412th Test Wing

War-Winning Capabilities … On Time, On Cost

IRIG 106 Synthetic Data Generation for Big Data Analytics Development

26 Oct 2022

Bob Baggerman
Avionics Test and Analysis Corp (ATAC)
404-502-0513

U.S. AIR FORCE

DISTRIBUTION STATEMENT A - Approved for public release; distribution is unlimited.
412TW-PA-22310

Integrity - Service - Excellence

Introduction

- Bob Baggerman
- Avionics Test and Analysis Corp (ATAC)

  Team member with JT4

  Office near Eglin AFB
- Email

  bob.baggerman@avtest.com

  bob@irig106.org
- Phone

  404-502-0513

Problem

- DoD is pursuing "Big Data Analytics" analysis and storage solutions for flight test data
  - Azure / AWS cloud services
  - TRMC CHEETAS software
- Developers lack realistic data set to develop with
  - Almost all flight test data sets are either classified or restricted in some way
  - Existing real world data set lack "interesting" features for developers to test search and analyze algorithms

Solution

Generate "Synthetic" Flight Test Data Files

- **Synthetic - Derived from unrestricted data**
  - **Previously recorded commercial flights**
  - **Aircraft dynamic simulation software**
- **Written into IRIG 106 Ch 10 format files**
- **No releasability issues**
- **Crafted to present interesting test cases**

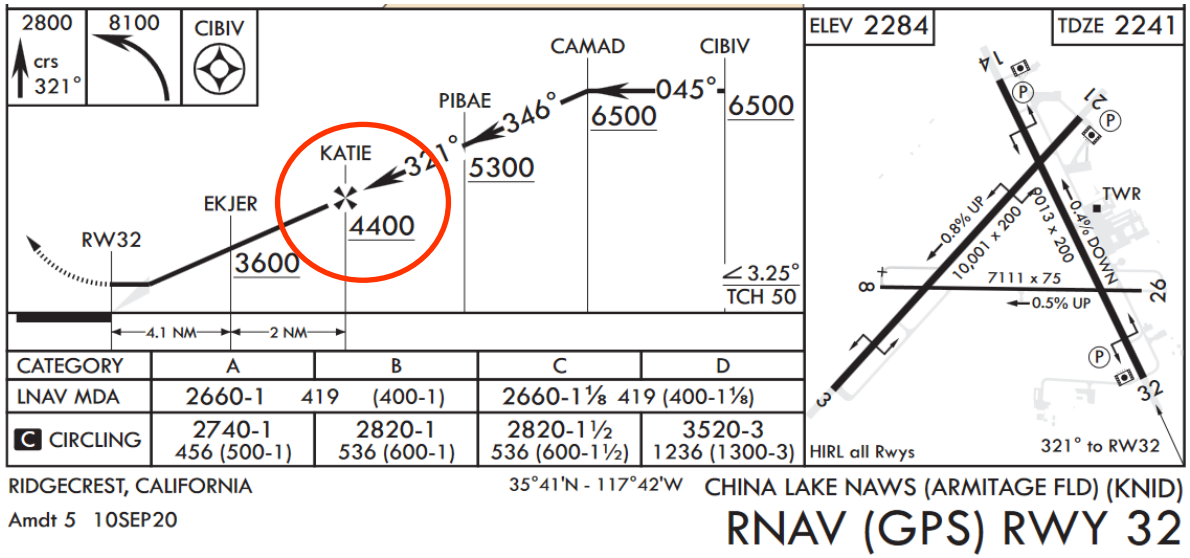
<u>Important Point!</u>

The System Under Test is the Big Data Analytics platform. Data sets are to support BDA development.

Examples of BDA Analysis

Nominal Flight Path

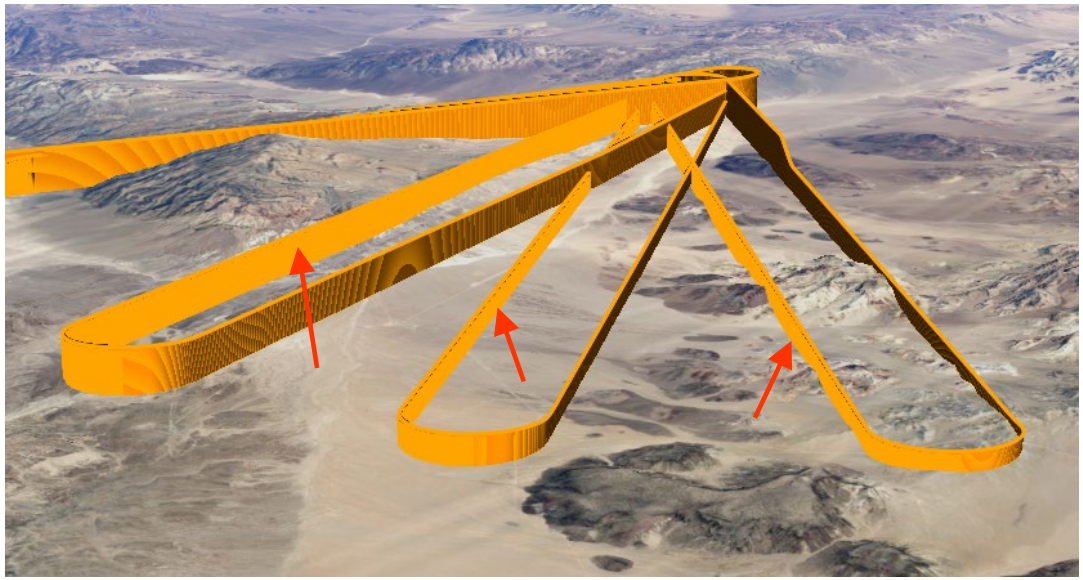
- **What is the average Altitude Error and Standard Deviation over the KNID Rwy 32 RNAV Final Approach Fix?**
- **What flights were more than 3 Standard Deviations from the correct Altitude at the FAF?**



