412th Test Wing

War-Winning Capabilities … On Time, On Cost

Data Analysis via Open Source Tools

May 2019

Phillip Rogers
Air Force Test Center
661-277-3737

Integrity - Service - Excellence

Overview

- Introduction
- Proprietary Languages
- Open Source Software
- Python
- Example Applications
- References
- Conclusion

Introduction

- Historically, proprietary software languages have been used for performing data analysis and developing data analysis tools.
- Presently, some data analysis is being performed using open source software and more tools are developed using open source software.
- We'll explore some of the reasons for this trend and show some examples of tools developed using open source software.

Overview

- Introduction
- **Proprietary Languages**
- Open Source Software
- Python
- Example Applications
- Conclusion

Proprietary Languages

- Proprietary software languages can offer many benefits.
  - Tailored to a specific domain.
    - Easier to code domain specific tasks
  - Single company manages language
    - Consistent coding style
    - Consistent documentation

Proprietary Languages

- However, proprietary software languages come with drawbacks
  - Vendor lock-in
    - Dependent on the vendor to continue providing product and services
    - Hard to switch to another vendor without incurring additional costs
  - Licensing
    - Annual costs to license the software (typically increase each year)
  - Specialized Libraries/Toolboxes
    - Sometimes incur an additional cost
  - Distribution
    - Tools developed using the Proprietary Language usually require each user to have:
      - A copy/licence of the Proprietary Software Language as well.
      - A copy/license of the same additional libraries
      - The same version of the software

Overview

- Introduction
- Proprietary Languages
- **Open Source Software**
- Python
- Example Applications
- Conclusion

Open Source Software

- Is DoD okay with Open Source Software?
  - "To effectively achieve its missions, the Department of Defense must develop and update its software-based capabilities faster than ever, to anticipate new threats and respond to continuously changing requirements. The use of Open Source Software (OSS) can provide advantages in this regard."
    - DoD CIO Memo "Clarifying Guidance Regarding Open Source Software (OSS)" 16 Oct 2009

- REF: http://dodcio.defense.gov/OpenSourceSoftwareFAQ.aspx

Open Source Software

- Is Open Source Software forbidden by DoD Information Assurance Policy?
  - "DoD has clarified policy on the use of open source software to take advantage of the capabilities available in the Open Source community as long as certain prerequisites are met." Three cases are acceptable:
    - A utility that has publicly available source code is acceptable.
    - A commercial product that incorporates open source software is acceptable because the commercial vendor provides a warranty.
    - Vendor supported open source software is acceptable.

- REF: http://dodcio.defense.gov/OpenSourceSoftwareFAQ.aspx

Overview

- Introduction
- Proprietary Languages
- Open Source Software
- **Python**
- Example Applications
- Conclusion

Python

- What is Python?
  - Developed under an Open Source Initiative (OSI) approved open source license
    - Freely usable and distributable, even for commercial use
  - High-level, general-purpose, interpreted programming language (similar to Java)
  - Thousands of third-party add-on packages
  - Two major code branches
    - Python 3.x - Newest branch, not compatible with 2.x
    - Python 2.x - Final 2.x version 2.7 released in mid-2010, no new major releases

- REF: https://www.python.org/

Python

- Who uses Python?
  - Top engineering schools
    - MIT
      - 6.0001 “Introduction to Computer Science Programming in Python”
    - Stanford
      - CME 193 “Introduction to Scientific Python”
      - CME 211 “Software Development for Scientists and Engineers”
    - CalTech
      - CS 1 “Introduction to Computer Programming”
      - PH 20 “Computational Physics”
    - Carnegie Mellon
      - CMU 15-112: Fundamentals of Programming and Computer Science

- REF: http://catalog.mit.edu/degree-charts/electrical-engineering-computer-science-course-6-2/
- REF: https://exploredegrees.stanford.edu/coursedescriptions/cme/
- REF: http://www.catalog.caltech.edu/current
- REF: https://www.cs.cmu.edu/~112/

Python

- Who uses Python?
  - Top science and research facilities
    - National Weather Service
    - Los Alamos National Laboratory
    - Lawrence Livermore National Laboratories
    - NASA
  - Top commercial graphics companies
    - Industrial Light & Magic
    - Walt Disney Feature Animation
  - Top websites
    - Google
    - YouTube.com

"Python allows us to produce maintainable features in record times, with a minimum of developers."
- Cuong Do, Software Architect, YouTube.com

- REF: https://wiki.python.org/moin/OrganizationsUsingPython
- REF: https://www.python.org/success-stories/

Python

- Why use Python for T&E?
  - Focused on readability and simplicity
  - Mature and capable enough to produce applications for scientific data analysis
  - Extensive add-on packages for technical computing and visualization (similar to MATLAB)
    - SciPy – array and matrix operations, signal processing, optimization, statistics
    - Scikits – image processing and machine learning
    - h5py – HDF5 file I/O
- REF: https://www.python.org/

Python

- Python and Scientific Computing
  - SciPy ecosystem of open-source packages for scientific computing
    - NumPy – numerical array and matrix operations
    - SciPy Library – signal processing, optimization, statistics
    - Matplotlib – publication quality 2D plots
    - pandas – table-based data structures
    - SymPy – symbolic mathematics
    - iPython – interactive interface
    - nose – code testing framework
- REF: http://www.scipy.org/about.html

Python

- Python and Graphical User Interfaces
  - Support for multiple GUI frameworks, including Qt5
  - Qt5 is a multi-platform GUI framework
    - Written in C++
    - All functionality available in Python through the PyQt5 or PySide2 python packages
    - Qt Designer is a tool that allows you to compose and customize widgets/dialogs via drag and drop
  - Previous proprietary software languages had some support for designing GUIs
    - But for full support, you had to resort to an underlying language, like Java
- REF: https://wiki.python.org/moin/GuiProgramming
- REF: https://doc.qt.io/qt-5/qtdesigner-manual.html

Python

- Python and Big Data Analytics
  - Apache Spark - Fast general-purpose cluster computing system on top of Apache Hadoop with Python API, modules include:
    - Spark Streaming - Real-time data streams
    - Spark SQL - Structured data and relational queries
    - MLib - Machine learning library
    - GraphX - Graph (i.e., connections via vertices and edges) processing

- REF: https://spark.apache.org/

Python

- Commercial Python distributions provide Python plus a number of add-on packages with simplified installation and no or low cost (similar to Red Hat for Linux)
  - Anaconda
    - Army, Air Force approved
    - Not sure about Navy approval
  - Enthought Canopy
    - Previous Army/Air Force approval
    - Not sure Navy approval

Python

- Where to get Python
  - Python
    - https://www.python.org/
  - Anaconda (Python + packages)
    - https://www.anaconda.com/
  - Enthought Canopy (Python + packages)
    - https://www.enthought.com/
  - Others

- DISCLAIMER: Make sure your Python distribution is authorized to run on your network!

