DATA REDUCTION

Data Representation

- Big / Little Endian – which bit / byte is first?
  - Most desktop computers have their least significant bit (LSB) as bit 0.
  - Many A/C manufacturers annotate their Interface Control Documents (ICDs) with the LSB as the highest bit in an 8-bit or 16-bit word.
  - Byte ordering in multi-byte words is often reversed with the low byte appearing first.
  - Network byte order is defined to always be big-endian meaning multi-byte numbers are ordered from the most significant byte to the least significant byte.

Data Representation

- Signed Integers typically use two's complement notation to represent positive and negative numbers.
- Final data values must take into account the value of the least significant bit (LSB).
- Units conversion uses the final data value and applies standard methods to obtain the desired result.

Data Representation

```
SIGNAL NAME:....... BARO REFERENCE ALTITUDE
SIGNAL TYPE:....... SERIAL DIGITAL
SOURCE:............
DESTINATION:.......

SIGNAL RANGE:...... -1500 TO +80,000 FEET MSL
TRANSMISSION RATE:. 6.25 PER SECOND

      DATA FORMAT
DATA BIT     DESCRIPTION
 1:....... SIGN BIT
 2:....... MSB (40,960 FEET)
 3:....... *
 4:....... *
 5:....... *
 6:....... *
 7:....... *
 8:....... *
 9:....... *
10:....... *
11:....... *
12:....... *
13:....... *
14:....... *
15:....... *
16:....... LSB (2.5 FEET)
```

Data Representation

- Coded parameters are typically embedded inside a 16- or 32-bit parameter.
- Converting a coded parameter into a meaningful value uses a lookup table to convert the numeric value to the corresponding code.
- ASCII is commonly used for text values encoded into 8-bits.

Data Representation

```
SIGNAL NAME:........ EGI OFP ID
SIGNAL TYPE:........ SERIAL DIGITAL
SOURCE:............. EMBEDDED GPS/INS SYSTEM
DESTINATION:........

SIGNAL RANGE:....... N/A
TRANSMISSION RATE:. 1.5625 PER SECOND
```

```
--------- MSP / WORD 05 --------
DATA BIT     DESCRIPTION
 1:........ MSB
 2:........ *
 3:........ *
 4:........ ** ASCII CODED
 5:........ *    OFP ID (CHAR 1)
 6:........ *
 7:........ *
 8:........ LSB
 9:........ MSB
10:........ *
11:........ *
12:........ ** ASCII CODED
13:........ *    OFP ID (CHAR 2)
14:........ *
15:........ *
16:........ LSB
```

```
--------- LSP / WORD 06 --------
DATA BIT     DESCRIPTION
 1:........ MSB
 2:........ *
 3:........ *
 4:........ ** ASCII CODED
 5:........ *    OFP ID (CHAR 3)
 6:........ *
 7:........ *
 8:........ LSB
 9:........ MSB
10:........ *
11:........ *
12:........ ** ASCII CODED
13:........ *    OFP ID (CHAR 4)
14:........ *
15:........ *
16:........ LSB
```

Data Representation

- Discrete values are typically single bit parameters that indicate some type of status information.
- Converting the bit position from an ICD to computer ordering is key to decoding the proper value.
- Discretes are often converted into coded values (ON / OFF) to aid in presentation.

