Flight Data – What Every Investigator Should Know

by

Michael R. Poole
Managing Partner, Flightscape
and
Simon Lie
Associate Technical Fellow, Boeing Air Safety Investigation

Mike Poole is a Professional Engineer with a current pilot's license. He represented Canada as the national expert panel member to the International Civil Aviation Organization's Flight Recorder Panel. He's worked for more than 20 years with the Transportation Safety Board of Canada where he was the head of the flight recorder and performance laboratory and the Flight Recorder Group Chairman on all major accidents, the most recent being Swiss Air 111. Mike is now one of the principles at Flightscape, a flight sciences company specializing in flight data analysis software.

Simon Lie is an Associate Technical Fellow and Air Safety Investigator at Boeing Commercial Airplanes. He has an academic background in composite structures and aero-elasticity, and spent 10 years of his career at Boeing troubleshooting pre-delivery issues on the B747, 767 and 777 aircraft. Simon led Boeing support of the numerous accident investigations including Singapore Airlines 747 Taipei 2000, China Airlines 747 Taiwan 2002, Flash Air 737 Egypt 2004, and Mandala Airlines 737 Indonesia 2005.

Abstract

This paper summarizes issues that every investigator or airline Flight Operations Quality Assurance (FOQA) analyst should know about flight data and, in particular, how the information can be misleading if not fully understood. The process from recovery to final report is presented from a high level perspective, citing actual examples from major international investigations where the data was initially misleading. The correct interpretation of flight data and/or audio data requires a full understanding of the entire signal path from measurement to recorder to investigator. Understanding the underlying principles of the process is paramount to successful interpretation. While not every investigator can have this level of expertise, it is important that they are aware of the types of problems that have been seen in the past but are generally little known outside of the flight recorder laboratory.

Major international investigations generate significant pressure to disseminate information both within the team and externally. Counteracting this pressure is the need to properly validate and understand the flight data. Balancing how much validation is sufficient before disseminating flight data is an important issue for the investigation team. This paper will also demonstrate an example of two flight animations of the same data with differing results and explain why.

Introduction

Flight data is becoming more readily accessible and is increasingly being used for investigation and airline safety programs. Modern aircraft record a huge amount of data compared to just a few years ago, but even in the most advanced aircraft recording systems, significantly less than 1% of the available data is actually recorded. The challenge of analyzing flight data is to recreate an accurate understanding of an event from that small percentage of the available data. To illustrate this challenge, the two photos shown in Figure 1 depict the same scene. The image on the right has been compressed to retain only 1% of the data.

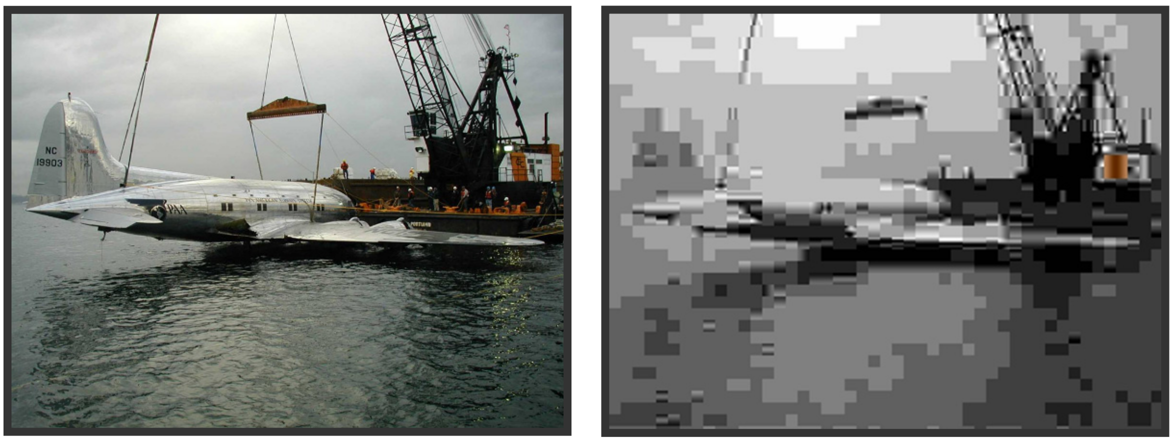


Figure 1 – Recovery of the Boeing S-307 that ditched in Elliott Bay near Seattle, Washington on 28 March 2002. The image on the right has been compressed to 1% of the file size of the original image on the left. The challenge of correctly interpreting the scene on the right is analogous to the challenge of interpreting recorded flight data which is usually much less than 1% of the data available onboard the aircraft.

It is not difficult to imagine that important details might be overlooked or misinterpreted if the image on the right is the only data available. However, given detailed a priori knowledge of the dimensions, weights and capabilities of the airplane, crane, barge and related operating procedures, the right hand image can provide an accurate picture of what happened. Similarly, the analysis of FDR data requires careful attention to detail to avoid misinterpretation; which is why most major investigative authorities consider flight data analysis as a specialty. It is important that investigators and airline FOQA analysts appreciate the 'provenance' of the flight data, especially when drawing substantive conclusions. There is a proliferation of flight data analysis tools which are becoming progressively more automated which in turn increases the potential to mislead. There are many examples where the correct interpretation of an FDR recording requires a full understanding of the provenance including the methods employed by the replay ground station.

The Science of Flight Data

The scientific evaluation of data requires an understanding of the origin of the data and how it was processed. No reputable scientific journal publishes research results unless the origin and history of the data has been reviewed and understood. The same principle applies to recorded flight data. Both of the authors have seen professional investigators reach mistaken conclusions when reviewing recorded flight data without fully understanding the origin and history. As parameters proliferate, even the naming of parameters can lead to confusion. Consider two different parameters that are recorded on certain 737 aircraft: *Selected Fuel Flow* and *Selected Heading*. In the former, the "Selected" indicates that multiple fuel flow readings from different sensors are available and this particular value has been judged to be the most accurate and thus has been selected for display to the flight crew. In the latter, "Selected" means the target value of heading chosen by the flight crew via the autoflight mode control panel. As these two examples demonstrate, scientific rigor requires a full understanding of the origin of flight data had how it was processed.