Overview

- Introduction
- Proprietary Languages
- Open Source Software
- Python
- **Example Applications**
- Conclusion

Example Applications

- We have developed several applications for various disciplines in Python
  - By porting an existing tool from a proprietary language
  - Or From scratch
- Generic Tools
  - Data Plot & Analysis
    - Generic tool for generate report-ready plots from a signal list.
    - Reads/writes Excel files, outputs plots in several formats
  - Event Detection Tool
    - Generic event detection tool built from a propulsion tool
      - » Used for detecting events during Airstart and Transient maneuvers.
    - Generates plots, data listings and populates a database
- Avionics
  - ADS-B Analysis Tool
    - Avionics tool for generating reports to validate ADS-B transmissions
    - Generates plots and data to populate a Word document

Example Applications

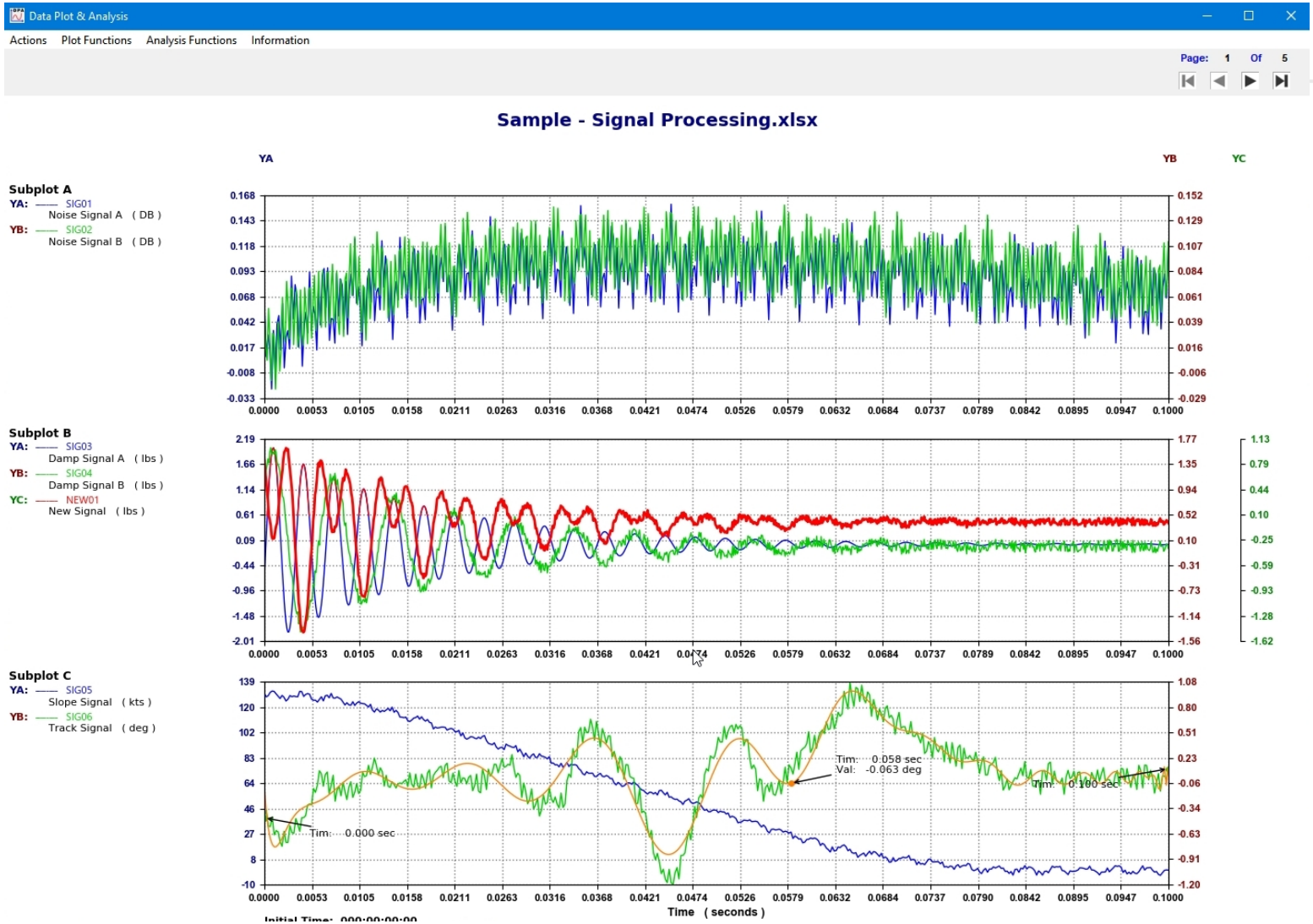
- Performance and Flying Qualities
  - P&FQ Toolbox
    - A library of routines (calculations) that provide the basis of analyses needed by P&FQ engineers
- Propulsion
  - Engine Inlet Compatibility Analysis Program (Inlet AP)
    - Analyzes Pressure and Temperature Probe data from engine inlets
    - Used by 412th TW (Edwards AFB) and AEDC (Arnold AFB)
    - Generates plots, videos, and populates a database
- Structures
  - Structures Analysis Bundle for Evaluation and Reporting (SABER)
    - Provides a single tool for performing Loads, Flutter, and Noise and Vibration analysis
    - Configurable to support any platform.
    - Still in Development

Example Applications

- Data Plot & Analysis
  - Reads a signal list file (xlsx) to generate report-ready plots
    - Handles several input data file formats
    - Capable of producing multiple page plots with multiple subplots and multiple axes per plot.

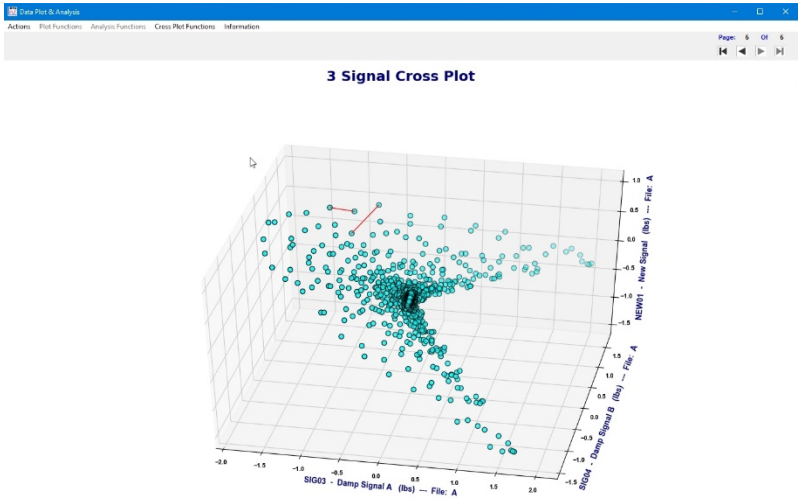
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	No.	ID	Name	Units	File	Active	Plot	SubPlot	Axis	Line Color	Line Style	Line Weight	Point Marker	Maker Size	Number of Markers	Scalar Factor	Bias Value	New Units	Limit Line Overlay	Equation
2	1	SIG01	Noise Signal A	DB	a	y	a	a	a	blue	solid	thin	none							
3	2	SIG02	Noise Signal B	DB	a	y	a	a	b	green	solid	thin	none							
4	3	SIG03	Damp Signal A	lbs	a	y	a	b	a	blue	solid	thin	none						s,>=,0.6,r,s,t,z	
5	4	SIG04	Damp Signal B	lbs	a	y	a	b	b	green	solid	thin	none							
6	5	SIG05	Slope Signal	kts	a	y	a	c	a	blue	solid	thin	none							
7	6	SIG06	Track Signal	deg	a	y	a	c	b	green	solid	thin	none							
8	7	SIG07	Sine Signal	DB	a	y	b	a	a	blue	solid	thin	none							
9	8	SIG08	Spike Signal	%	a	y	b	a	b	green	solid	thin	none							
10	9	SIG09	Ground Speed	kts	a	y,y	c,c	a,b	a,a	blue,blue	solid,n	thin,thin	none,circle	0,5	0,10					
11	10	SIG10	Decel Rate	ft/s^2	a	y,y	c,c	a,c	b,a	red,blue	solid,solid	thin,thin	none,none							
12	11	SIG11	Hydro Press	psig	a	y,y	c,c	a,d	c,a	green,blue	solid,solid	thin,n	none,plus	0,10	0,20					
13	12	SIG12	Random A	ft	a	y	d	a	a	b	s	t	n							
14	13	SIG13	Random B	ft	a	y	d	a	a	r	s	t	n							
15	14	SIG14	Random C	ft	a	y	d	a	a	g	s	t	n							
16	15	NEW01	New Signal	lbs	a	y	a	b	c,a	r	s	m	n							(S0004-S0003)/2