Data Representation

```
SIGNAL NAME:........ CURRENT LATCHED BIT, WORD 12, WARNING DISCRETES
SIGNAL TYPE:........ SERIAL DIGITAL
SOURCE:............. EMBEDDED GPS/INS SYSTEM
DESTINATION:........

SIGNAL RANGE:....... N/A
TRANSMISSION RATE:. 1 PER SECOND

      DATA FORMAT
DATA BIT     DESCRIPTION
 1:........ SP HISTORY STORE INCOMPLETE
 2:........ GEM DUAL PORT RAM MEM TIMEOUT WARNING
 3:........ SP CPU FUNCTIONAL FAULT
 4:........ GPS READY TIMEOUT WARNING
 5:........ ACCEL BIAS WARNING
 6:........ GYRO BIAS WARNING
 7:........ CLOCK FREQUENCY WARNING
 8:........ BARO BIAS WARNING
 9:........ ALTITUDE DIVERGENCE WARNING
10:........ ACCEL SCALE FACTOR WARNING
11:........ ZEROIZE PERFORMED WARNING
12:........ SYNCHRO REFERENCE (26 VAC) LOST
13:........ ETI VERIFICATION FAIL
14:........ SYNCHRO FAILED TO COMPLETE PROCESSING
15:........ BACKUP POWER ON WARNING
16:........ BATTERY LESS THAN 19 VOLTS WARNING
```

Data Representation

- Engineering unit conversion consists of translating a raw data type into a final value properly scaled and converted into the desired units.
- Angle measurements are often taken in semicircles and must be converted into Degrees (180 Deg / semicircle) to make them meaningful.

Data Representation

```
SIGNAL NAME:....... DISPLAY TRUE HEADING (SELECTED SOLUTION)
SIGNAL TYPE:....... SERIAL DIGITAL
SOURCE:............ EMBEDDED GPS/INS SYSTEM
DESTINATION:.......

SIGNAL RANGE:...... +/- 1 SEMICIRCLE
TRANSMISSION RATE:. 3.125 PER SECOND

      DATA FORMAT
DATA BIT     DESCRIPTION
 1:....... SIGN
 2:....... MSB (1/2 SEMICIRCLE)
 3:....... *
 4:....... *
 5:....... *
 6:....... *
 7:....... *
 8:....... *
 9:....... *
10:....... *
11:....... *
12:....... *
13:....... *
14:....... *
15:....... *
16:....... LSB (1/32,768 SEMICIRCLE)

REMARKS

POSITIVE SENSE IS CLOCKWISE WITH RESPECT TO TRUE NORTH.
```

Data Representation

- Final data types are derived from both the precision and range of the base measurement.
- Care must be taken to properly identify base units (Nautical vs Statute miles) when converting to another unit (6076 ft / nm vs 5280 ft/sm).

Data Representation

SIGNAL NAME:....... DISTANCE TO SELECTED DESTINATION
SIGNAL TYPE:....... SERIAL DIGITAL
SOURCE:............ EMBEDDED GPS/INS SYSTEM
DESTINATION:.......

SIGNAL RANGE:...... 0 TO 6,553.5 NAUTICAL MILES
TRANSMISSION RATE:. 3.125 PER SECOND

DATA FORMAT

DATA BIT	DESCRIPTION
1:........	MSB (3,276.8 NAUTICAL MILES)
2:........	*
3:........	*
4:........	*
5:........	*
6:........	*
7:........	*
8:........	*
9:........	*
10:........	*
11:........	*
12:........	*
13:........	*
14:........	*
15:........	*
16:........	LSB (0.1 NAUTICAL MILES)

