i.e., nonscheduled Part 135); and Part 129 applied to foreign commercial operators.

Although accidents among large transport-category aircraft receive substantial media attention, most aviation accidents are associated with general aviation operations. In 2002, 1,700 accidents involved general aviation flights as compared to 108 that involved commercial flights (40 Part 121, 9 scheduled Part 135, and 59 nonscheduled Part 135).

General aviation can be described as any civil flight that is not covered under Parts 121, 129, or 135. General aviation operations can include any type of aircraft (small fixed-wing aircraft are the most common, but transport-category aircraft repositioning between commercial flights, balloons, registered ultralight aircraft, helicopters, and gliders can also be general aviation operations). Excluded from general aviation are military or government public use operations; unregistered or illegal aircraft operations; and foreign-registered aircraft.

## Phases of Accident Investigation

When an accident or incident occurs, a decision is made at the NTSB regarding the type of investigation to be performed. The type of investigation is generally indicative of its scope: Categories include major, field major, field, limited, incident, and public use aircraft investigations. For example, a major investigation is conducted from headquarters and typically involves multiple investigative groups (e.g., power plants, human performance, air traffic control) and numerous organizations that may be parties to the investigation (e.g., operators, manufacturers, pilot unions).

A field investigation, conducted by an investigator

from one of the regional offices, may involve only a few participating organizations, such as the aircraft manufacturer and the Federal Aviation Administration (FAA). These participants are called parties to the investigation. Limited investigations generally do not involve parties to the investigation and rely primarily on data collected by FAA inspectors and pilot reports.

The NTSB employs approximately 60 regional aviation accident investigators who participate in the creation of new accident/incident records. At least one staff member performs quality control checks at each of the ten regional offices located throughout the United States. Preliminary reports are entered into the central database within 5 business days of an accident or incident. Investigators then complete a more detailed factual report, which is transmitted to headquarters. The factual report is available on the NTSB website shortly thereafter. In most cases, the probable cause is available approximately one month after the factual report. Most investigators perform 7 to 15 field investigations per year, plus an additional 25 to 50 limited investigations.

For each accident record, a standard set of data is routinely collected. Additional information may be collected during the investigation depending on the circumstances of the accident. For example, accidents that initially appear to be related to pilot error may include more detailed training information than for accidents caused by mechanical failure.

#### History and Scope of the NTSB Aviation Accident Database

Although the NTSB was established in 1967, its collection of accident records dates back to 1962 for airlines and 1964 for general aviation, commuters, and air taxi operations. These accidents were investigated by the NTSB's predecessor agency, the Civil Aeronautics Board.

Prior to 1982, investigators completed an accident reporting form and sent it to headquarters for processing and storage. Beginning in the mid-1970s, headquarters personnel keypunched coded data and short narratives that were stored in a database management system supported by a Department of Transportation computer. By the late 1970s, the NTSB had transferred the database to its own computer system.

Between 1982 and 1984, data collection forms were expanded to include a longer narrative, more detailed sequence of events and supplemental forms covering issues such as occupant survival information, injured ground personnel, and crash kinematics.

Data entry and management was centralized at the NTSB headquarters until the early 1990s when the NTSB placed mini-computers in each of its regional offices for data entry and transitioned to a mainframe (VAX) system for data storage. During this period, data coding was streamlined and many supplemental data collection forms were no longer entered into the database.

Most recently, in 2000, the NTSB transitioned all data from 1982 forward from the hierarchical data structure maintained in the System1032 RDMS on a VAX mainframe to a fully relational SQL Server database known as the Accident Data Management System (ADMS).

Reports generated from accident data between 1982 to the present are available at <http://www.nts.gov/avdata/>. Accident data prior to 1982 is currently stored in both System1032 RDMS and MSAccess 95, but will eventually be assimilated into the existing relational database. Docket materials for this time period are currently stored on microfiche.

The current ADMS system contains records of approximately 40,000 accidents and approximately 2,000 new event records are added to the database each year. Each record contains over 650 fields of data relating to personnel (e.g., pilots, flight crew, passengers), aircraft and the environment.

#### Accident Database Structure

Figure 1 shows the logical relationships for data that exist in the current NTSB Aviation Accident Database, and Table 1 provides a brief description of each data table that is available to the public. At its highest level, the database is organized around the *events* table (i.e., accidents or incidents). Associated with events are date, location, weather, etc. Because an event may involve multiple aircraft (as would be the case in collisions), the *aircraft* table is structured under the *events* table. Several other tables fall under the *aircraft* table, including *engines*,

*flight crew*, and *occurrences*. In each case, they indicate a one-to-many relationship (e.g., one aircraft may have several engines.)

Finally, note that three tables are prefaced with “dt” (e.g., *dt\_aircraft*). These indicate “detail tables” that contain responses to questions that may have

multiple responses. For example, “runway conditions” may include any or all of the following: wet, grooved, rough, ice-covered, soft, holes, vegetation, etc.