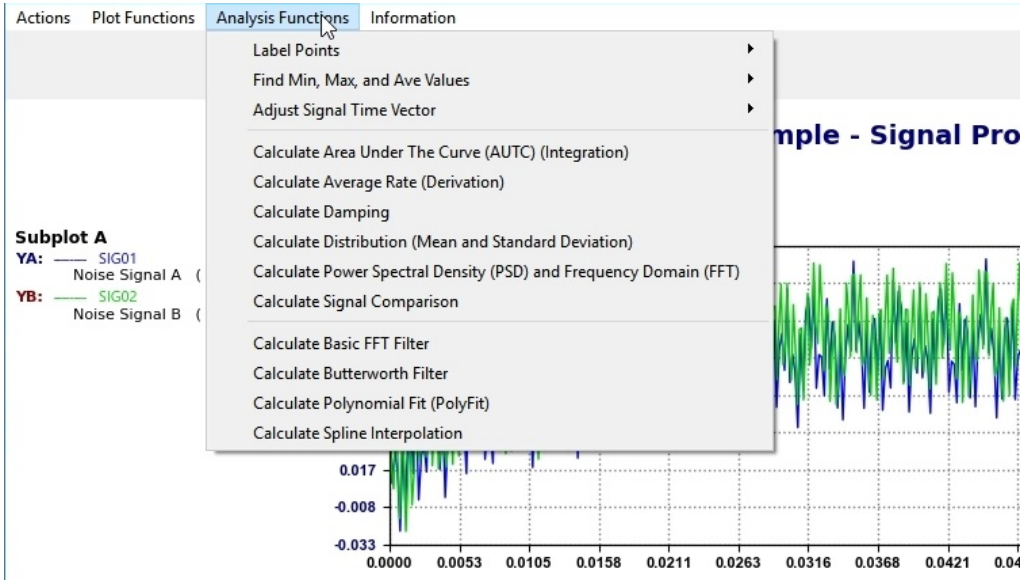
Example Applications

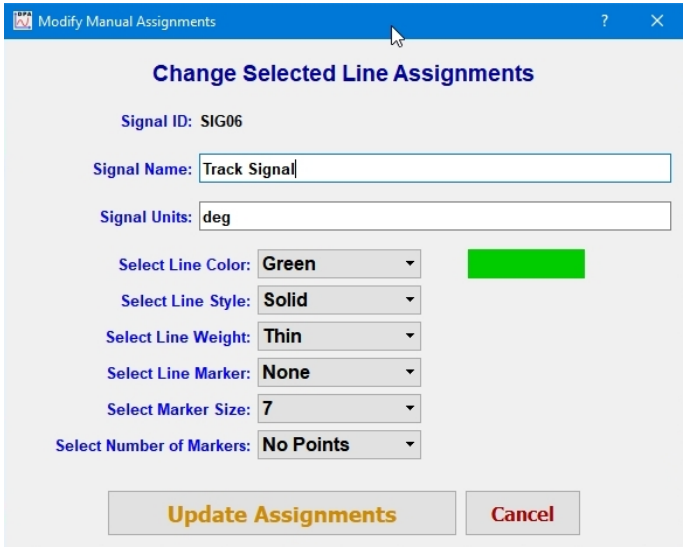


Example Applications

- Data Plot & Analysis
  - Built-in Data and Analysis tools
  - Plot and line customization options
  - 3D Cross plot support