Provenance

According to the Oxford English Dictionary, *provenance* is 'a record of the ultimate derivation and passage of an item through its various owners'. Adapted for the context of recorded flight data, the definition becomes 'a record of a physical measurement or system state and the changes to that record as it passes through various system components until it is interpreted for an investigation.' Consider the "Selected Heading" example above. The flight crew uses the heading window on the mode control panel (MCP) to choose a heading they wish to fly. The MCP transmits this value to the flight control computer (FCC). The FCC uses the value for computing the correct flight director and autopilot behavior. In addition, the FCC transmits the value to the digital flight data acquisition unit (DFDAU). Continuing to follow the signal chain, we find that the DFDAU stores the values it receives from the FCC until it is scheduled to be written to the FDR. The FDR writes the value to either magnetic tape or (we can hope) solid-state memory as a sequence of ones and zeros. The data is subsequently extracted and converted from raw binary format back into engineering units (i.e., degrees). The converted value is represented as a plot, table, animation, or possibly another format. Finally, the data representation is interpreted by the accident or incident investigator. Figure 2 shows the signal chain in pictoral format.

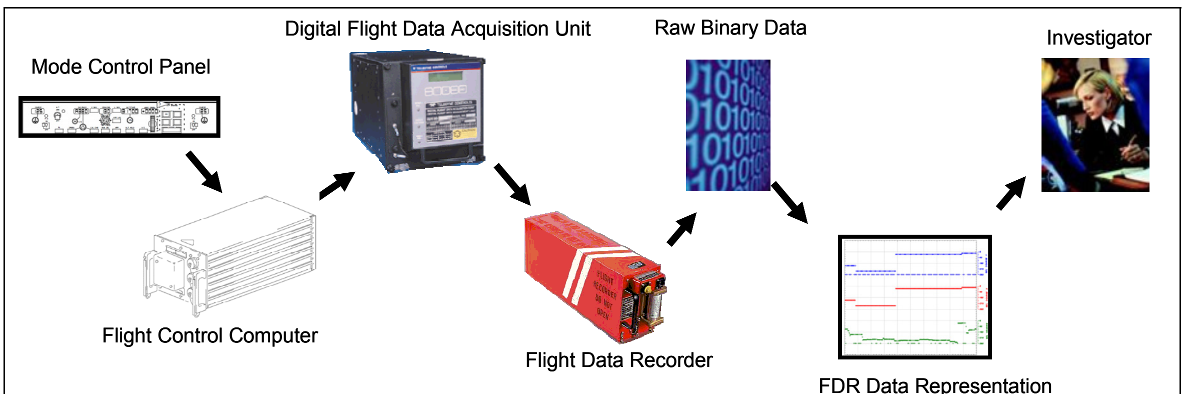


Figure 2 – A graphical depiction of the provenance of the selected heading *parameter recorded on a 737-300 FDR. The data originates in the mode control panel and passes through a number of distinct transformations before being utilized during an investigation.*

In theory, each parameter may have its own unique signal chain. In practice, parameters that have the same source often share the same chain – but not always. At each step of the chain, there is the potential for a change to the signal. Therefore, each step must be fully understood as both intended and unintended changes can affect the results.

B737-700 Example

On January 3, 2004, about 02:45:06 UTC, 04:45:06 Local time, Flash Airlines flight FSH604, a Boeing 737-300, Egyptian registration SU-ZCF, crashed into the Red Sea shortly after takeoff from Sharm el-Sheikh International Airport (SSH) in South Sinai, Egypt. The flight was a passenger charter flight to Charles de Gaulle Airport (CDG), France with a stopover in Cairo International Airport (CAI) for refueling. Flight 604 departed from Sharm el-Sheikh airport with 2 pilots (Captain and First Officer), 1 observer, 4 cabin crew, 6 off-duty crew members and 135 passengers on board. The airplane was destroyed due to impact forces with the Red Sea with no survivors.

The airplane had departed from Sharm el-Sheikh runway 22R and was airborne at 02:42:33 UTC, approximately 2 ½ minutes prior to the crash, and had been cleared for a climbing left turn to intercept the 306 radial from Sharm el-Sheikh VOR station located just north of runway 22R.[1]

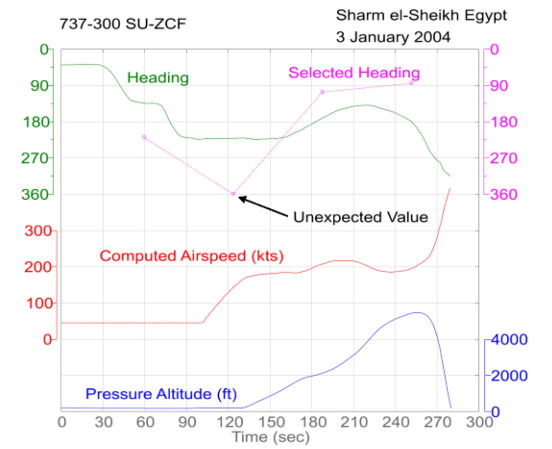


Figure 3 – Altitude, airspeed, heading, and selected heading parameters during the accident flight. Selected heading is recorded once every 64 seconds. The other three parameters are recorded every second. The value at time 123 seconds is unusual as a value equal to the runway heading (220°) would be expected at this phase of flight. The unorthodox arrangement of the heading and selected heading scales (from higher values to lower values) is intentional. By convention, parameter values that result in or result from a right turn tend towards the bottom of the page.

The FDR and CVR were subsequently recovered from a depth of over 1000 meters and provided data used during the investigation. The airplane began the left turn but then rolled out of the left turn and into a right bank that eventually reached 110° right bank. A recovery attempt was made but was not completed before the airplane descended into the Red Sea.

The FDR recorded that the departure was flown with the use of the captain's and first officer's flight directors in heading select mode. In this mode, the flight director provides roll guidance to turn the airplane towards and hold a "selected heading" set by the flight crew on the mode control panel. Accordingly, investigative attention turned to the recorded values of selected heading on the FDR. Figure 3 depicts the airplane heading, selected heading, altitude and airspeed during the accident flight. Heading, computed airspeed and altitude are recorded each second. Selected heading is recorded once every 64 seconds. Standard practice calls for setting the selected heading equal to runway heading during takeoff. At time 59 sec, before the airplane

[1] Flash Airlines Flight 604 Final Report, Egyptian Ministry of Civil Aviation 2006, page 2.

turns onto the runway, the recorded value of selected heading[2] was 220° (runway heading) as expected. At time 123 sec, just prior to rotation, the recorded value was 360°. Later during the flight, the recorded values were to the left of the airplane heading, as would be expected during a left turn. The 360° value was unusual as the expected value would still be runway heading at this point of the takeoff roll. The recorded selected heading data could have indicated an unusual procedure by the flight crew, a malfunction of the mode control panel or flight control computer, or something else. Thus, one focus of the investigation was to understand the actual reason for the unusual reading.