Examples of BDA Analysis

Flight Segments for Analysis

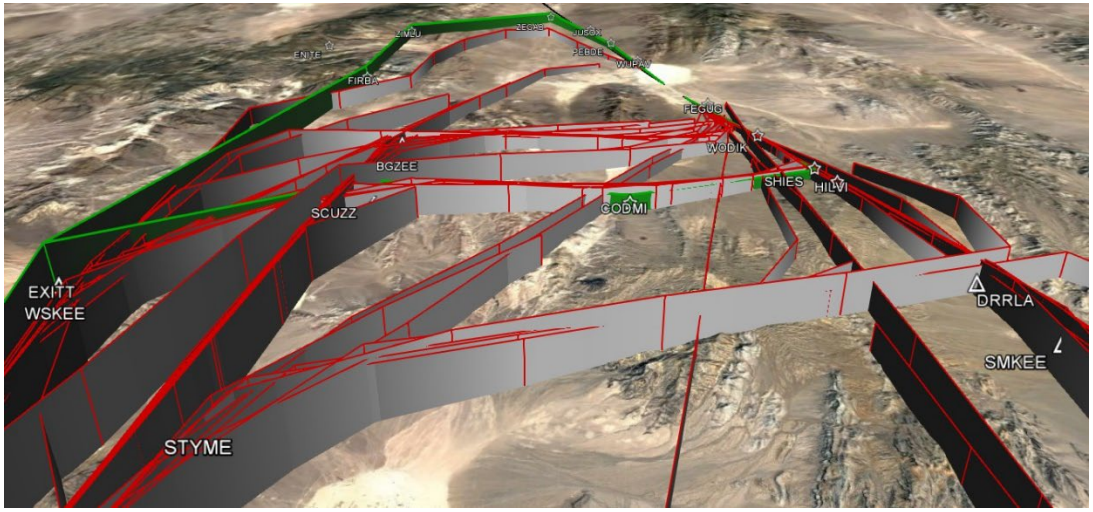
- What flight segments were flown on the Echo Range on Headings 180 to 270 between 3000' and 6000' MSL?



Examples of BDA Analysis

Flight Segments for EW Analysis

- What flight segments were flown on the Nellis Range between 5/1/2020 and 5/14/2020 where the RWR detected an SA-6?
- What was the Average and Standard Deviation of Detection Range to the Target?



Types of Synthetic Data

- **Contrived Data – Unrealistic flight test data but data types and values useful for testing correct decoding and conversion of IRIG 106 values**
  - **Carefully hand crafted**
  - **Presents edge cases**

- **Repurposed Data – Convert previously recorded flight data into IRIG 106 format**
  - **Large NASA repository of recorded flight data**

- **Synthesized Data – Mimic realistic flight test data but with very controlled flight conditions**
  - **Computer generated flight dynamics**

Contrived Data

- Generated from a content definition data file.
  - Format defined in IRIG 106 Chapter 10 Programming Handbook (RCC Document 123-16) Appendix P “XML Mapping”
- Content definition data file contents are written by hand in XML format
  - Only a few well-crafted data types and fields are necessary to validate a software data decoder or processor
- Content definition file converted to IRIG 106 Ch 10 data file
  - FLIDAS from Data Bus Tools GmbH

Contrived Data Example


```xml
<?xml version="1.0" encoding="UTF-8"?>


  <cns:Packet ChannelID="16" RTC="1">
    <cns:ARINC429Data>
      <cns:ARINC429Message Subchannel="1" Raw31="0x048D14F0" GapTime="10000"
      <cns:ARINC429Message Subchannel="1" Raw31="0x150C84F0" GapTime="10000"
      <cns:ARINC429Message Subchannel="1" Raw31="0x048D16F0" GapTime="10000"
      <cns:ARINC429Message Subchannel="1" Raw31="0x248D16F0" GapTime="10000"
      <cns:ARINC429Message Subchannel="1" Raw31="0x248D160F" GapTime="10000"

      <cns:ARINC429Message Subchannel="1" Raw31="0x000000F0" GapTime="10000"
      <cns:ARINC429Message Subchannel="1" Raw31="0x1FFFFCF0" GapTime="10000"

      <cns:ARINC429Message Subchannel="2" Raw31="0x000000F0" GapTime="10000"
      <cns:ARINC429Message Subchannel="2" Raw31="0x000004F0" GapTime="10000"
      <cns:ARINC429Message Subchannel="2" Raw31="0x1FFFFCF0" GapTime="10000"
      <cns:ARINC429Message Subchannel="2" Raw31="0x0FFFFCF0" GapTime="10000"
      <cns:ARINC429Message Subchannel="2" Raw31="0x100000F0" GapTime="10000"

      <cns:ARINC429Message Subchannel="3" Raw31="0x      F0" GapTime="10000"
      <cns:ARINC429Message Subchannel="3" Raw31="0x      F0" GapTime="10000"
      <cns:ARINC429Message Subchannel="3" Raw31="0x      F0" GapTime="10000"

    </cns:ARINC429Data>
  </cns:Packet>
</cns:ch10>
```