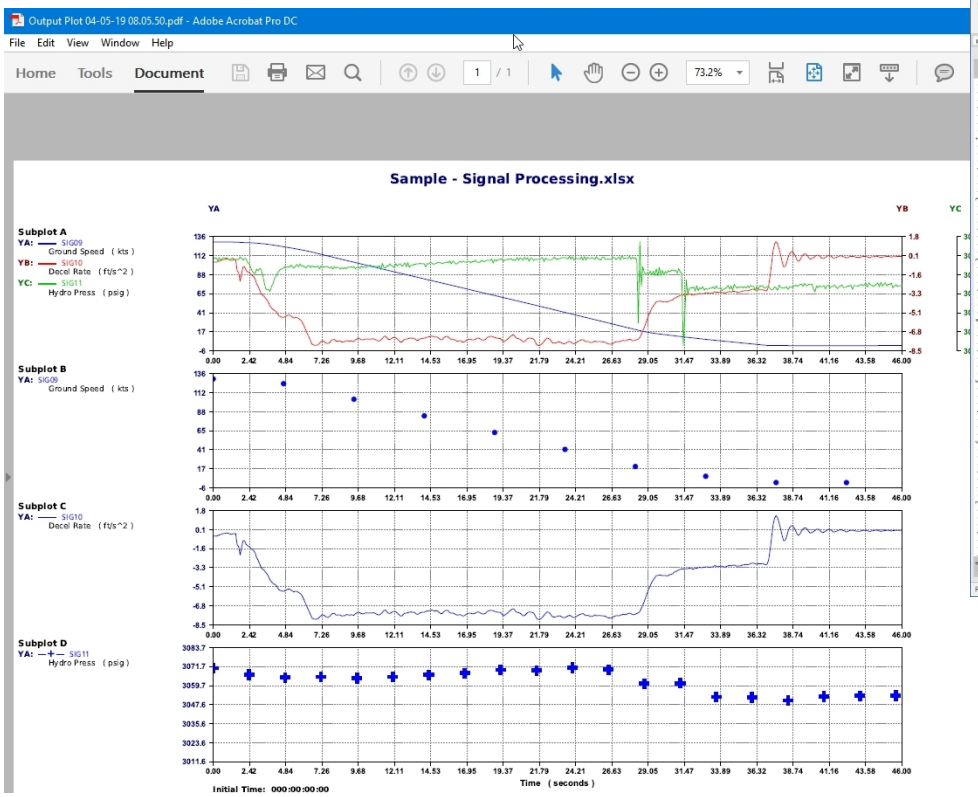


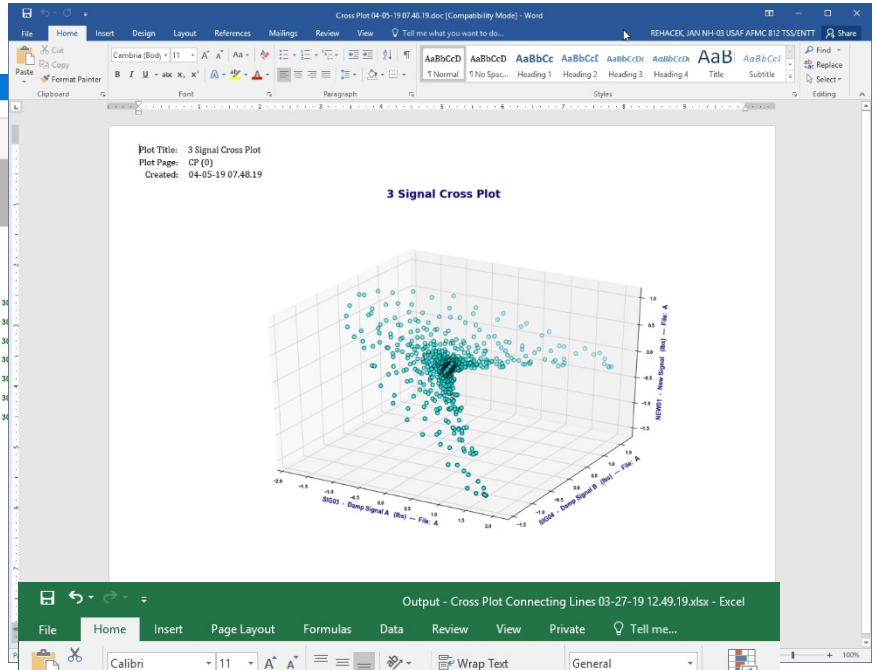


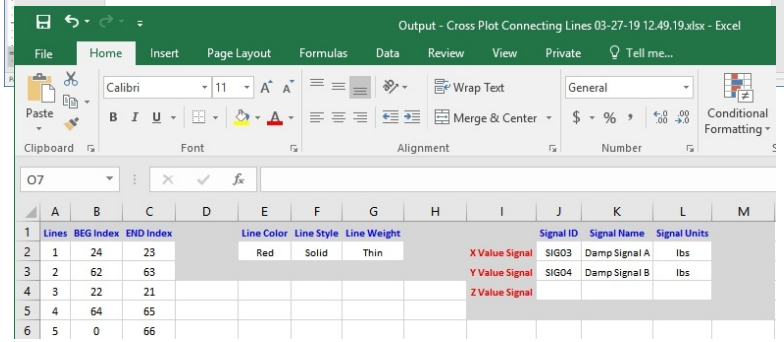


Example Applications

- Data Plot & Analysis
  - Outputs to Word, PDF, and several picture file formats
  - Outputs settings to Excel