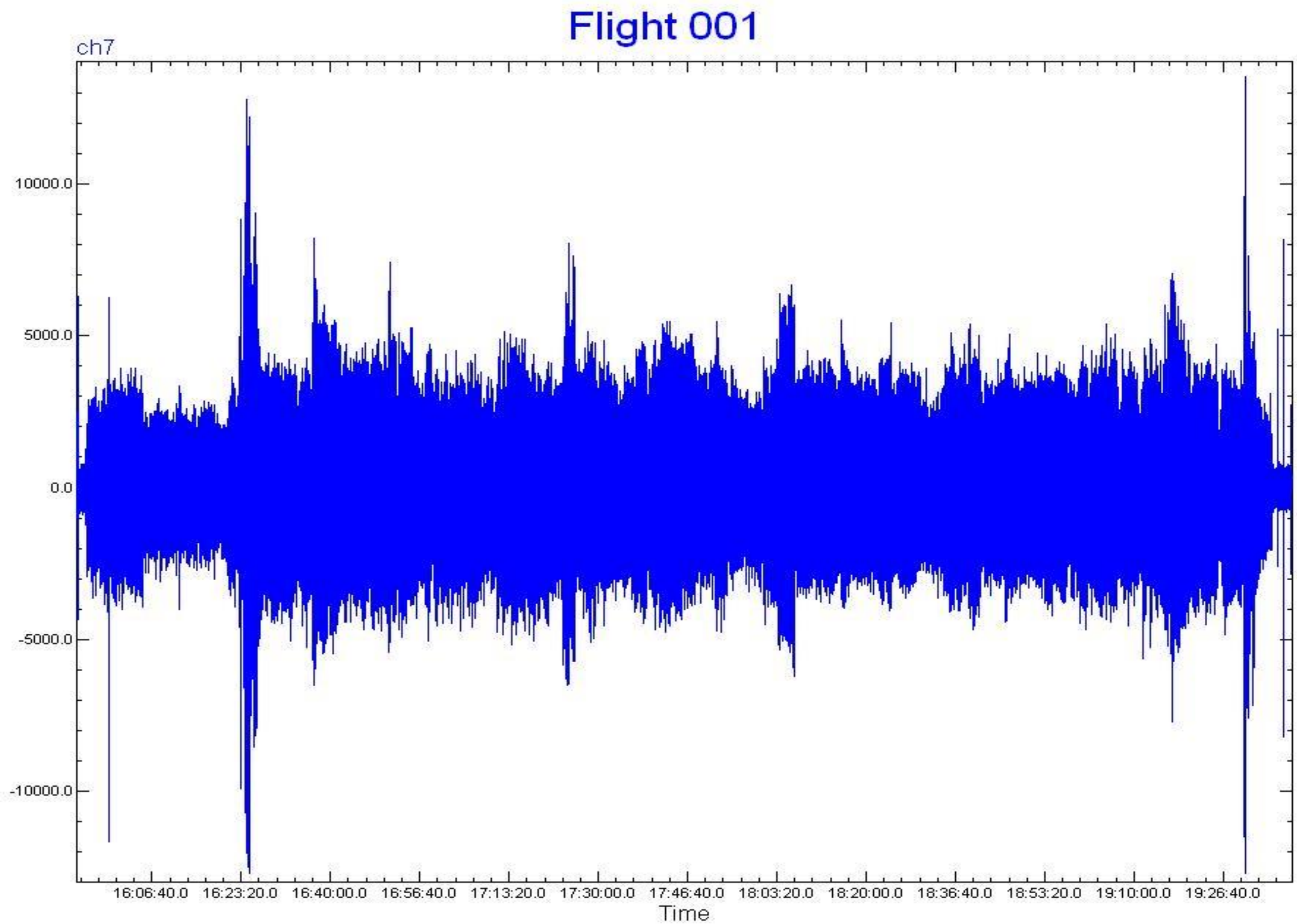
Data Reduction

- Data Reduction is the process whereby you convert raw measured data into meaningful units over a specific time interval.
- Many different techniques are employed in the data reduction process in order to present the results in a form that helps to determine the success or failure of the test.

Data Reduction

- Sampling is the process of selecting a reduced number of data points for the purpose of observing overall trends.
- Care must be taken when sampling to ensure the data is accurately presented and not skewed by the post processing.
- Example: Acoustic data is typically gathered at high rates (50000 sps) to preserve the inherent frequency information. A sampled plot (100 sps) shows overall levels and indicates overall data quality.

Data Reduction



Data Reduction

- Conditional parameters depend on the value in another parameter.
- Example: GPS data is only valid when the receiver has lock on at least three satellites. While position data is typically reported at a constant rate it is only good when the “GPS Valid” bit is set to 1.