When examining unexpected FDR data, a common practice is to use the entire 25 hour record to determine if the unusual behavior has been present on previous flights. Figure 4 depicts the same four parameters from an earlier flight recorded on the FDR. The recorded values of selected heading generally followed the actual heading (as expected), but there were repeated instances where the two differed and the selected heading was recorded as 360°. During some of the times that the 360° values were recorded, the airplane was flying on a heading of approximately 315° with the autopilot engaged in heading select mode. With the selected heading 45° to the right of the actual heading the airplane would have been expected to begin a right turn towards 360°. However, no such behavior was observed in the recorded data. Figure 5 depicts a portion of the same flight as shown in figure 4 at a different time scale. The unexpected 360° degree values can be seen to alternate with values coincident with the actual airplane heading.

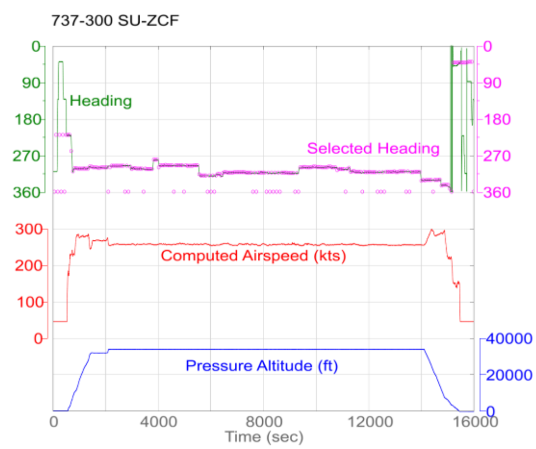


Figure 4 – Altitude, airspeed, heading, and selected heading parameters during an earlier flight. Although selected heading generally follows actual heading as would be expected, there are repeated instances where unexpected values of 360° are recorded.

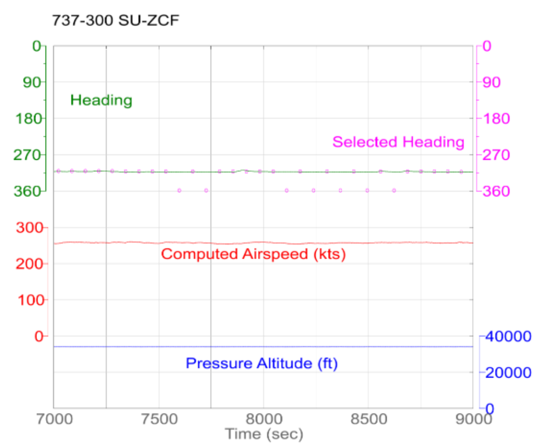


Figure 5 – A portion of the same flight depicted in figure 4 (note change in time scale). The unexpected 360° values of selected heading alternate with values coincident with the actual airplane heading (expected values).

A common practice among DFDAU manufacturers is to use alternating patterns to indicate errors in the FDR data. For example, "stale data", occurs when a source stops transmitting data to the DFDAU or the transmitted data is not received by the DFDAU. Consultation with the manufacturer of the DFDAU confirmed that the

[2] In this article, reported values of selected heading are rounded to the nearest whole degree. Although the resolution of the recording system for this parameter is 0.088°, only whole degrees are selectable on the mode control panel.

alternating pattern observed in the FDR data from the accident flight was an error code indicating "stale data" which originated in the DFDAU. The stale data error code is an alternating sequence of 4095_{10} counts (i.e. 111111111111_2) and the last value received[3]. For selected heading, 4095_{10} counts converts to 360°, therefore the stale data error code consists recorded values of 360° alternating with the last value received. If the inquiry had ended here, one might conclude that the FCC had malfunctioned as evidenced by the apparent lack of selected heading transmission to the DFDAU. Such a conclusion would be incorrect.

In addition to the 25 hours of FDR data available from the accident airplane, the Egyptian MCA provided 25 hours of FDR data from the sister ship. An examination of that data confirmed the same behavior – selected heading occasionally alternated between an expected value and 360°. As shown in Figure 6, the same behavior was also discovered in the selected course #1 parameter[4] on both airplanes but not in the selected course #2 on either airplane. Based on these discoveries, the possibility arose that some sort of design characteristic was responsible for the observed data. Perhaps there was some difference in the way the selected heading and selected course #1 parameters were processed compared to the selected course #2 data that would explain the anomaly.

Figure 6 – Similar to selected heading, unexpected values of 360° were also observed in the selected course #1 parameter but not in selected course #2. All three parameters are recorded once every 64 seconds. FDR data from the accidents airplane's sister ship exhibited the same unexpected values in the selected heading and selected course #1.

Accordingly, the inquiry focused on how the DFDAU detected stale data. According to the DFDAU manufacturer, stale data is detected as follows:

- The DFDAU uses an 8-bit counter to track the number of data samples it has received from the source (in this case the FCC).
- When scheduled to write a value to the FDR, the DFDAU compares the value of the counter to the value of the counter the last time a sample was sent to the FDR.
- If either the counter value or the data value is different, the DFDAU concludes the data is fresh. If both the counter value and the data value are the same, the DFDAU concludes the value is stale. After three consecutive stale

[3] The subscript notation used in this article indicates the radix or base of the number. Thus binary numbers have the subscript 2 and decimal numbers have the subscript 10. By convention, the radix itself is expressed in base 10.

[4] Selected course is the VOR radial selected by the flight crew on the mode control panel. In data originating from the left FCC, selected course #1 is the captain's course and selected course #2 is the first officer's course. In data originating from the right FCC, the two are swapped. The DFDAU can receive data from both FCC, but only records data from a single FCC. A different parameter on the FDR indicates which FCC is the source of the data on the FDR – demonstrating yet another reason to ensure the provenance of the data is well understood.

samples, the DFDAU begins writing the stale data error code until either the counter value or data value change.