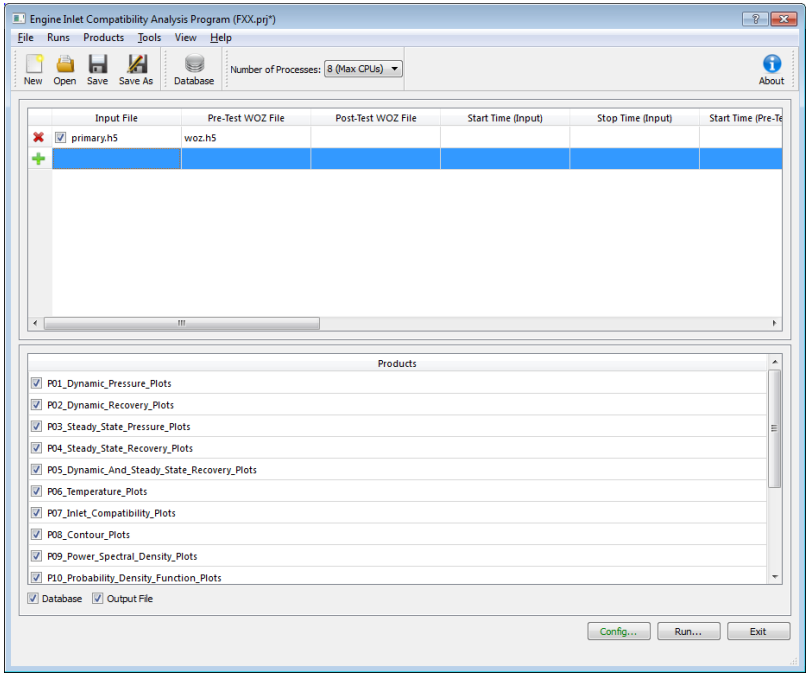


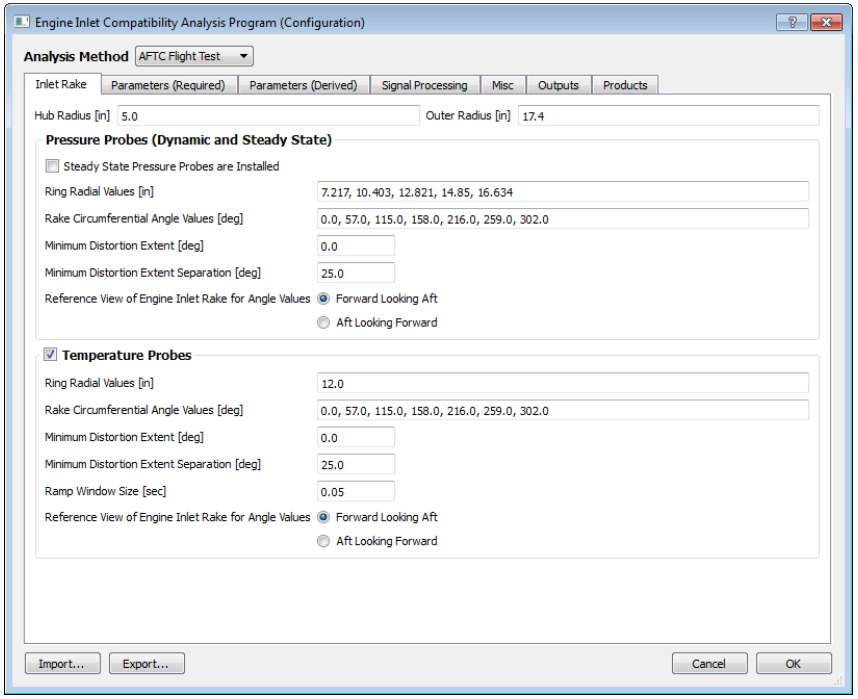




Example Applications

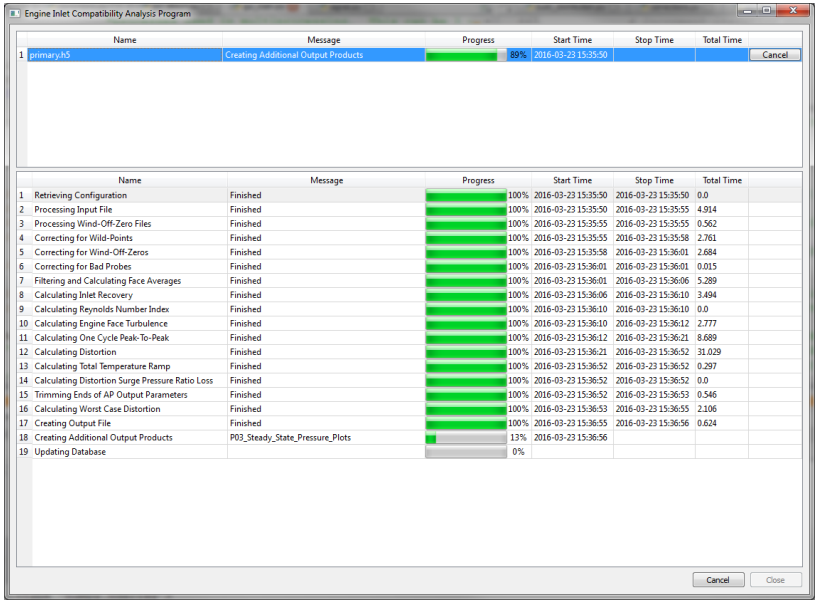
- Inlet AP
  - Configurable to support any platform
  - Capable of analyzing multiple test runs, producing products (plots, videos, etc.), and databasing results
  - Used at Edwards and Arnold

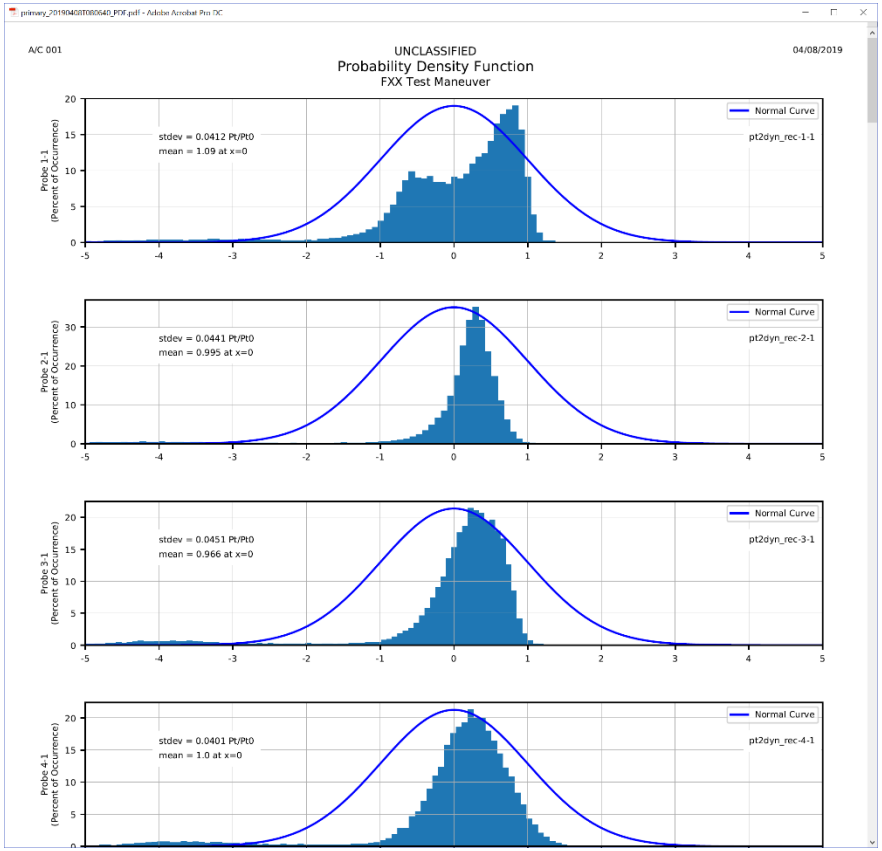




Example Applications

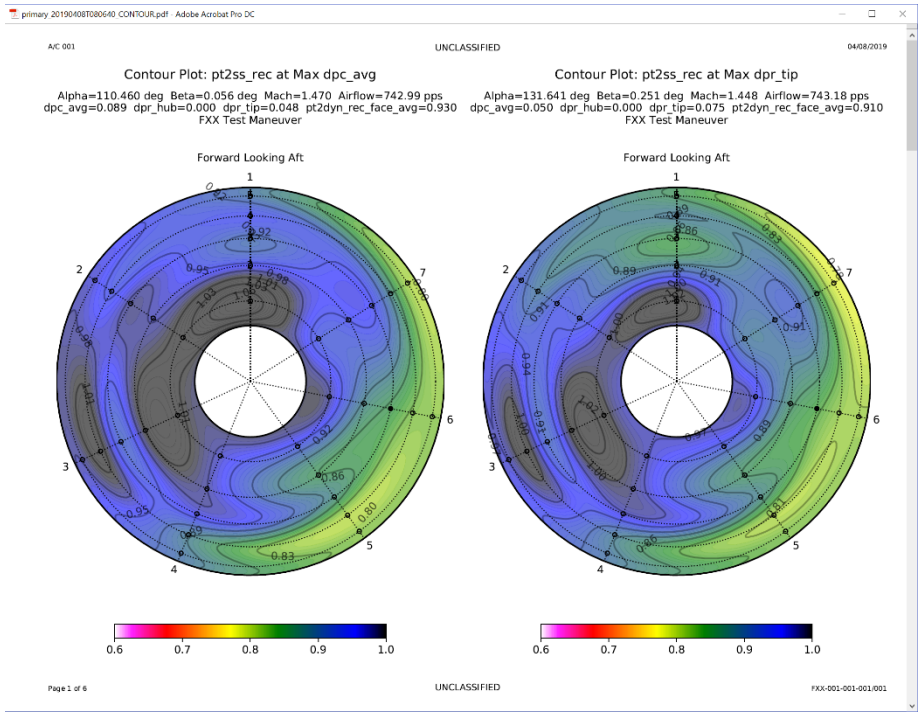
- Inlet AP
  - Batch analysis can be run on multiple CPU cores to streamline the process.
    - Progress can be tracked on each batch process