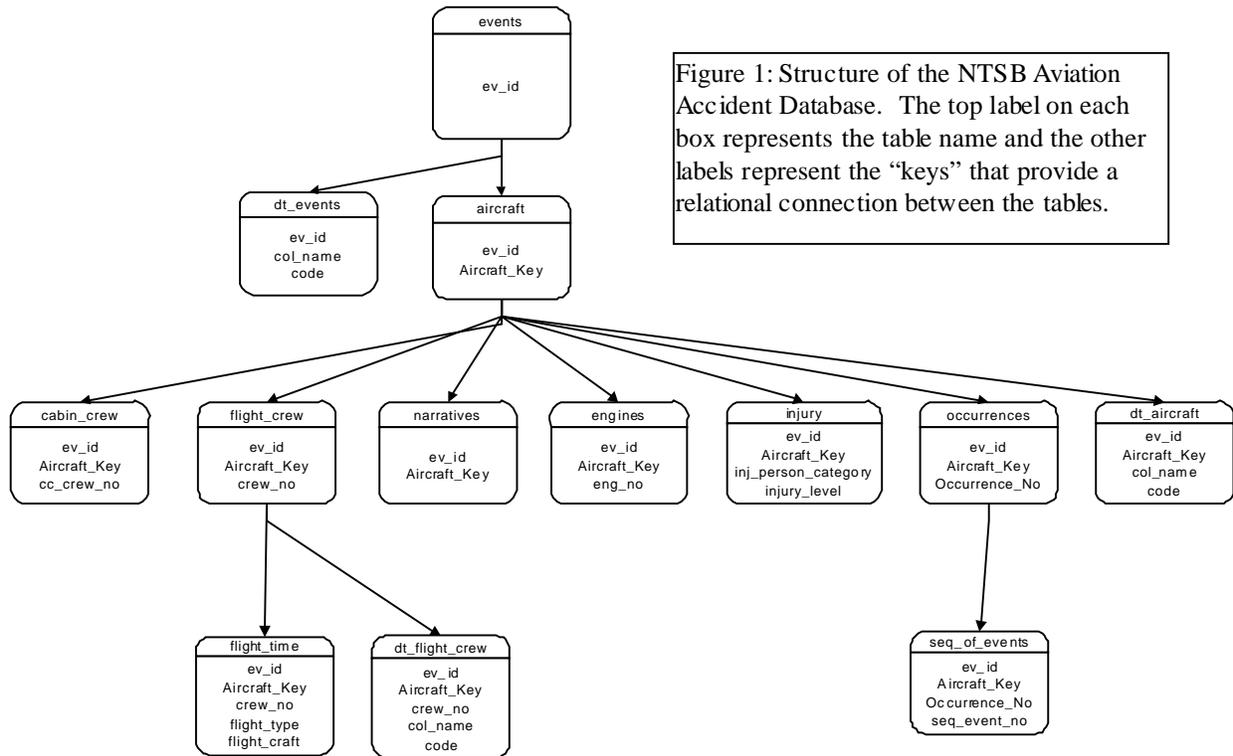


Figure 1: Structure of the NTSB Aviation Accident Database. The top label on each box represents the table name and the other labels represent the “keys” that provide a relational connection between the tables.

Table 1: Data tables available to the public in the NTSB aviation accident database

Table	Definition
<b>aircraft</b>	Pertains to individual aircraft that were involved in an event (e.g., make, model, operator).
<b>cabin_crew</b>	Pertains to members of the cabin crew involved in an event (e.g., injury level, use of seatbelts).
<b>dt_aircraft</b>	Contains several aircraft-related subcategories (e.g., types of landing gear).
<b>dt_events</b>	Contains event-related subcategories (e.g., property damage, sources of weather briefings).
<b>dt_flight_crew</b>	Contains several flight crew-related subcategories (e.g., pilot ratings).
<b>engines</b>	Pertains to engines on aircraft that were involved in an event (e.g., manufacturer, power units).
<b>events</b>	Pertains to the accident event (e.g., date, time, location).
<b>flight_crew</b>	Pertains to members of the flight crew involved in an event (e.g., crewmember role, injury level).
<b>flight_time</b>	Pertains to how much experience (i.e., time-in-flight) individual crewmembers have flying various aircraft in various types of flight.

<b>injury</b>	Includes the number of injured people aboard an aircraft by type of person (e.g., pilot, cabin crew, passenger) and level of injury. (Note: On-ground injuries are located in the events table.)
<b>narratives</b>	Contains narrative text describing the event (e.g., preliminary narrative, probable cause).
<b>occurrences</b>	Pertains to individual occurrences within an event (e.g., type of occurrence, phase of flight).
<b>seq_of_events</b>	Describes what happened during each occurrence using a “sentence structure” that employs person, modifier, and subject codes.

Two tables, *occurrences*, and *seq\_of\_events*, contain information pertaining to **what** happened in the accident, **when** it happened and **why** it happened.