Consulting with the FCC manufacturer, it was determined that selected heading and selected course #1 were transmitted to the DFDAU at a rate of 20 Hz. Thus, the DFDAU received selected heading data once every 50 ms and transmitted it once every 64 seconds – a ratio of 1280 to 1. In contrast, selected course #2 was transmitted by the FCC (and received by the DFDAU) at a rate of 10 Hz for a ratio of 640 to 1.

Figure 7 depicts the behavior of the 8-bit counter when utilized for selected course #2 (receive to transmit ratio of 640 to 1). The capacity of an 8-bit counter is 0-255 or 256 distinct values. During normal operation, the 8-bit counter will reach its maximum value and "roll over" back to zero at least twice and possibly three times as the 640 samples are received by the DFDAU between each sample transmitted to the FDR. Regardless of the value of the counter when a sample is transmitted, the counter will be at a different value (that differs by approximately 128 from the previous value) when the next sample is transmitted. The result is that the DFDAU can correctly determine if the data is fresh or stale.

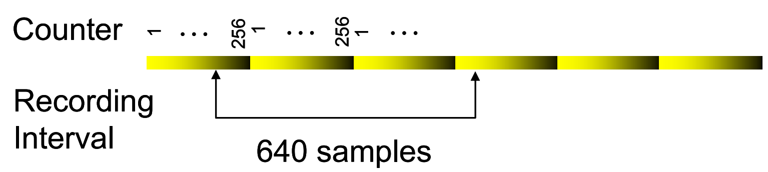


Figure 7. Behavior of an 8-bit counter when the receive-to-transmit interval is 640 to 1 as is the case for selected course #2. The capacity of an 8-bit counter is 256[5]. During normal operation, the counter will "roll over" twice (or possibly three times) between each transmission. Regardless of the value of the counter when a sample is transmitted, the counter will be at a different value (that differs by approximately 128 from the previous value) when the next sample is transmitted with the result that the DFDAU correctly detects that the data is fresh.

Applying the same analysis to selected heading and selected course #1 yields a different result. Figure 8 depicts the situation when the ratio of receive-to-transmit interval is 1280 to 1. In this case, the counter rolls over exactly five times between each transmittal to the FDR. Normal operation will result in the counter value being the same when successive samples are transmitted to the FDR. If the parameter value has not changed, the DFDAU will incorrectly detect that the data is stale, even though the correct number of samples (1280) has been received.

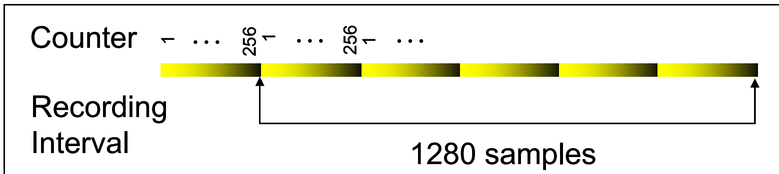


Figure 8. Behavior of an 8-bit counter when the receive-to-transmit interval is 1280 to 1 as is the case for selected heading and selected course #1. The counter will roll over exactly five times between each transmission. The result is that the DFDAU incorrectly detects that the data is stale during normal operation.

---

5 The actual values an 8-bit counter can encode are 0 through 255, a total of 256 distinct values. Software engineers may find fault with figures 7 and 8 because the values portrayed for the counter are 1-256 rather than 0-255. This choice was made for the benefit of that portion of society which starts counting at one rather than zero.

The anomalies in the selected heading and selected course #1 parameters occurred frequently but not in every instance during which the above conditions are met. The last step in the inquiry determined that the exact receive-to-transmission ratio depended upon the relative timing between the FCC internal clock and the DFDAU internal clock, known as jitter[6]. Occasionally, the DFDAU would detect 1279 or 1281 samples instead of 1280 in which case the data would be treated as fresh.

Once the behavior of the stale data detection algorithm was understood, it was a simple matter to correct the FDR data to accurately reflect the selected heading values transmitted by the FCC. The DFDAU will only detect stale data if the parameter value itself is unchanged. Therefore, it was possible to conclude that the selected heading transmitted by the FCC that resulted in the 360° value recorded on the FDR must have been the same as the previously recorded value – 220, the runway heading. The investigation concluded that the anomaly in the stale data detection capability of the DFDAU was responsible for the unexpected value of selected heading recorded on the FDR and that the actual value of selected heading at this time was 220°. The corrected value shown in Figure 9 depicts the data used for the analysis portion of the investigation.

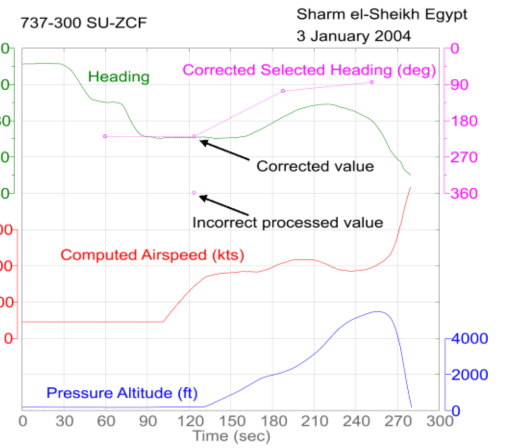


Figure 9 – The same data as depicted in figure 1 corrected to account for the anomaly discovered in the way the DFDAU processed the selected heading data. This was the data used for the analysis portion of the investigation.

As often occurs, this investigation uncovered a finding not related to the accident itself – that the DFDAU did not correctly process data when the receive interval-to-transmit interval ratio was a multiple of 256[7]. A full understanding of the provenance of the FDR data allowed for the correct interpretation of that data for subsequent use in the analysis of the accident.

---

[6] Another instance of jitter is often seen in the recording of UTC seconds. UTC seconds are typically recorded at an interval of four seconds (e.g. successive samples may be 3, 7, 11, 15, 19…). Occasionally, a recorded value will appear to be off by one second resulting in a pattern such as 3, 7, 10, 15, 19….

[7] As a result of this finding, the DFDAU manufacturer was notified of the anomaly so that necessary corrective actions could be taken.