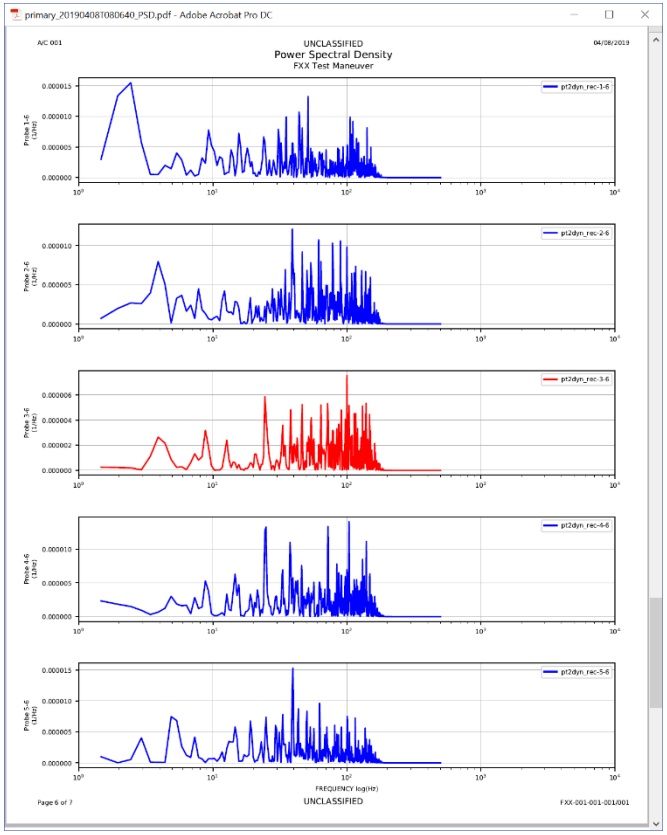




Example Applications

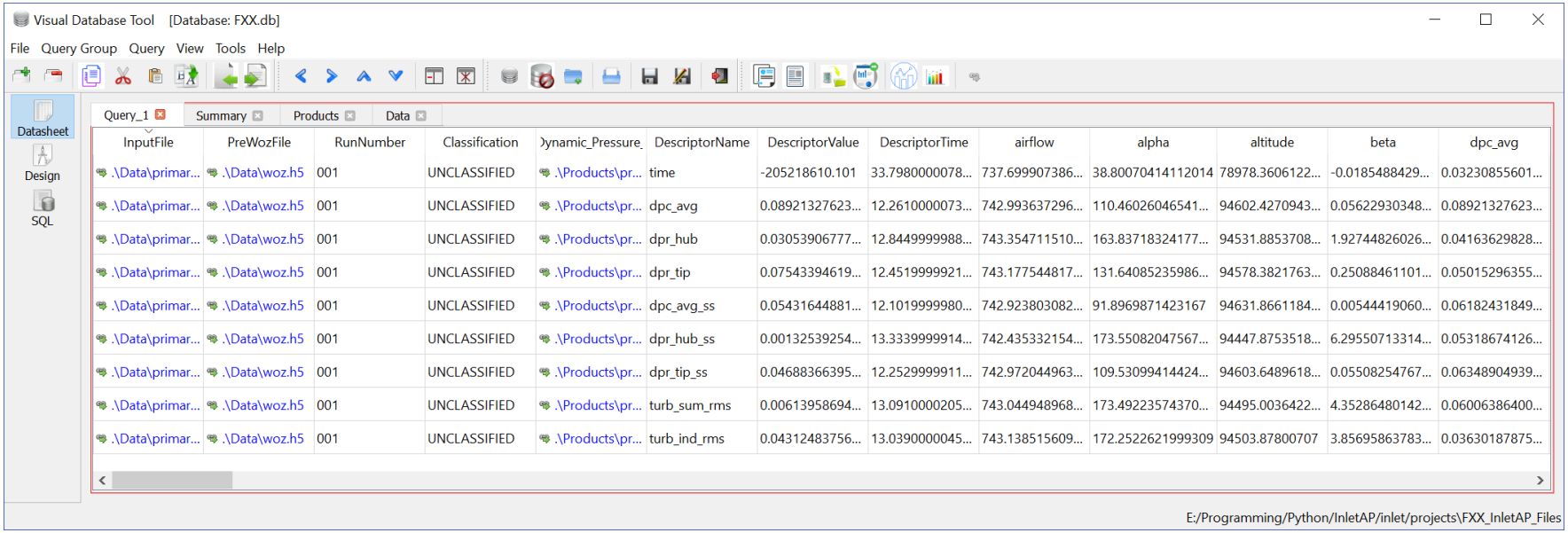
- Inlet AP
  - Wide range of customizable plots and videos
    - Supports marking bad probes on plots





Example Applications

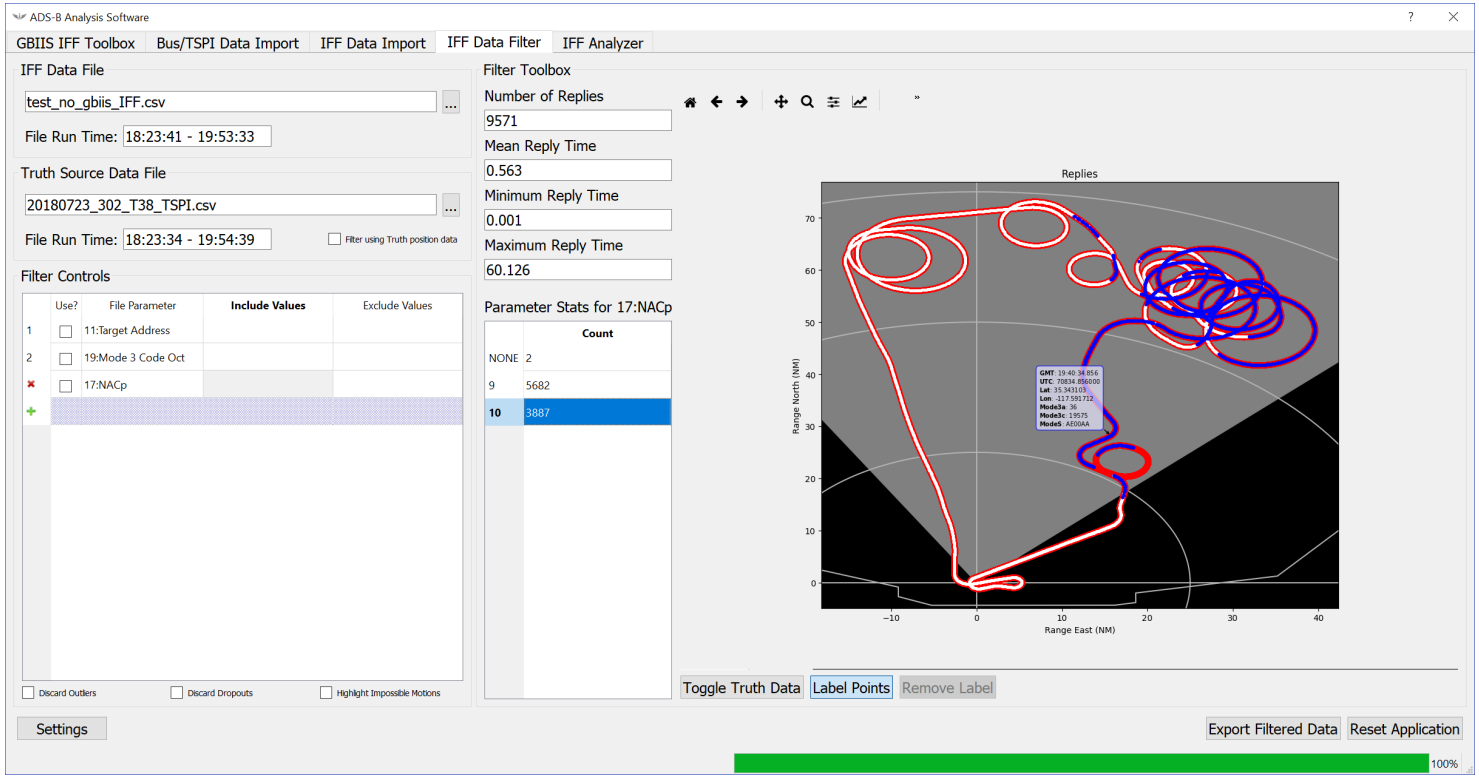
- Inlet AP
  - Capability to produce a customizable database
    - View database with hyperlinks to input/output data and products created.