Repurposed Data Sources

- NASA Sample Flight Data
  - Recorded regional jets

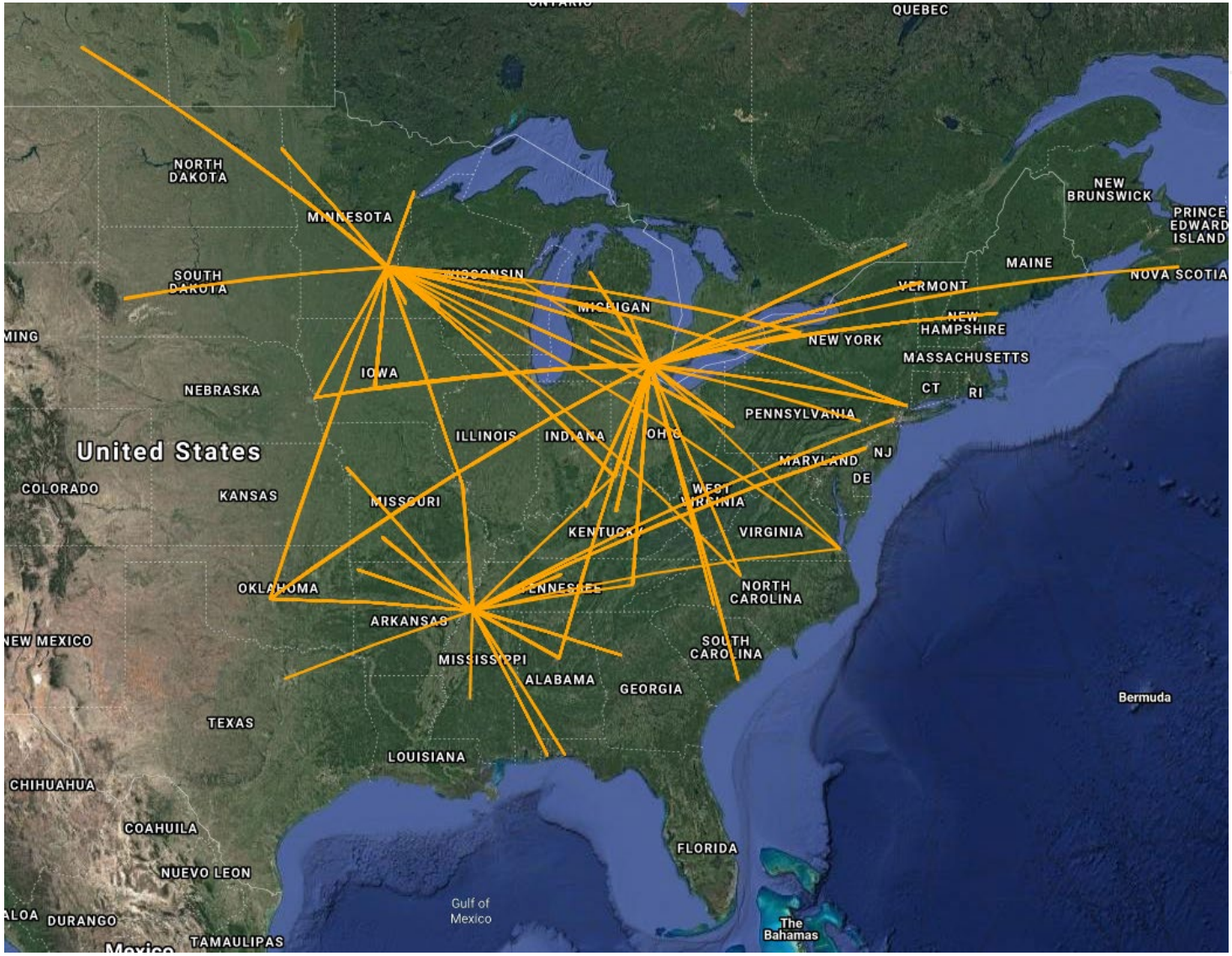
- ADS-B not used
  - Lots of available data but very limited data scope
- Unclassified USAF flights (e.g. T-38) not used
  - Unclassified but not publicly releasable
- "Live Player" not used
  - Too labor intensive to generate data
  - Not deterministic enough

Repurposed Data

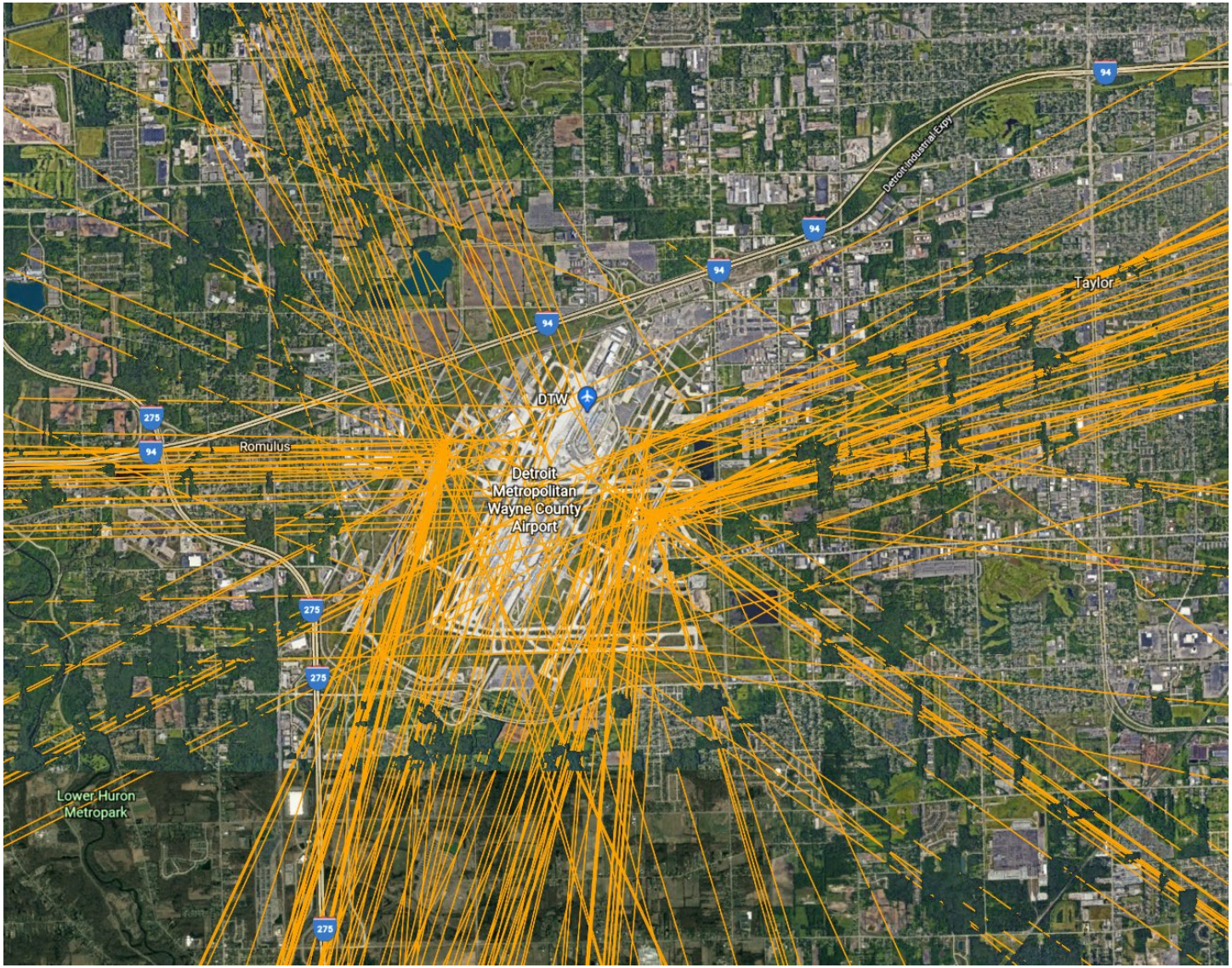
- **NASA Sample Flight Data**
  - **From NASA DASHlink**
  - **Actual data recorded onboard a single type of regional jet operating in commercial service over a three-year period**
  - **De-identified aggregate flight recorded data**
  - **Detailed aircraft dynamics, system performance, and other engineering parameters**
    - **186 recorded parameters**
  - **Matlab format, converted to CSV**
  - **220,000 flights!**