Flight Data Processing

As shown in Figure 10, flight data can be processed and used in two fundamentally different ways. One way is to convert the binary ones and zeros into engineering units data and produce CSV (Comma Separated Variables) files which are similar to what you would see in an Excel spreadsheet. The other, which is rare, is to interactively work with the binary data on a demand basis. The first method (CSV) is the most frequently used in the airline FOQA industry because the CSV format can be interchanged with nearly any computer system and software package. The second method (Binary) requires software applications that have a built-in engineering unit conversion process that allows interaction with the actual raw flight data. The CSV method can introduce several problems that can, in some cases, make it unacceptable for an accident investigation owing to the inability to trace the provenance of the data and high potential to produce misleading results.

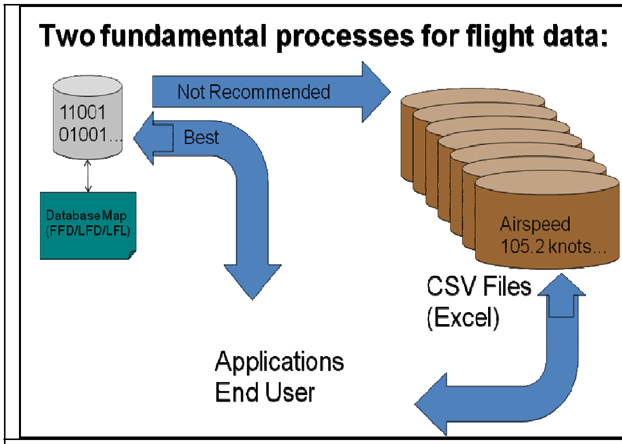
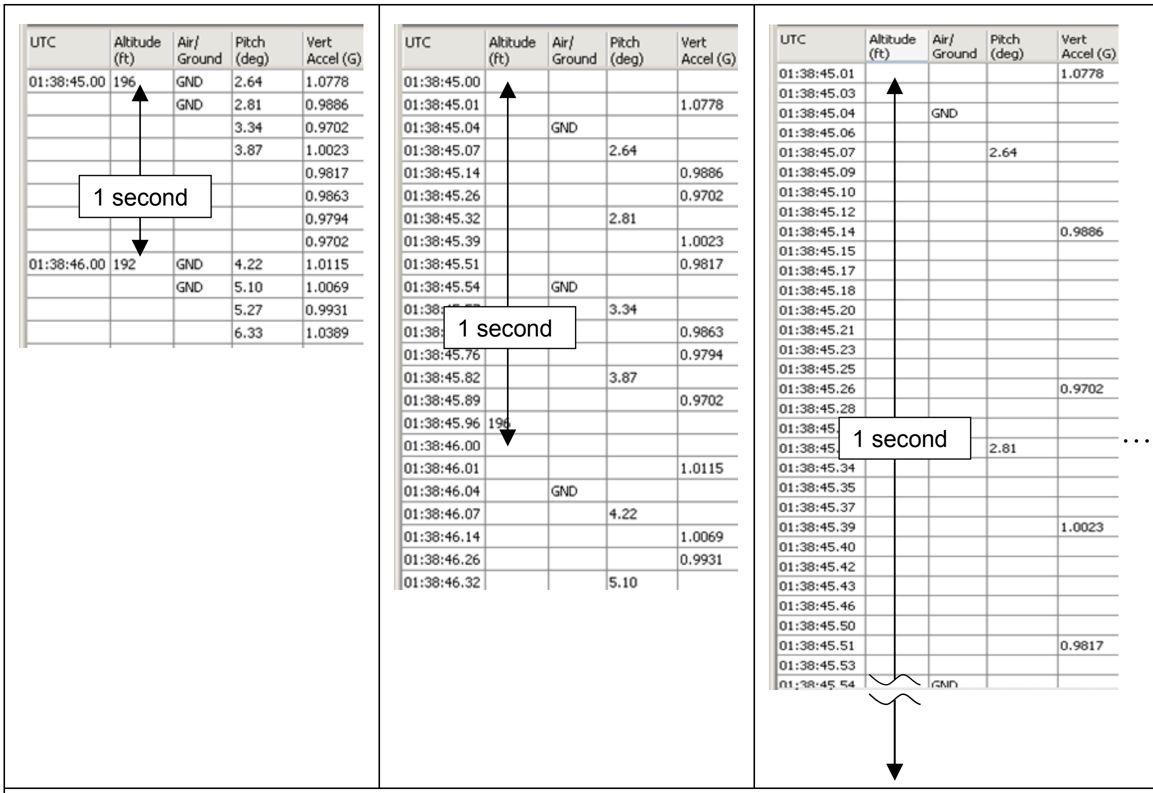


Figure 10 – Flight data must be converted from raw binary format to engineering units for use in plotting and animations. The most common method is to perform the conversion one time and store the results as engineering units, often in CSV files. The conversion process itself can affect the relative timing and precision of data. Unfortunately, the conversion details are almost never provided – thus a portion of the provenance of the data is lost. A better method is to work interactively with the raw data performing explicit conversions only when needed.

By way of example, Figure 11 depicts three CSV files generated with different options from the same source binary file. Each table lists the parameters across the top against time for each line of data. The table on the left discretizes time based on the highest sample rate of the data in the table (in this case the 8 Hz vertical acceleration parameter). This is the most compact and most frequently used method to convert flight data. It is also the most commonly used method to transmit converted data. The center table shows an expanded time base that includes the actual sample time for each entry in the table. This method makes the table larger in length but shows the relative timing between the parameters selected. The table on the right demonstrates what happens when the same principle is used to generate a table that includes parameters from each of the 64 words recorded each second[8].

[8] First generation digital recorders use a recording rate of 64 twelve-bit words per second (wps). Later recorders use 128, 256, 512, and even 1024 wps recording rates. The length of the table on the left side of Figure 3 would grow accordingly.