Data Reduction

- Trigger conditions are parameters that “trigger” a set of events or values.
- Example: Missile launch data is constantly computed but is only relevant in the time immediately preceding the point at which the pilot squeezes the pickle button. Some data values change meaning depending on the state of a sensor such as the radar.

Data Reduction

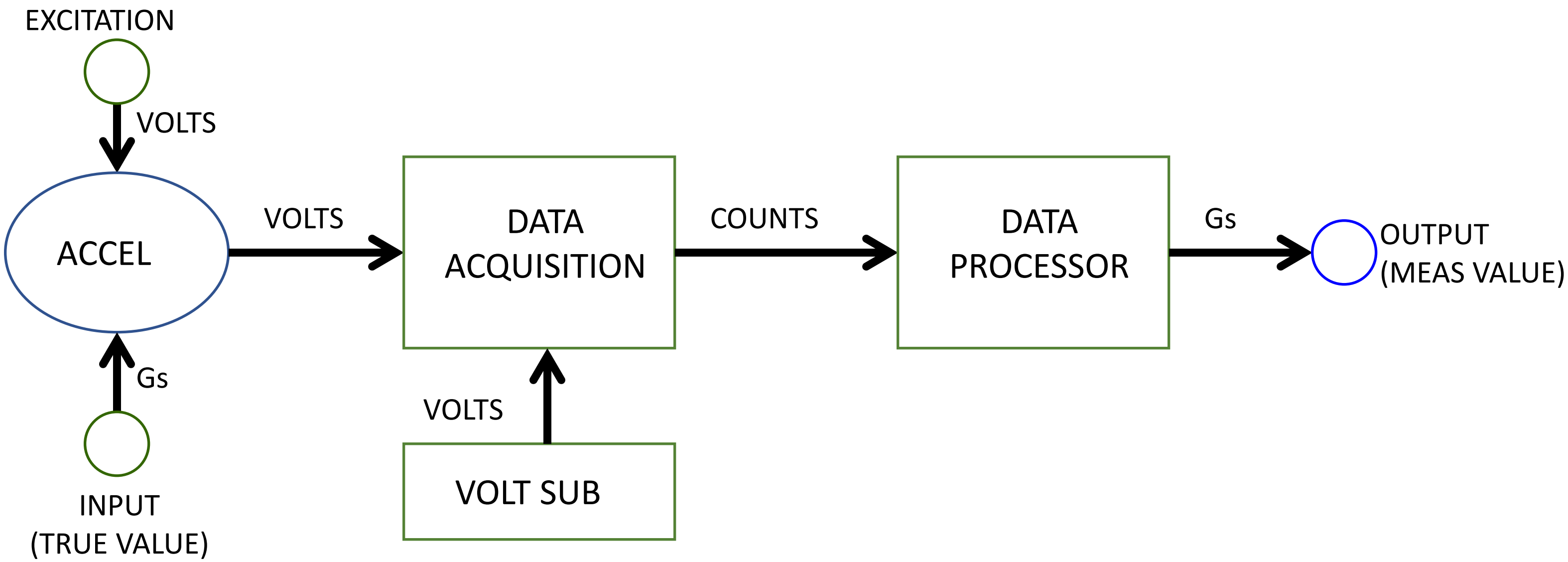
- Data appears to the instrumentation system in either a synchronous or asynchronous fashion.
- Example: Inertial data (position data) appears on the avionics bus at a prescribed rate (synchronous). Target position data only appears when a target has been acquired (asynchronous).

Analysis Software

- Many software programs have the ability to perform data analysis tasks. Typical desktop applications such as Microsoft Excel perform quite well in presenting time ordered data both in tabular and graphical form. Some limitations (65535 rows) prevent Excel from being used with large data sets.