https://c3.nasa.gov/dashlink/projects/85/

NASA Data Set 1 Summary



NASA Data into Detroit



Synthesized Data

- Start with synthetic aircraft position and other state
  - Position / Attitude / Velocity / Acceleration
  - IAS / Power / Flap Position / etc…
- Derive additional state from this aircraft state
  - Accelerometers / Strain Gauges / etc.
  - Threat engagement
    - RWR threat display
    - Missile Warning display
- Generate IRIG 106 Chapter 10 data file
  - 1553 Channels
  - PCM Channels
  - Video Channels
  - ARINC 429

Synthesis Software

- Suite of Software Programs
  - GOTS
    - BlueMax6 Flight Dynamics Simulation
  - COTS
    - X-Plane Flight Simulator
  - Custom C++
    - Preprocessing
    - Simulation Engine / Ch 10 File Generation

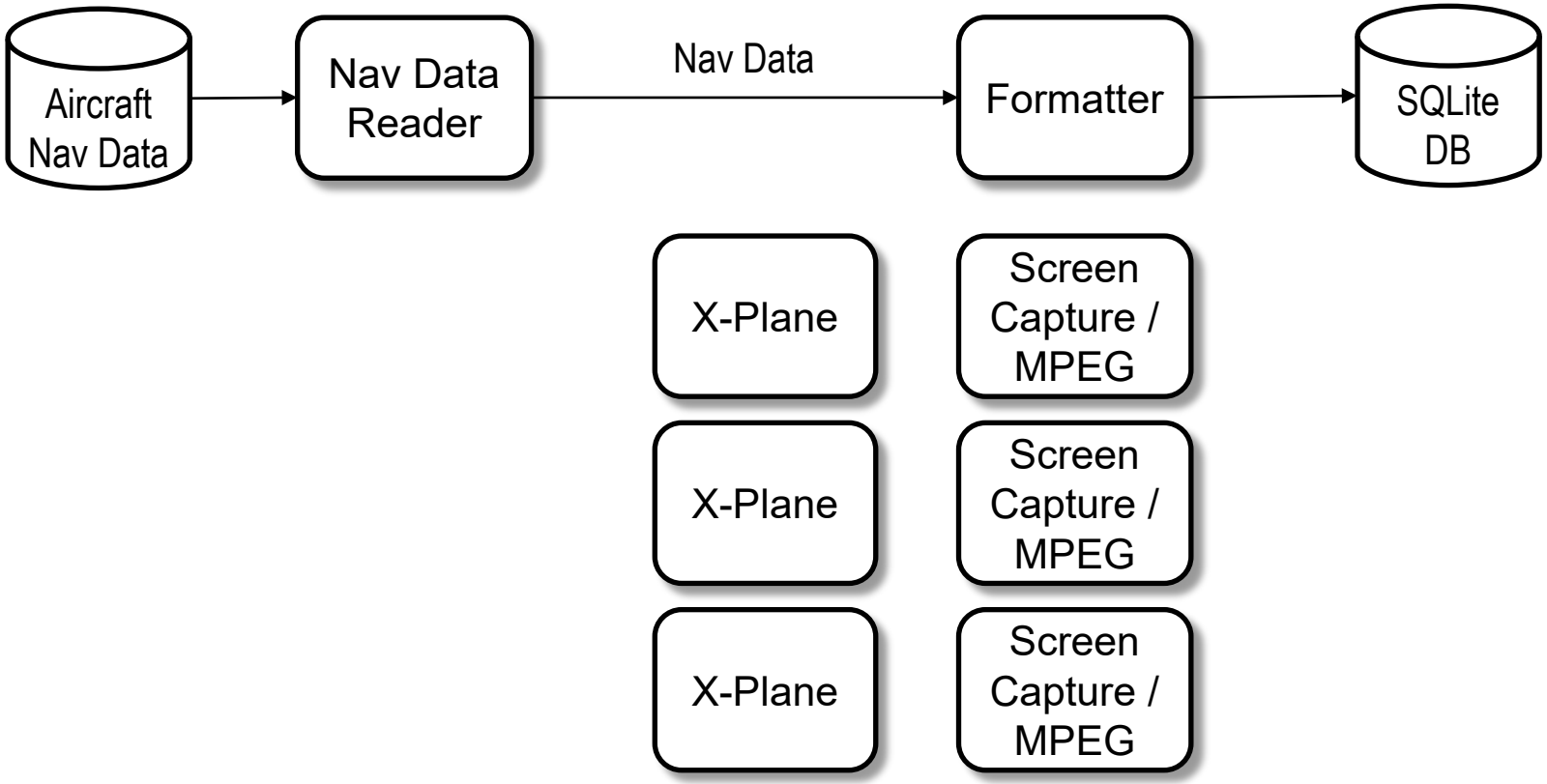
Synthesis Software

Simulation Engine / Ch 10 Generation

- C++ / Windows
- Command line driven to support batch operation
- Architecture
  - Data Source modules feed state machine
  - Data Formatter modules feed Ch 10 writing
  - Simple time slice simulation engine
- Source and Ch 10 ICD available on github