UTC	Altitude (ft)	Air/ Ground	Pitch (deg)	Vert Accel (G)
01:38:45.00	196	GND	2.64	1.0778
		GND	2.81	0.9886
			3.34	0.9702
			3.87	1.0023
				0.9817
				0.9863
				0.9794
				0.9702
01:38:46.00	192	GND	4.22	1.0115
		GND	5.10	1.0069
			5.27	0.9931
			6.33	1.0389

UTC	Altitude (ft)	Air/ Ground	Pitch (deg)	Vert Accel (G)
01:38:45.00				
01:38:45.01				1.0778
01:38:45.04		GND		
01:38:45.07			2.64	
01:38:45.14				0.9886
01:38:45.26				0.9702
01:38:45.32			2.81	
01:38:45.39				1.0023
01:38:45.51				0.9817
01:38:45.54		GND		
01:38:[illegible]			3.34	
01:38:[illegible]				0.9863
01:38:45.76				0.9794
01:38:45.82			3.87	
01:38:45.89				0.9702
01:38:45.96	196			
01:38:46.00				
01:38:46.01				1.0115
01:38:46.04		GND		
01:38:46.07			4.22	
01:38:46.14				1.0069
01:38:46.26				0.9931
01:38:46.32			5.10	

UTC	Altitude (ft)	Air/ Ground	Pitch (deg)	Vert Accel (G)
01:38:45.01				1.0778
01:38:45.03				
01:38:45.04		GND		
01:38:45.06				
01:38:45.07			2.64	
01:38:45.09				
01:38:45.10				
01:38:45.12				
01:38:45.14				0.9886
01:38:45.15				
01:38:45.17				
01:38:45.18				
01:38:45.20				
01:38:45.21				
01:38:45.23				
01:38:45.25				
01:38:45.26				0.9702
01:38:45.28				
01:38:45.[illegible]				
01:38:45.[illegible]			2.81	
01:38:45.34				
01:38:45.35				
01:38:45.37				
01:38:45.39				1.0023
01:38:45.40				
01:38:45.42				
01:38:45.43				
01:38:45.46				
01:38:45.50				
01:38:45.51				0.9817
01:38:45.53				
01:38:45.54		GND		

Figure 11 – Three methods of displaying the same flight data in tabular format. The table on the left list discrete times based on the highest sample rate data in the displayed data (in this case 8 Hz vertical acceleration). It is the most compact format but the relative timing between the different parameters is lost. The center table lists the actual sample time for each observation of the four parameters listed. The table on the left demonstrates how the length can expand as more parameter sample times are added.

From figure 11, it can be seen that the length of the table is growing in size from the left to right. A first generation FDR recording at 64 wps will record 5,760,000 words over 25 hours. The equivalent to the table on the right side of figure 11 would therefore require 5,760,000 lines. To display the entire data set in Microsoft Excel would require 180 spreadsheets due to the 32,000 line limit in Excel, assuming you have enough columns to list all the parameters, which is not usually possible. A 512 wps QAR recording for 25 hours would require 1440 Excel spreadsheets. Some QAR have up to 400 hour capacity. Thus it is impractical to replicate the original binary file in a CSV format. As a result most data processed and transmitted in the CSV format often compress the data such that each line of the file represents one second and includes only a subset of the actual recorded parameters.

The CSV method requires a number of compromises. The creator of a CSV file must decide which parameters are to be evaluated by the investigators and which are to be ignored. At best, the investigators are forced to request additional parameters once they begin their work and focus on a particular area (often resulting in a significant delay). At worst, the investigators may not realize that additional parameters were recorded and instead rely solely on the data in the original CSV file. Further, because the CSV file does not include the conversion algorithms themselves, it does not allow the provenance of the data to be checked and verified. For these reasons, Boeing does not accept CSV data from airlines requesting an evaluation of an event. Instead, Boeing requires the original raw binary data.

A320 Example 1 – CSV versus Binary

The following example highlights the problems with accepting and using a processed engineering units CSV file instead of working interactively with the binary data. During the investigation of an A320 tail strike, it was determined that both pilots operated the side sticks simultaneously. When operated simultaneously, the A320 performs an algebraic sum of the two inputs to command the aircraft[9]. Figure 12 shows the flight data for the 4 Hz side stick inputs as CSV data (green line) and the original binary data at their precise sample times (yellow dots). As expected, the green line is shifted slightly to the left by a fraction of a second since the CSV file is truncated to the nearest quarter second. Had this example been done with 1 Hz data, the time shift would be more pronounced. The difference between the two processing methods may seem very minor – surely nothing serious that would affect the analysis. A more thorough analysis of the event reveals the problems introduced by the time shift in the CSV file.