The *occurrence\_code* field in the *occurrences* table defines **what** happened during the accident. A total of 54<sup>1</sup> occurrence codes are used to outline the events of any given accident. Because aviation accidents are rarely limited to a single event, each accident is coded as a sequence of occurrences with as many as five different occurrence codes. Examples of occurrences are loss of engine power, in flight encounter with weather, and midair collision.

The *phase\_of\_flight* field within the *occurrences* table defines **when** each occurrence took place. There are 50 distinct phases of flight investigators use to describe the operational chronology of occurrences, such as standing, taxi-to-takeoff, and emergency descent.

In addition to coding accident occurrences, and phases of flight, accident investigators document explanatory causes, factors and findings to discern the cause-and-effect relationships in the accident sequence or **why** the accident happened. These data are extracted from a series of subject, person, and modifier codes within the *seq\_of\_events* table.<sup>2</sup> An excerpt from the brief report of an accident including occurrences, phases of flight and explanatory causes, factors and findings is presented

<sup>1</sup> Of the 54 codes, two – “missing aircraft” and “undetermined” – do not represent operational events.

<sup>2</sup> The relevant fields are *subj\_code*, *person\_code*, and *modifier\_code*. The *cause\_factor* code documents whether the statement is a cause, factor or finding.

below.

*Occurrence #1: Loss of engine power*

*Phase of Operation: Cruise*

*Findings:*

1. (Cause) Fluid Fuel – contamination, other than water
2. (Cause) Fluid fuel – contamination, water

*Occurrence #2: In flight collision with object*

*Phase of Operation: Emergency descent/landing*

*Findings:*

3. (Factor) Object – trees

*Occurrence #3: In flight collision with terrain*

*Phase of Operation: Emergency landing*

*Findings:*

4. (Factor) Terrain condition – none suitable

Within each occurrence, findings related to the accident sequence are identified. Findings may be further designated as causes or factors of the accident. The term “cause” is used to denote the origin of the accident event, and the term “factor” is used to describe situations or circumstances that contributed to the accident. An example of an accident cause in the preceding example is fuel contamination, which led to an emergency landing. Environmental conditions were cited as factors because the landing was made worse due to trees and unsuitable terrain. Just as accidents often include a series of events, many factors may combine to cause an accident. For this reason, a single accident record can include multiple cause and factor codes. In many cases, an accident record will have a single cause and multiple factors.

A resulting “probable cause” is coded as the combination of all causes, factors, and findings associated with the accident. For the example used previously, the probable cause was, “A loss of engine power while in cruise flight caused by a contaminated fuel supply and the subsequent inflight collision with trees and unsuitable terrain during the

emergency descent and landing.”

#### Accessing the NTSB Accident Database

To download data from the NTSB’s Aviation Accident Database, researchers may use the following website: <<ftp://www.nts.gov/avdata/>>. This site contains self-extracting ZIP files of datasets in Microsoft Access Version 95 (.mdb) format, covering accident information from 1982 to present.<sup>3</sup> One 1982-until-present data set and data from individual years are updated on the first of each month. Other updates occur weekly to add the most current accident data.

To access NTSB accident data without downloading an entire database, researchers may prefer to use a database query tool that allows users to search for sets of accidents using commonly known information, such as date, location, and category of aircraft. The query results in a list of accidents matching the query and links to brief narrative descriptions for individual accidents.<sup>4</sup>

#### Other Sources of Aviation Accident Data

In addition to the Aviation Accident Database, researchers may be interested in the NTSB accident docket, or the recommendations database. The NTSB Docket Management System (DMS) contains factual information related to the accident, such as, pilot reports, photographic or schematic information, and investigative group reports concerning metallurgy, operations, human performance, structures, systems, and others. Work is in progress to provide public access to the DMS via the NTSB website at <<http://www.nts.gov/>>. Automated data files beginning in 1993 are available in the current docket system. Presently, researchers who wish to obtain docket information on a given accident may contact the NTSB Public Inquiries Division at (202) 314-6551.

The NTSB also maintains an archive of safety recommendations that have been issued to the FAA, manufacturers, operators, etc. A searchable database of recommendation letters issued after

September 14, 1998, is currently available at <<http://www.nts.gov/Recs/letters/letters.htm>> and older letters are continuously being added to the database.

#### Accident Data Products

The most easily accessible summary of accident data can be retrieved from the NTSB website at <<http://www.nts.gov/aviation/Stats.htm>> This set of 12 tables is updated annually and includes the total number of accidents by year, accident rates, and fatalities for both commercial and general aviation accidents.

In addition to these tables, the NTSB prepares two annual statistical reviews, the Annual Review of Air Carrier Operations and the Annual Review of General Aviation, which provide a more detailed analysis. These reports and selected datasets from which they were prepared are available as printed publications and on the NTSB website.

---

<sup>3</sup> Data prior to 1982 is maintained in both System 1032 and MSAccess 95. A small number of records for accidents occurring before 1982 appear in ADMS, as a result of information that became available after this database was implemented.

<sup>4</sup>The database query tool is located at <[http://www.nts.gov/nts/query.asp#query\\_start](http://www.nts.gov/nts/query.asp#query_start)>