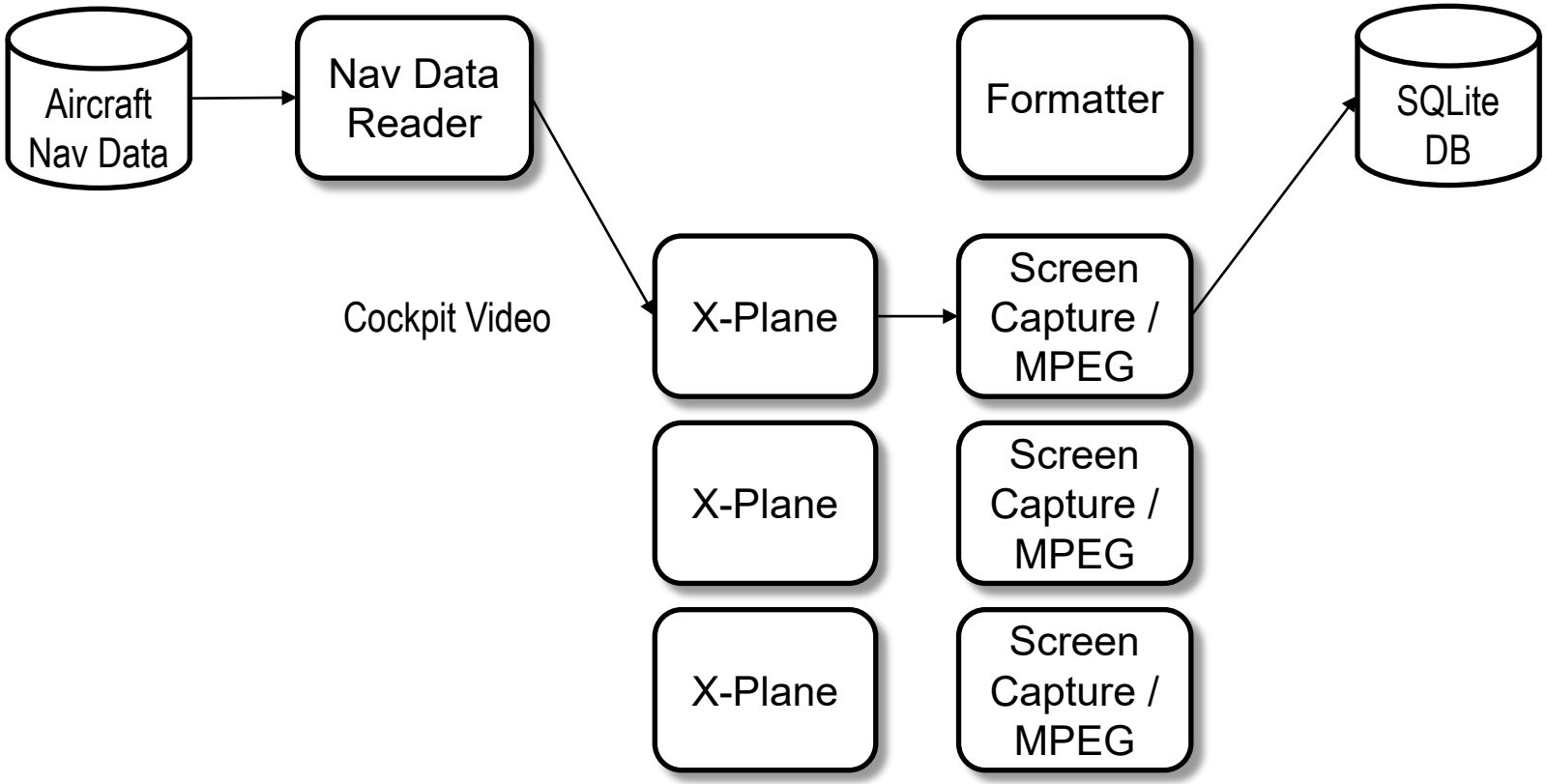
Synthesis Software

Preprocessing



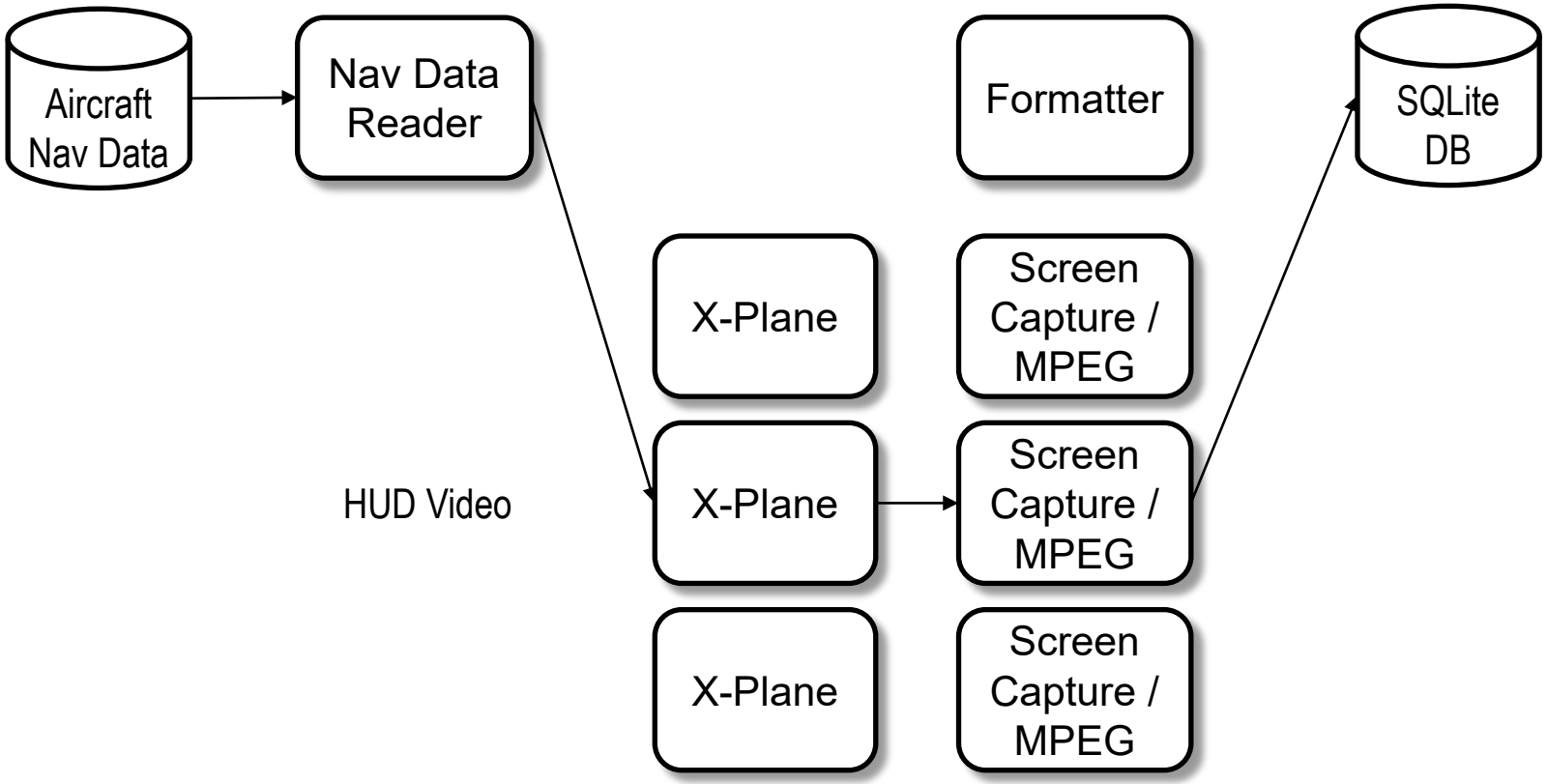
Synthesis Software

Preprocessing



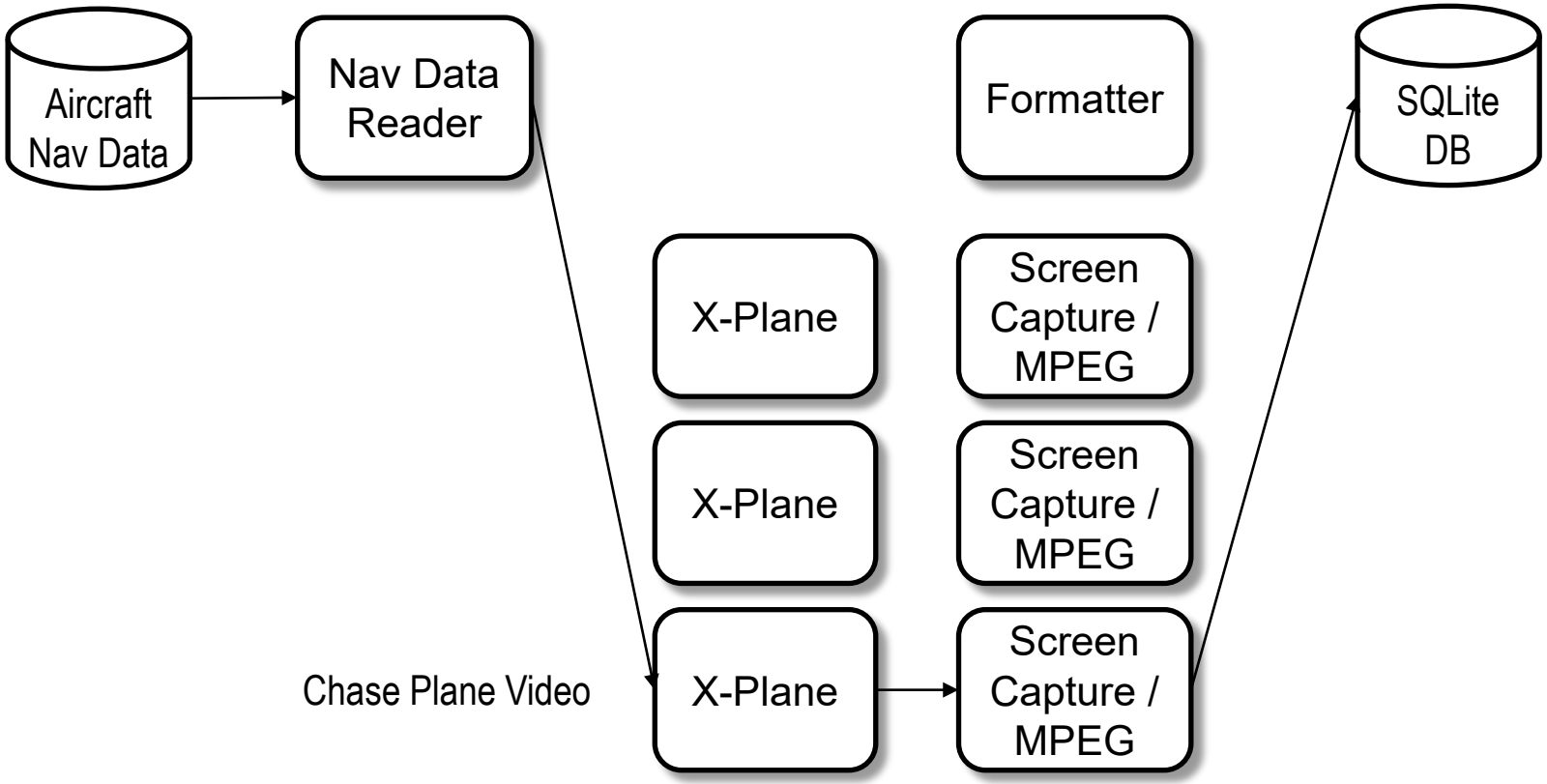
Synthesis Software

Preprocessing



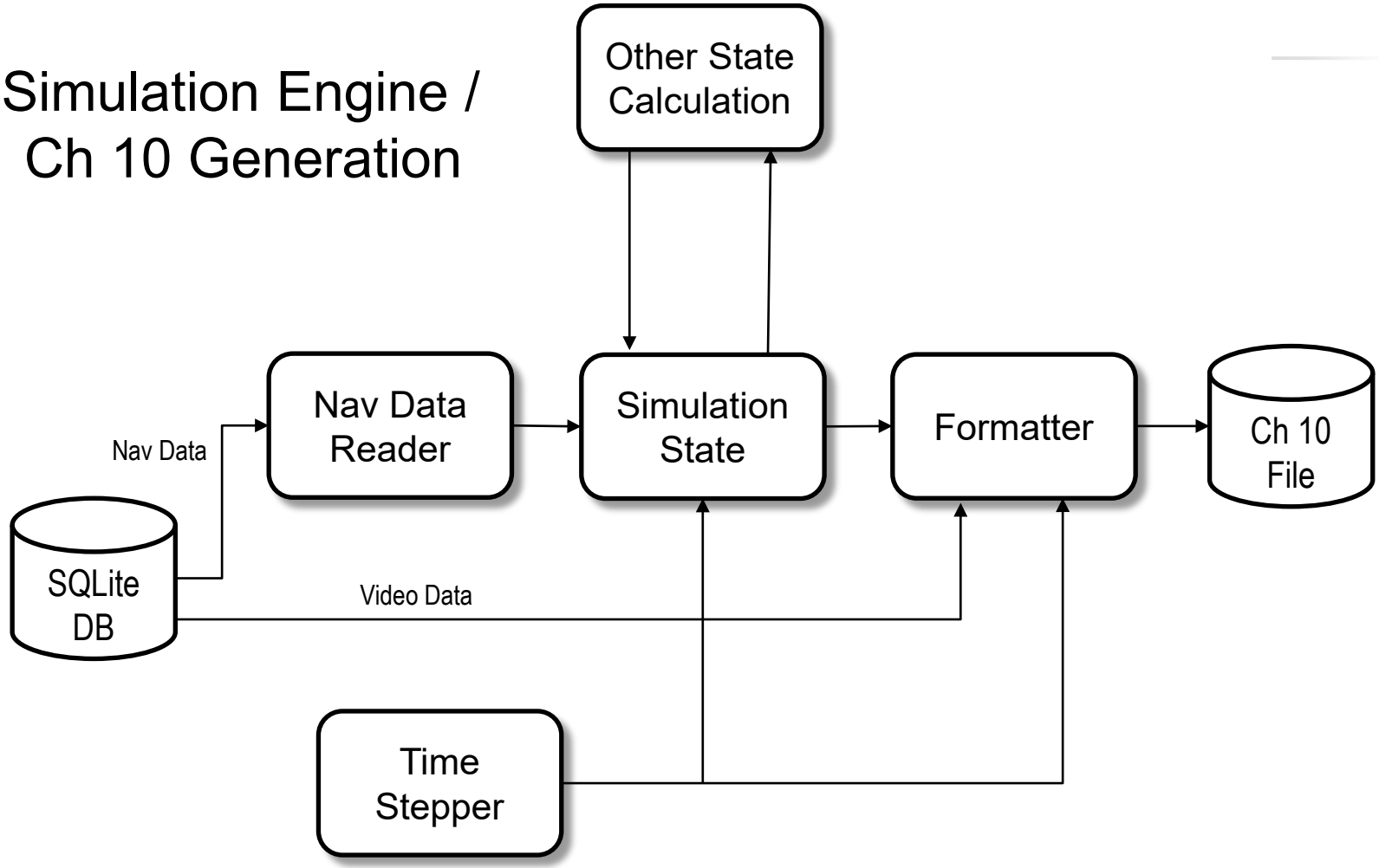
Synthesis Software

Preprocessing



Synthesis Software

Simulation Engine / Ch 10 Generation



BlueMax6 Data Source

- BlueMax6
  - GOTS software from DSIAC
  - Software simulation
  - Flight scenario described at a very high level
  - Generates realistic flight dynamics data
    - Lat / Lon / Alt / TAS / AOA / Throttle / …
    - 67 available flight parameters

BlueMax6 Example

Scenario File

```
CmdAltitudeMSLf 2300
CmdGearPosition 2200
CmdAirspeedMach BestRateOfClimb
CmdFlapPosition Auto
CmdSegmentEndMode Acquisition
CmdFlySegment

WriteMessage Low Pass Takeoff
CmdTrueHeading 154.5
CmdGroundRangeNm 2
CmdAltitudeMSLf 2300
CmdThroPosition 300
CmdFlapPosition 0
CmdSlatPosition 0
CmdFlySegment

WriteMessage Location 1
CmdWaypointLL 35.XXXXXX:N 117.XXXXXX:W
CmdWaypointNavMode Direct
CmdAltitudeMSLf 6000
CmdAirspeedKtas 300
CmdFlySegment
```