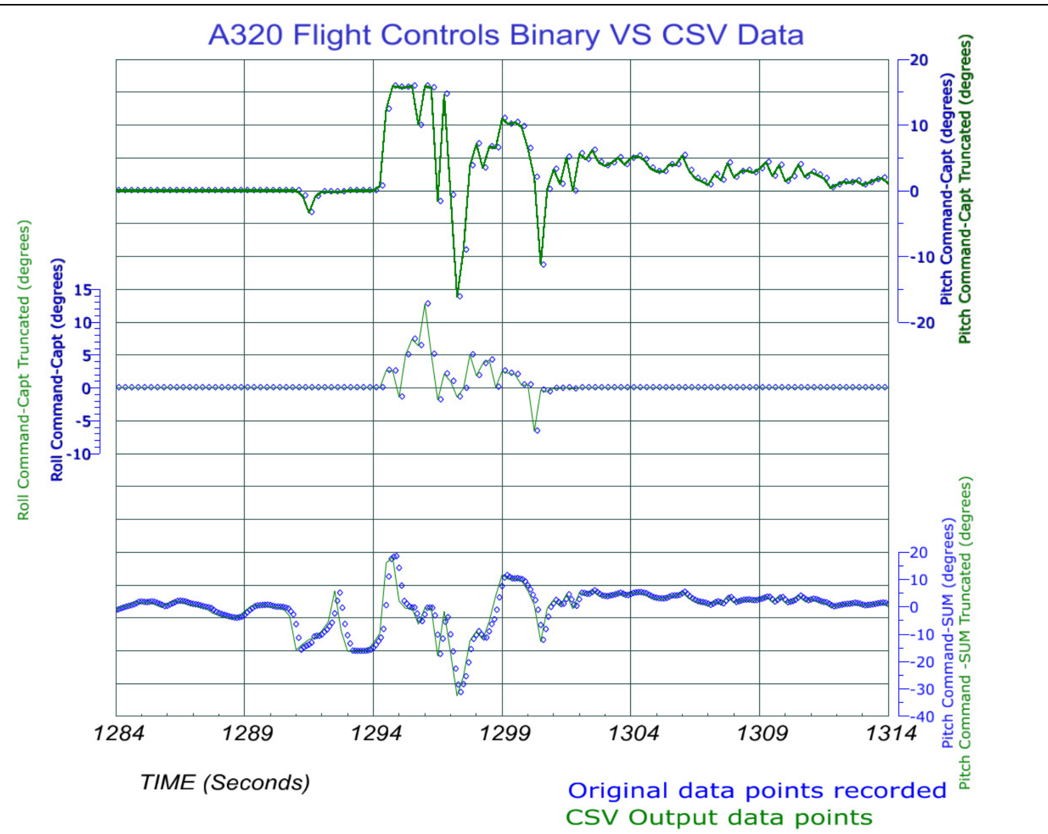


Figure 12 – Comparison of A320 side stick position data recorded at 4 Hz. The original binary data points are shown as blue circles. The same data processed by the CSV method is shown as green lines. Note the time shift as the CSV method positions the 4 Hz data to the nearest ¼ second rather than at the actual time recorded.

[9] The algebraic-sum algorithm is used unless one pilot presses the priority button to eliminate the other side stick position from the control algorithm. Neither priority button was used in this event.

A flight animation was developed using both the CSV data and the binary source data side-by-side to demonstrate the effect of the time shift. Figure 13 contains screenshots showing the recorded side-stick position at various points during the animation. The first officer is the pilot flying. The binary data is shown in the lower left quadrant and the CSV data is shown in the lower right quadrant of each image. The yellow diamond is the captain's input and the green diamond is the first officer's input. When both are operated together, the white hollow diamond shows the algebraic sum. In Figure 13, the error between CSV and binary becomes dramatic. At key points in the animation, the behaviour of the side-sticks, and in particular the first officer's pitch inputs, is considerably different. While the binary data still has issues regarding resolution, accuracy, latency and other provenance issues, it is considerably better than the CSV data which has all those issues as well as the time shift resulting from the processing. It is not difficult to imagine that an analysis of the event based solely on the CSV data could result in erroneous conclusions.

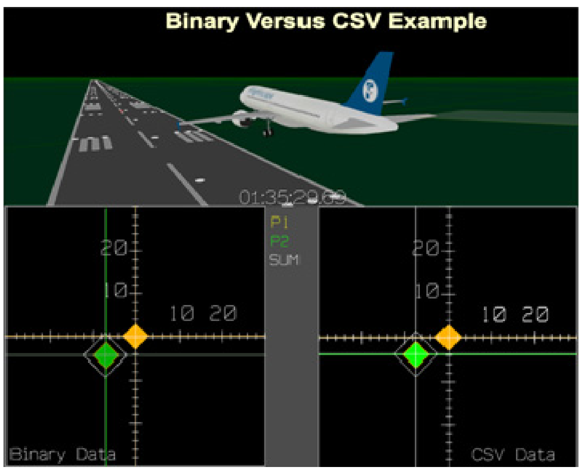
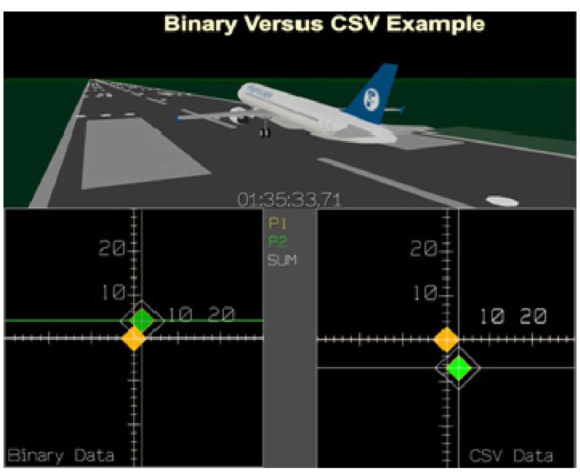
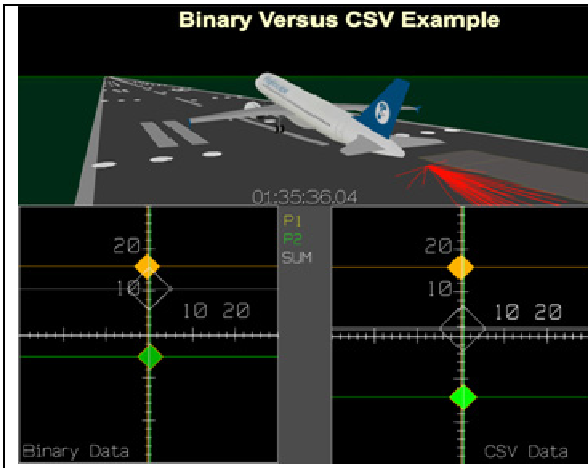


Figure 13 – Still frames taken from an animation of an A320 tail strike event during which both pilots operated their side-stick controllers simultaneously. The lower portion of each screen depicts the location of the side-stick controllers. Position based on the binary data is shown in the left quarter and position based on CSV data is shown the in right quarter. The yellow diamond in the lower portion of each screen depicts the motion of the captain's controller, the green the first officer's, and the white hollow diamond the algebraic sum. The first image is taken approximately 6 seconds prior to the tail strike.



Approximately 3 seconds later the first officer's pitch input appears to be opposite directions in the binary and CSV data. In actuality the apparent reversal is an artifact of the time shift of the CSV data and the frequency of the first officer's inputs.



At the time of the tail strike, differences can be seen in the first officer's pitch input and in the algebraic sum. A performance evaluation of the airplane's response based on the CSV data will not yield the same results as one based on the binary data.

A320 Example 2 – Recovering Data 'Dropouts'

The second A320 example highlights the need to work with the binary data to recover what is known as data 'dropouts'. These are frequent with magnetic tape-based recordings which are susceptible to mechanical vibrations such as occur during turbulence or a crash sequence. Data dropouts can also occur within solid state recordings, primarily due to power interruptions.