Measurement Uncertainty

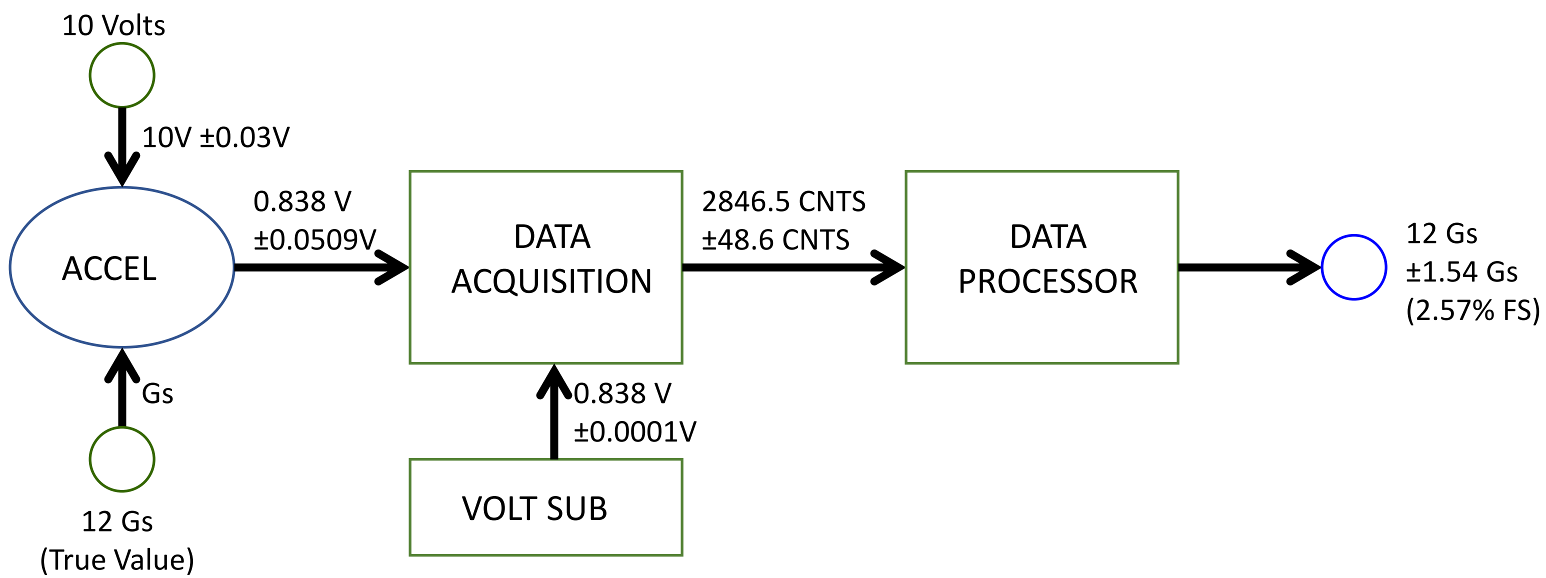
- Now we will look at the uncertainty of an acceleration measurement. The uncertainty describes the bound where the true value of the measurement can be found. In cases where this bound is critical, an uncertainty analysis is performed.



- The methodology for doing an uncertainty analysis can be found in IRIG-122 Uncertainty Analysis Principles and Methods.