InputFile	PreWozFile	RunNumber	Classification	Dynamic_Pressure	DescriptorName	DescriptorValue	DescriptorTime	airflow	alpha	altitude	beta	dpc_avg
.\Data\primar...	.\Data\woz.h5	001	UNCLASSIFIED	.\Products\pr...	time	-205218610.101	33.7980000078...	737.699907386...	38.80070414112014	78978.3606122...	-0.0185488429...	0.03230855601...
.\Data\primar...	.\Data\woz.h5	001	UNCLASSIFIED	.\Products\pr...	dpc_avg	0.08921327623...	12.2610000073...	742.993637296...	110.46026046541...	94602.4270943...	0.05622930348...	0.08921327623...
.\Data\primar...	.\Data\woz.h5	001	UNCLASSIFIED	.\Products\pr...	dpr_hub	0.03053906777...	12.8449999988...	743.354711510...	163.83718324177...	94531.8853708...	1.92744826026...	0.04163629828...
.\Data\primar...	.\Data\woz.h5	001	UNCLASSIFIED	.\Products\pr...	dpr_tip	0.07543394619...	12.4519999921...	743.177544817...	131.64085235986...	94578.3821763...	0.25088461101...	0.05015296355...
.\Data\primar...	.\Data\woz.h5	001	UNCLASSIFIED	.\Products\pr...	dpc_avg_ss	0.05431644881...	12.1019999980...	742.923803082...	91.8969871423167	94631.8661184...	0.00544419060...	0.06182431849...
.\Data\primar...	.\Data\woz.h5	001	UNCLASSIFIED	.\Products\pr...	dpr_hub_ss	0.00132539254...	13.3339999914...	742.435332154...	173.55082047567...	94447.8753518...	6.29550713314...	0.05318674126...
.\Data\primar...	.\Data\woz.h5	001	UNCLASSIFIED	.\Products\pr...	dpr_tip_ss	0.04688366395...	12.2529999911...	742.972044963...	109.53099414424...	94603.6489618...	0.05508254767...	0.06348904939...
.\Data\primar...	.\Data\woz.h5	001	UNCLASSIFIED	.\Products\pr...	turb_sum_rms	0.00613958694...	13.0910000205...	743.044948968...	173.49223574370...	94495.0036422...	4.35286480142...	0.06006386400...
.\Data\primar...	.\Data\woz.h5	001	UNCLASSIFIED	.\Products\pr...	turb_ind_rms	0.04312483756...	13.0390000045...	743.138515609...	172.2522621999309	94503.87800707	3.85695863783...	0.03630187875...

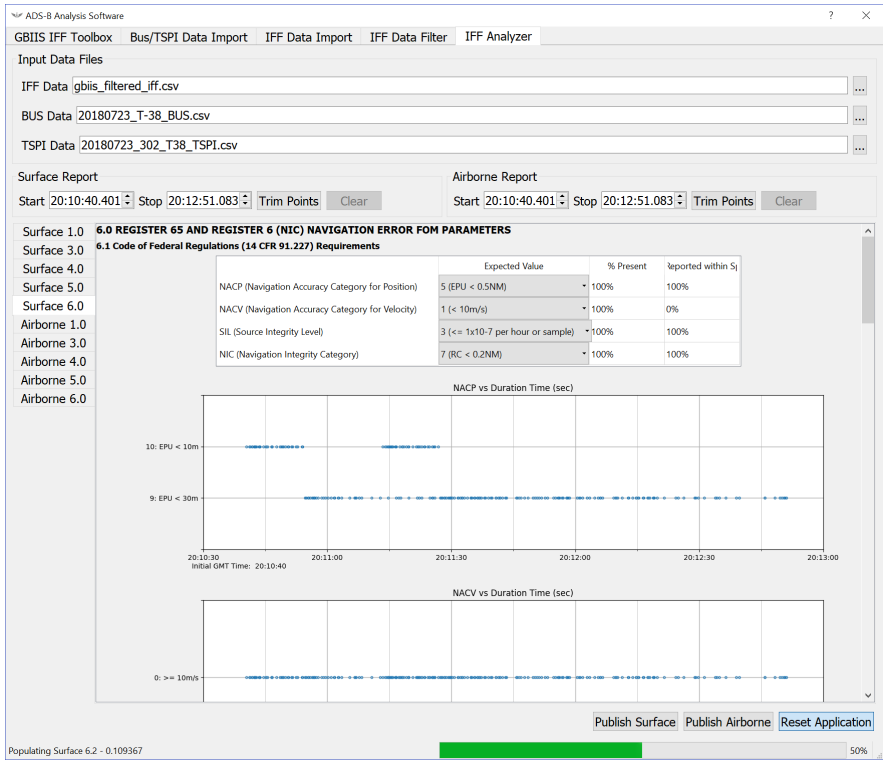
Example Applications

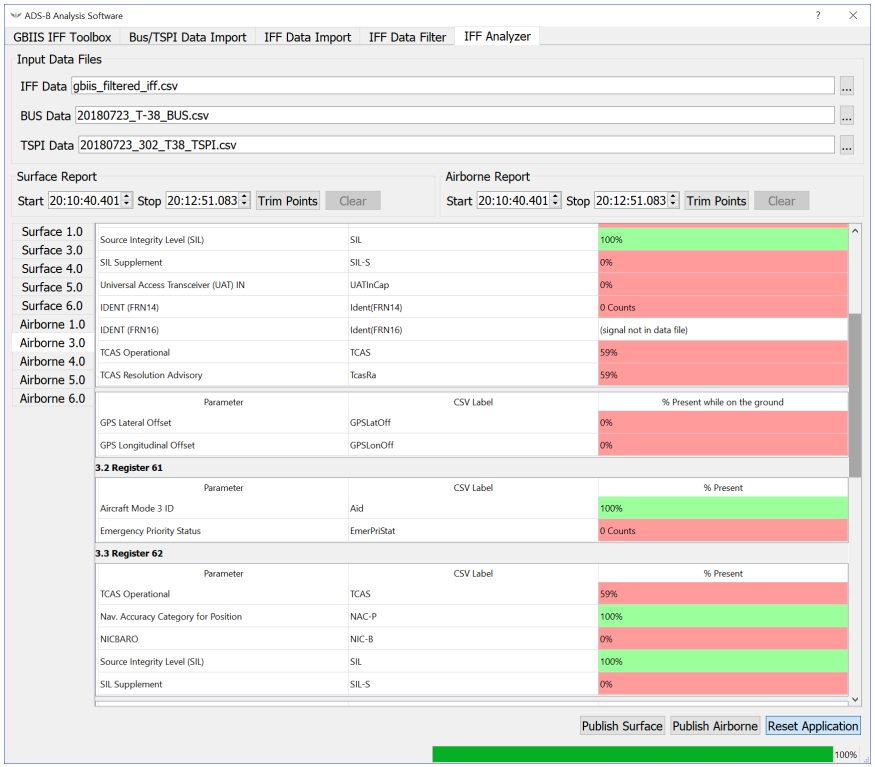
- ADS-B Analysis Software
  - Validates ADS-B transmission data against BUS/TSPI data
  - Able to filter ADS-B data to find the appropriate data
  - Fills in a Word template document to aid in report generation.