On 14 February 1990, an A320 crashed in Bangalore, India with the loss of 92 lives. The investigation was conducted by the High Court of Karnataka, in accordance with Indian laws and regulations. The Court initially decided to have the recorders replayed at a local operator's readout facility but the operator did not have A320's in their fleet at the time and it was unable to playback the FDR. The Court eventually decided to go to the Transportation Safety Board of Canada (TSB) where a request was made for a 'readout' of the data only with no analysis. The TSB took the position that some analysis of the data was necessary in order to ensure that the data produced was correct. As discussed in this paper, the processing of flight data is fraught with the potential for error as well as choices in the manner in which it will be produced. The best way to ensure the data has been processed correctly is to examine and analyze it. If the data appears unusual or unexpected, start by ensuring the signal chain or provenance of the data is fully understood. It is important to understand that readouts using different systems or different settings on the same system can produce different results. As noted above, different results are possible depending on the options selected (for example timing options on the generation of a CSV file).

The aircraft crashed short of the runway in daylight VFR conditions. At the moment of first impact, the initial playback 'lost' one second of data due to vibration or jarring of the magnetic tape-based digital FDR. The initial investigation from the field concluded that a mild initial touchdown on the golf course was followed by a bounce and a second devastating impact. During the bounce, one engine spooled up and the other spooled down in response to the TOGA (take-off go-around) command. Early beliefs were that if both engines had spooled up, it may have been possible to avoid the accident. In fact, some newspapers reported that Canadian experts believed that the engine failed to respond and likely caused the accident (apparently because the readout had occurred in Canada).

The software system originally developed at the TSB is unique in that not only can it interactively work with binary data, it also has an extensive 'wave-bit editor' which enables investigators to study the waveform recorded on the magnetic tape which represents the digital ones and zeroes. Figure 14 is a sample screenshot[10] which shows how the bits are interpreted from the raw waveform recorded on the magnetic tape. The bits are recorded at a constant rate in bit cells of constant length. A zero is encoded as 180° of a sine wave within the allocated time while a one is encoded as 360° in the same interval. As shown on the third line of figure 13, a one followed by a zero was incorrectly interpreted as a single long one. The error is readily detectable by the FDR specialists and can be manually edited using the analysis software.

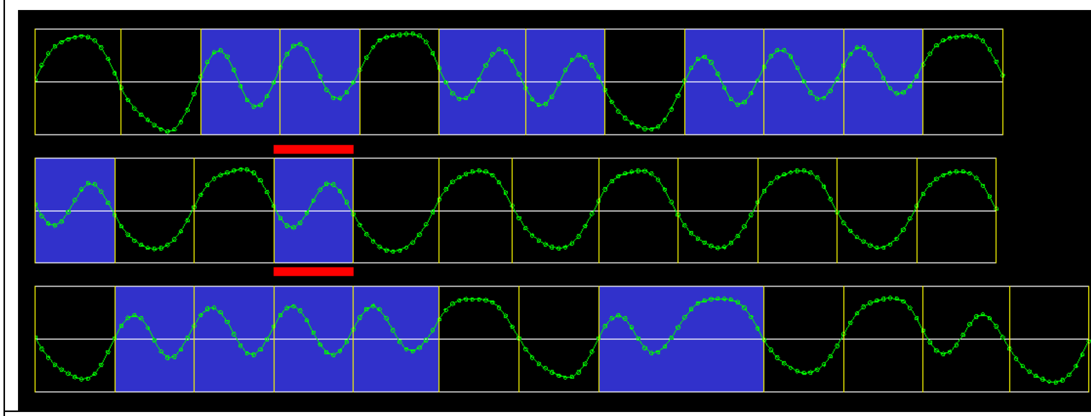


Figure 14 – Digital FDR data recorded on magnetic tape consists of a sinusoidal waveform. A one or zero is encoded as either 360° or 180° of sine wave, respectively, within a fixed-length bit cell. The actual location of the cells is not known and must be deduced from the waveform by the readout software. The variation in bit cell length evident above is the result of tape stretch, variations in tape speed during either recording or playback and other inconvenient realities inherent in mechanical devices. For ease of analysis, the readout system has drawn boxes around deduced location of each bit cell and color coded the background as black or blue representing zero and one bits respectively. The red highlight indicates the position of the "cursor". On the third line of the figure, the large blue block has been incorrectly interpreted as a single one bit rather than a one followed by a zero. When such errors occur, an entire second[11] worth of data is lost as all the subsequent bits are shifted.

The FDR data from the Bangalore accident contained several similar problem areas in the waveform around the time of the first impact. Manual review and correction required a full day to correct the "lost" one second of data. However, once done, the corrected data revealed that the vertical acceleration peaked at over 6 G, the maximum range of the recording. This changed the initial conclusion that the engine may have *caused* the accident to the first contact *being* the accident. In fact, the first contact was so hard it was the 'crash' and caused the engine to spool down due to damage to the electronic controls. Attention to the provenance of the data, specifically the re-evaluation of the conversion from waveform to binary bitstream, redirected the entire focus of the investigation.

---

[10] The example data is not from the Bangalore accident which was an NRZ format recording.

[11] Each second, a predefined synchronization code is written to the FDR. The readout system searches for and uses these "sync codes" to identify the beginning of each second of data thus limiting the effect of missing (or extraneous) bits in the data stream.

It should be noted that readout facilities used for FOQA or maintenance purposes typically do not have any ability to view the waveform and correct errors such as the ones that affected the data from the Bangalore accident. Had the Court gone to an operator for the readout or had the contribution of the TSB been limited to the readout without the necessary analysis to confirm the provenance of the data, it is likely the specific one second of data at the time of first impact would not have been recovered leading to erroneous conclusions[12].

Summary

Every investigator or analyst who uses flight data should know that the correct interpretation of flight data requires a full understanding of the provenance of the data. Each step in the signal chain from measurement to transmission, recording, decoding, conversion, and the final representation can introduce unintended changes and thus the potential for error. The examples discussed above demonstrate unintended changes introduced onboard the accident aircraft and during the subsequent recovery and conversion processes. Determining the provenance of each parameter and understanding the capabilities and replay processes within the analysis software is a necessary step in the interpreting flight data.

---

[12] In part as a result of this accident, ICAO updated Annex 13 which now provides that states which conduct substantial analysis or provide substantial technical support (such as FDR readout) be permitted to appoint an Accredited Representative to the investigation. Annex 13 Appendix D was updated to include guidelines for States without readout facilities that highlights the difference between an ‘airline’ and ‘investigation’ facility and the need to be able to work with and interactively edit FDR source data at the bit level.