Measurement Uncertainty

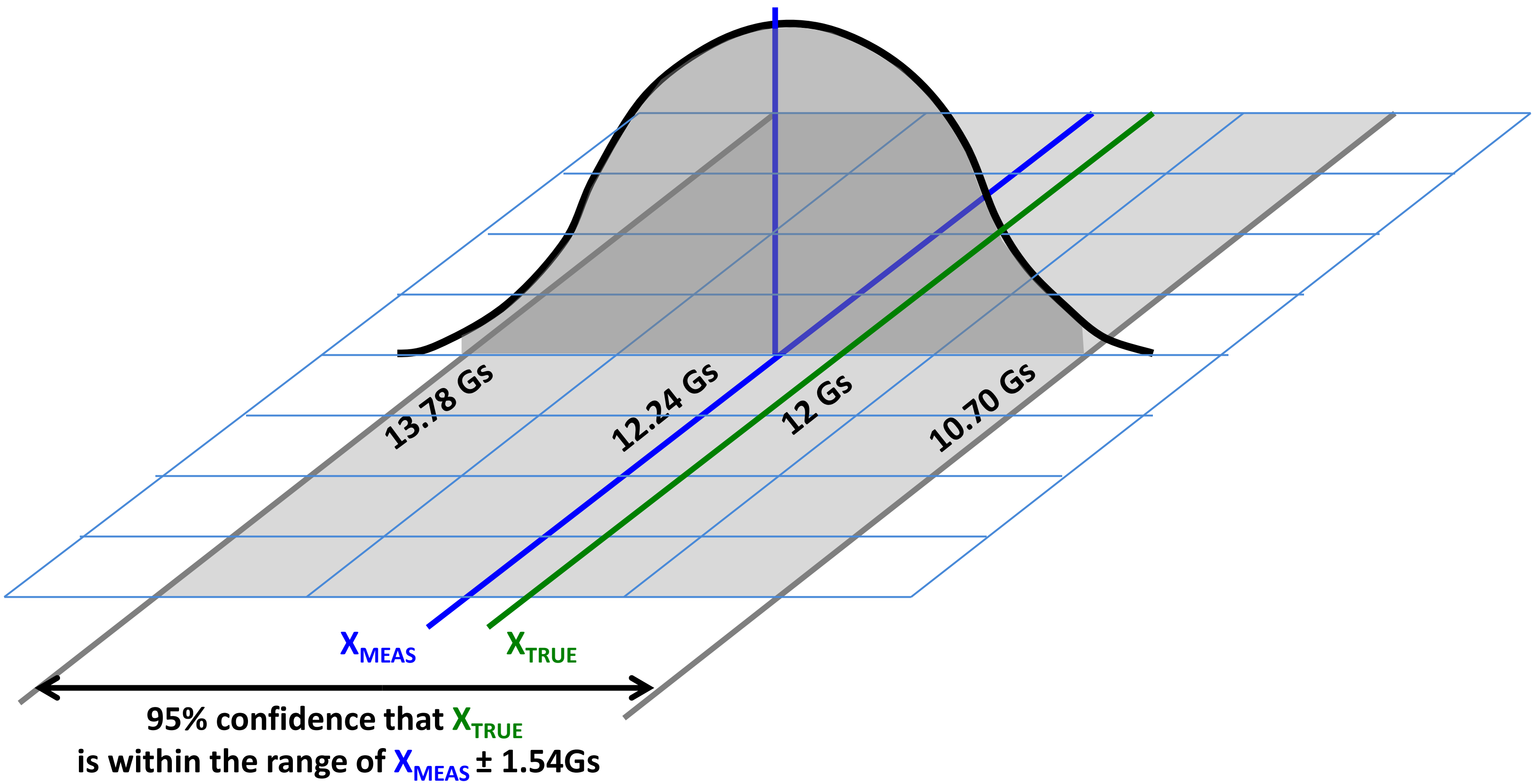
- After the error sources are identified and quantified, the numbers are run through the software which takes care of all the math to properly combine the uncertainties. A simple root-sum-squared method is not used.



- The document describing how the numbers were arrived and the assumptions made for this particular analysis was 14 pages long. The result is a 95% confidence that the true value is within ±1.54 Gs of the measured value (±2.57% FS). The original requirement for the example accelerometer measurement was ±3%.

Measurement Uncertainty

- This graphic depicts the measurement and the unknown true value.