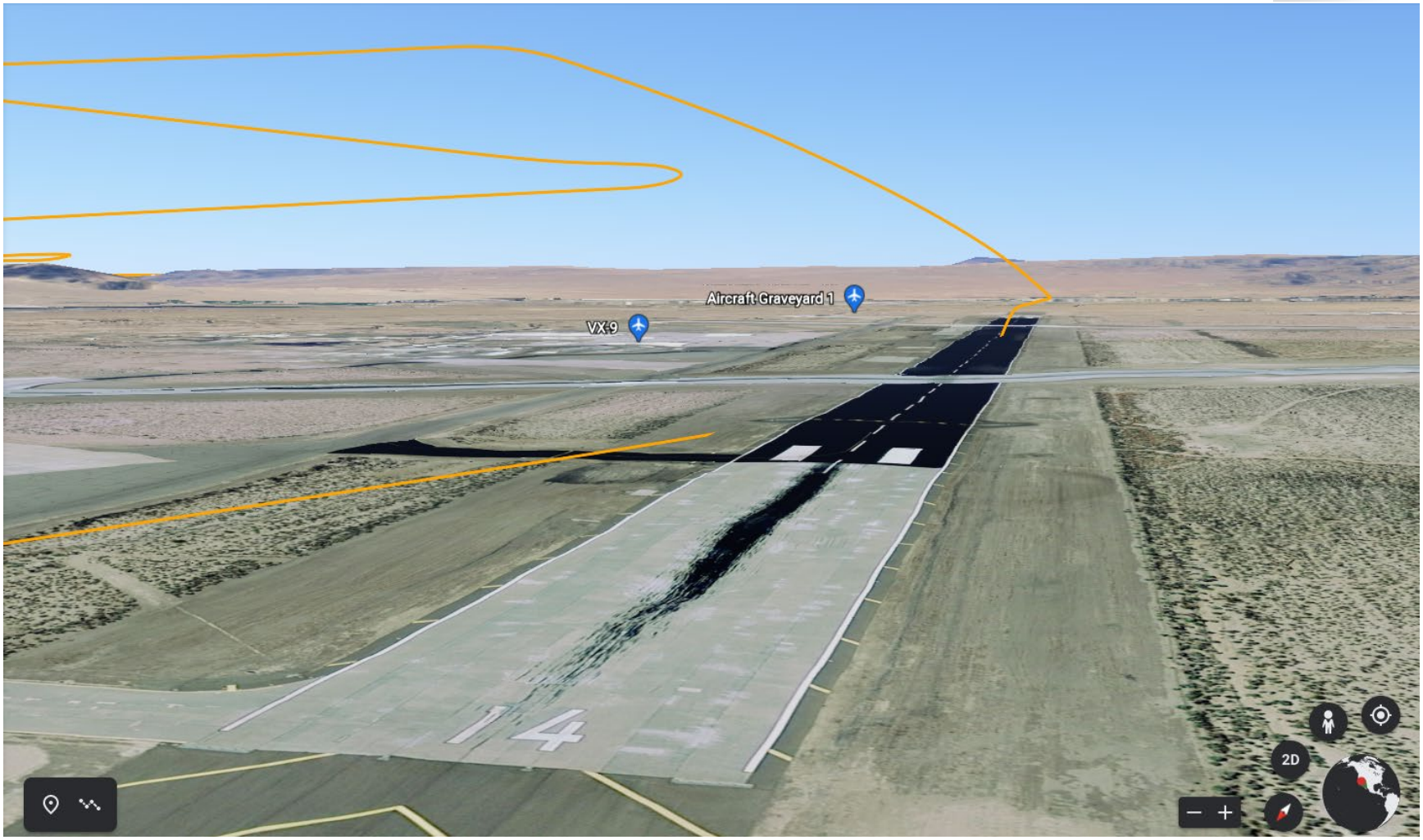
BlueMax6 Example

Output Data File

actime	aclatd	aclond	acaltf	acktas
182.6800000000000068	35.6815813184698172	-117.6469659063139801	4214.0491672053231014	234.9775112918847242
182.7199999999999989	35.6816175507618780	-117.6469954548966825	4214.5785340285856364	234.9733499230588336
182.7599999999999909	35.6816538554664646	-117.6470248673700780	4215.1050067184696672	234.9691469844862297
182.8000000000000114	35.6816902323893785	-117.6470541431704930	4215.6286798426426685	234.9651945717821491
182.8400000000000034	35.6817266807135312	-117.6470832829917441	4216.1492098506369075	234.9609685909008476
182.8799999999999955	35.6817632003089855	-117.6471122861824625	4216.6667279553757908	234.9569890259908505
182.9199999999999875	35.6817997905321675	-117.6471411530263680	4217.1810291864694591	234.9527846554025814

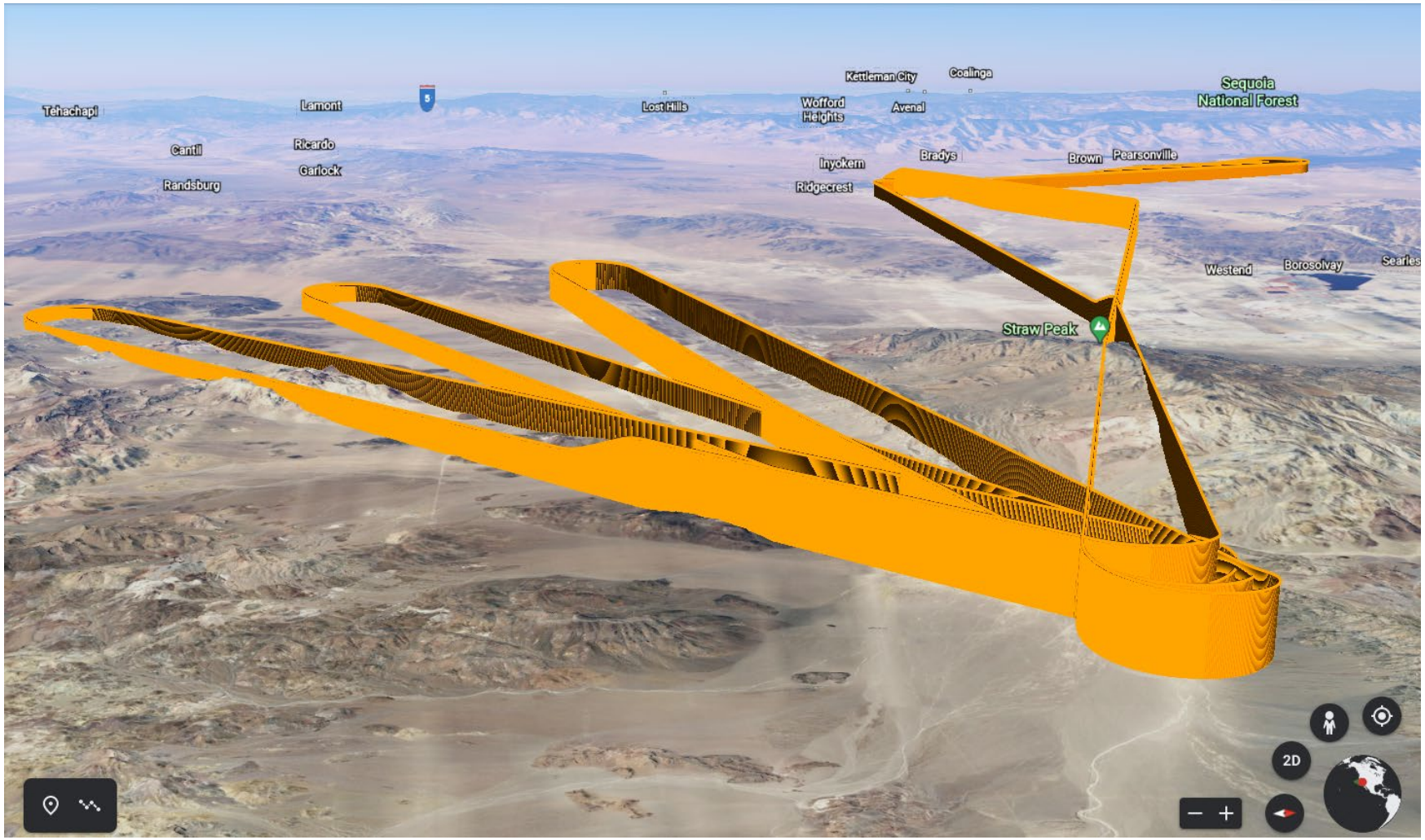
BlueMax6 Scenario

Takeoff China Lake Rwy 14



BlueMax6 Scenario

Wagon Wheels over Echo Range



Ch 10 Playback

Conclusion

- Big Data Analytics development needs
  - Unrestricted data
  - Data that offers useful test cases
- Synthetically generated data meets these needs
  - Different types for different needs
- Future Plans
  - More derived simulation state
  - More Ch 10 data types
  - Improved processing for video data