Example Applications

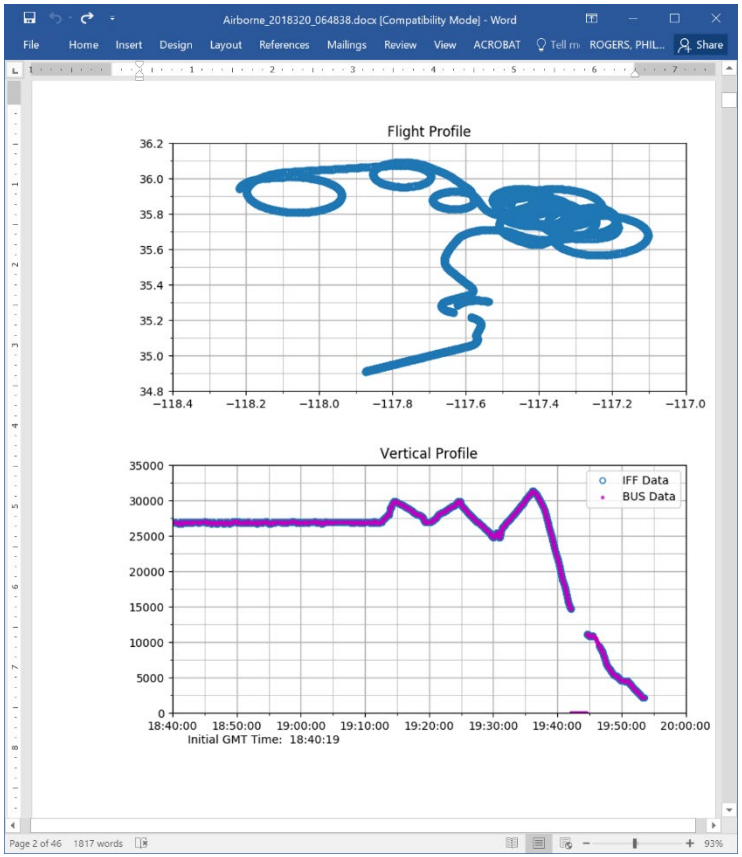
- ADS-B Analysis Software
  - Generates and previews the various report sections that include plots and data tables

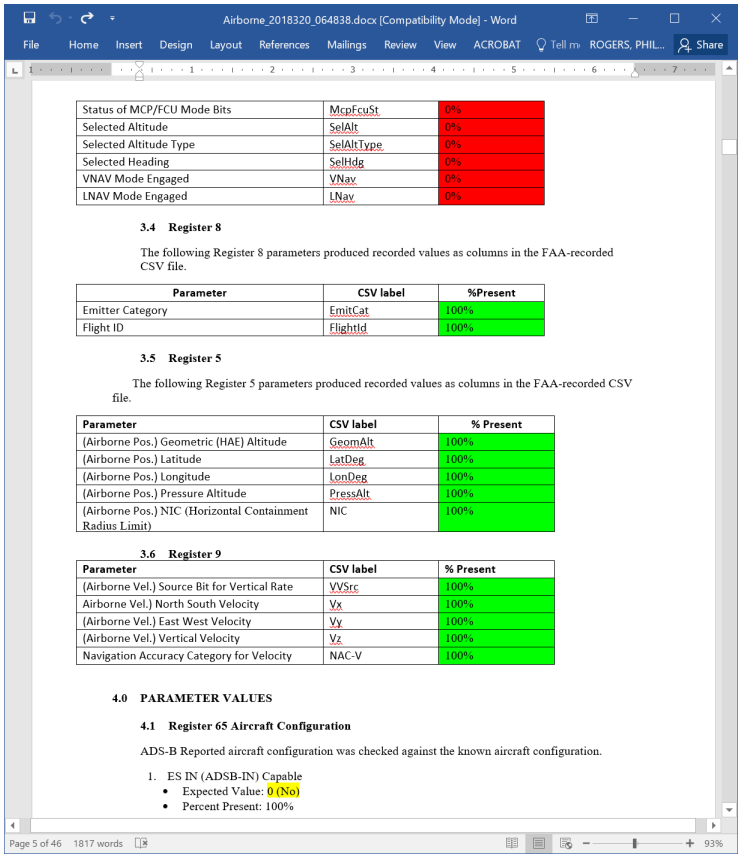




Example Applications

- ADS-B Analysis Software
  - Auto populates a Word document template with the calculated values and plots





Status of MCP/FCU Mode Bits	McpFcuSt	0%
Selected Altitude	SelAlt	0%
Selected Altitude Type	SelAltType	0%
Selected Heading	SelHdg	0%
VNAV Mode Engaged	VNav	0%
LNAV Mode Engaged	LNav	0%

3.4 Register 8

The following Register 8 parameters produced recorded values as columns in the FAA-recorded CSV file.

Parameter	CSV label	%Present
Emitter Category	EmitCat	100%
Flight ID	FlightId	100%

3.5 Register 5

The following Register 5 parameters produced recorded values as columns in the FAA-recorded CSV file.

Parameter	CSV label	% Present
(Airborne Pos.) Geometric (HAE) Altitude	GeomAlt	100%
(Airborne Pos.) Latitude	LatDeg	100%
(Airborne Pos.) Longitude	LonDeg	100%
(Airborne Pos.) Pressure Altitude	PressAlt	100%
(Airborne Pos.) NIC (Horizontal Containment Radius Limit)	NIC	100%

3.6 Register 9

Parameter	CSV label	% Present
(Airborne Vel.) Source Bit for Vertical Rate	VVSrc	100%
Airborne Vel.) North South Velocity	Vx	100%
(Airborne Vel.) East West Velocity	Vy	100%
(Airborne Vel.) Vertical Velocity	Vz	100%
Navigation Accuracy Category for Velocity	NAC-V	100%

4.0 PARAMETER VALUES

4.1 Register 65 Aircraft Configuration

ADS-B Reported aircraft configuration was checked against the known aircraft configuration.

1. ES IN (ADSB-IN) Capable
   - Expected Value: 0 (No)
   - Percent Present: 100%

Overview

- Introduction
- Proprietary Languages
- Open Source Software
- Python
- Example Applications
- **Conclusion**

Conclusion

- Python has been used to develop a wide array of data analysis tools and more are being developed
- Python is a robust scientific computing and visualization platform
- Python is financially competitive with commercial competitors due to its open source licensing
- Python is approved for DoD usage

- Consider using Python for your T&E computing requirements