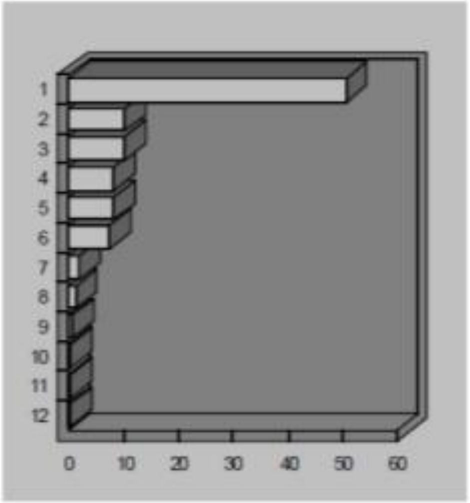


Measurement Uncertainty

- If the 3% uncertainty requirement was exceeded, then the sources of the error would have to be identified and see if it could be reduced. For this example, over 50% of the error in the accelerometer was due to temperature. From there, the large error sources can be identified and improved upon. If 2% was requested, probably another accelerometer would have to be procured with better temperature specs.

Pareto Diagram: (for accelerometer only)

Rank	Error Component	Type	Weight (%)
1	Vostherm	B	50.697
2	Non-linearity and Hysteresis	B	10.139
3	Thermal Sensitivity	B	10.009
4	Magnetic Susceptiblility	B	8.341
5	Transverse Sensitivity	B	8.341
6	Voltage Excitation Sensitivity	B	7.507
7	Voltage offset from excitation	B	1.927
8	Thermal transient error	B	1.251
9	Base Strain	B	0.834
10	Nominal Sensitivity	B	0.400
11	Resolution	B	0.373
12	Residual Noise	B	0.182



Analysis Software

- MATLAB is one of the most commonly used analysis tools in the test community today. Many analysis routines have been written using MATLAB’s programming language.
- http://www.mathworks.com
- Python has replaced MATLAB in many instances due to its popularity in commercial and data science applications.
- Anaconda provides a full-featured Python distribution
- https://anaconda.com

Analysis Software

- ILIAD is a general purpose tool for telemetry programming, pre-flight data validation, real-time display, post-test analysis and archival test information storage.
- http:///www.goiliad.com

Analysis Software

- DPS is a general purpose data reduction tool that converts data from a recorded format (IRIG 106 Chapter 10) into engineering units in an ASCII (comma separated value or tab delimited) format.
- http://software.avtest.com

Analysis Methods

- Determining whether or not an item under test meets specification requirements is a typical Developmental Test (DT) method.
- Analysis methods and algorithms are typically specified prior to the beginning of test as a part of the overall test plan.
- Part of the test preparation process includes the purchase and development of software to support test activities.

Analysis Methods

- Anomaly investigation typically begins when something goes wrong or doesn't work as expected during a test.
- Tools and methods for this type of analysis must include low-level data examination and visualization to track it through the system and understand what happened.

Lessons Learned

- Understand time
  - Much of what happens in the world of instrumentation relies on accurate time tagging to recreate what actually occurred.
  - You must understand the source of your time tag, how your data is time tagged, what happens to your data if the time source fails, how accurately you need to time tag your data.

Lessons Learned

- Understand data rates
  - Many of the measured parameters within a data system are sampled at different rates.
  - Comparing values in a time correlated manner must take this into consideration and adjust accordingly – i.e. you may have to interpolate between two values of a slow parameter to make a valid comparison to a high sample rate one.

Lessons Learned

- Understand conditional parameters
  - Make sure you identify required parameters that are conditional in nature and only use the data when the condition is met.
  - This also applies to asynchronous data that only appears when some specific mode or event happens.

Lessons Learned

- Know your user’s real requirements
  - Creating a plot of every point in a high sample rate signal may not be what the user really wants / needs.
  - Sampling data that with be processed further such as time / frequency domain could skew or alter the final results.

Lessons Learned

- Plan on bad data
  - Very few real tests have 100% perfect data. Plan on having to recover from a lost time source or data gaps.
  - Develop tools that will assist in recreating good data from a corrupted source. Extract usable data and deliver what you can.