

X-Files:

A Data Challenge for Space Domain Awareness

<https://xfiles.llnl.gov>

Robert Armstrong, Dan Merl, Joshua Meyers,
Kerianne Pruett, Eddie Schlafly, Michael Schneider

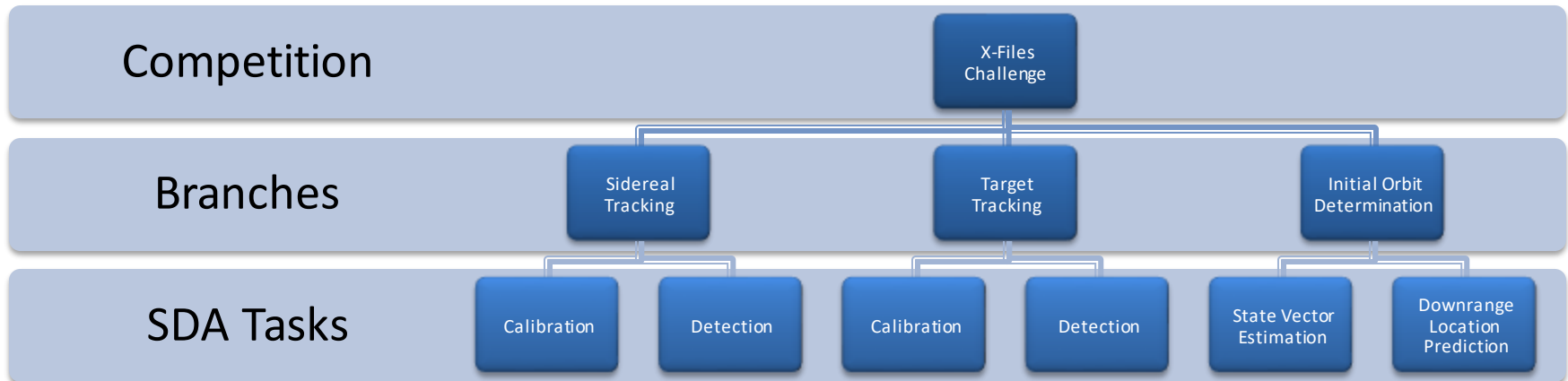
September 2021
Data version: v2.2

xfiles@llnl.gov



X-Files: A data analytics competition for SDA

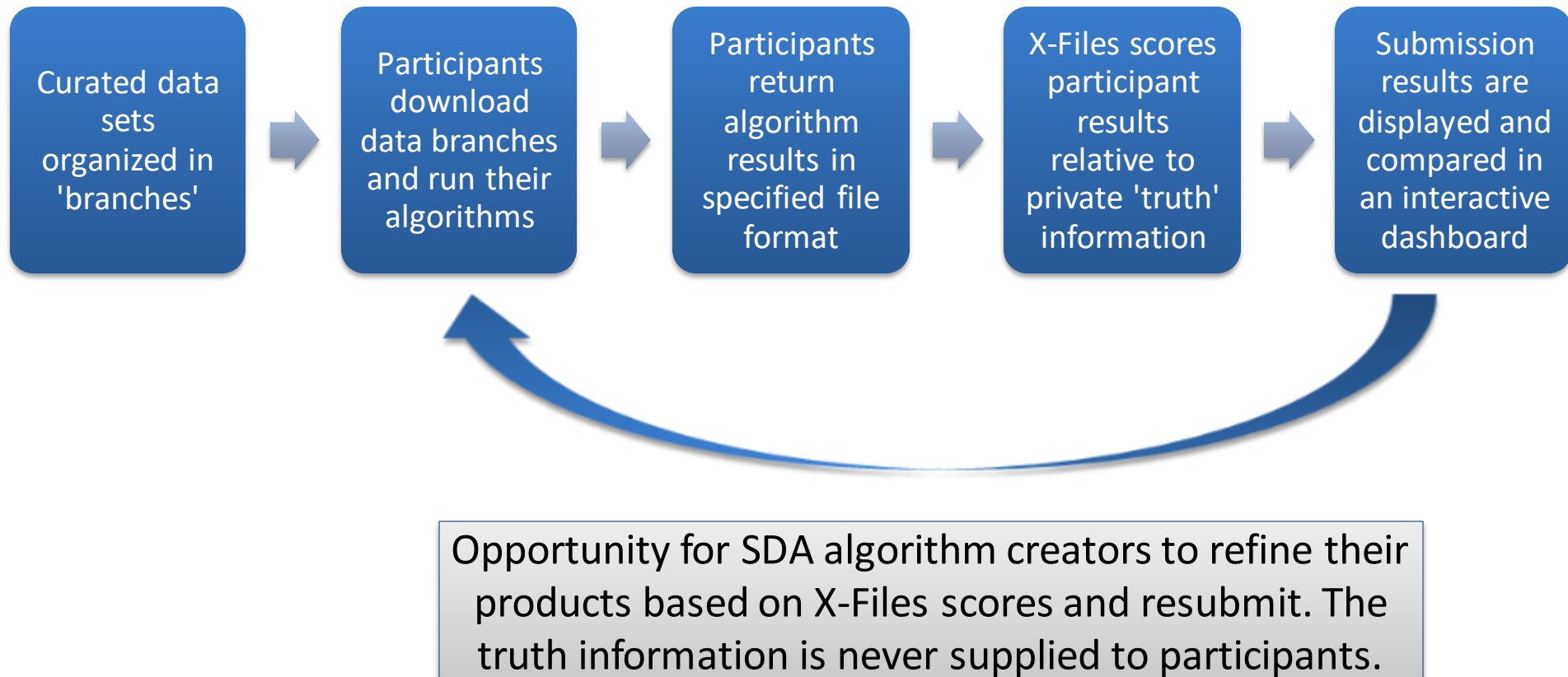
- The challenge requires blind analysis of curated data sets ('branches') for distinct tasks in an optical LEO SDA pipeline:



- X-Files participants submit formatted analysis results that get scored and populated onto the X-Files dashboard
 - <https://xfiles.llnl.gov>
- Adding your score to the X-Files leaderboard is a great way to show off your SDA algorithms and get attention for your product!

A community competition to advance SDA data analytics

The X-Files participation process



Download data, run your algorithm, and submit results to earn a score

Branches in X-Files v2.2

A. Target Tracking

- Simulated images of ground-based target tracking observations for testing locating SROs in images.

B. Sidereal Tracking

- Simulated images of ground-based sidereal tracking observations for testing locating SROs in images.

C. IOD – Target Tracking

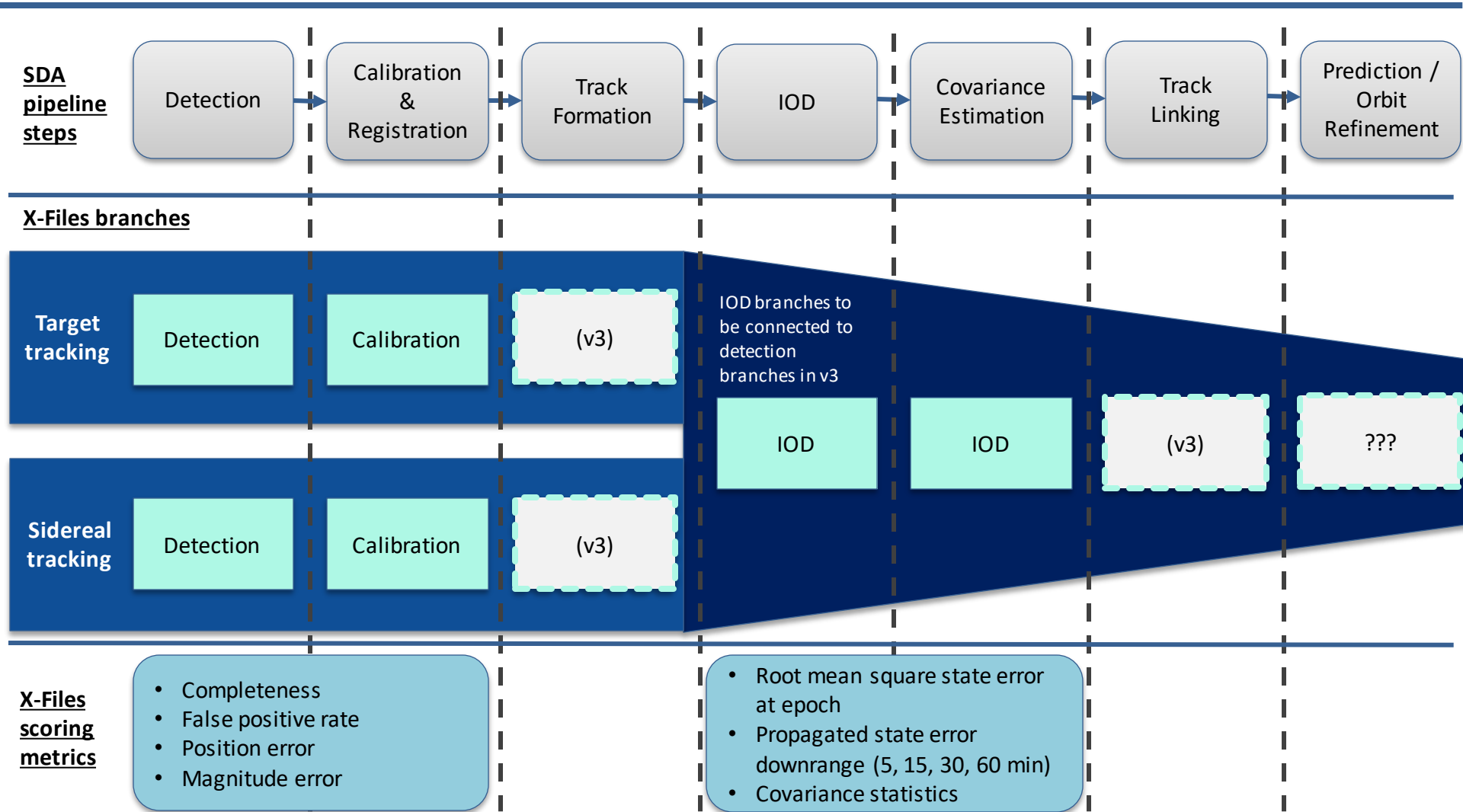
- Tracks of optical, angle-only observations of varying lengths for testing IOD predictions.



Example Sidereal branch image

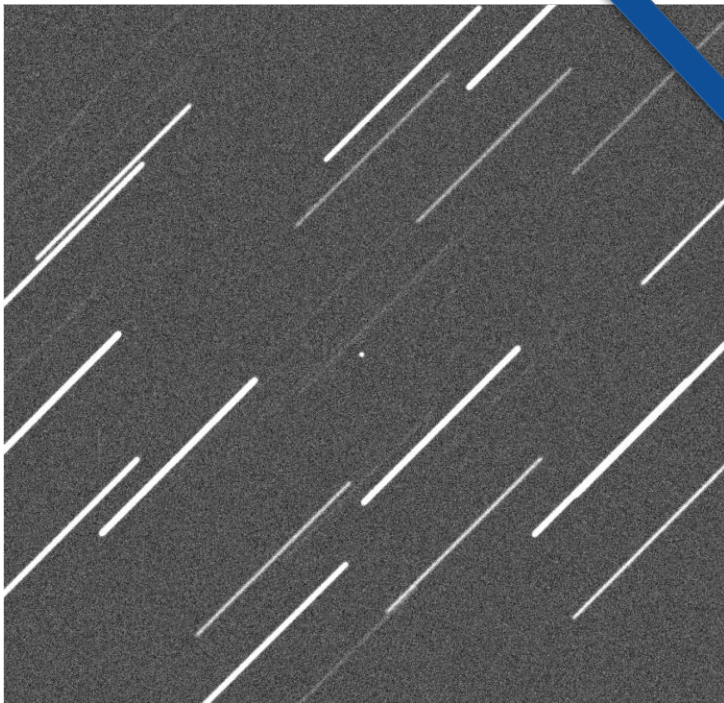
Participants may run as many algorithms on as many branches as they like

The X-Files 'branches' test distinct components of an SDA data analysis pipeline



Target/Sidereal Branch Structure

1) Competitor is supplied set of images (per branch)



Example Target branch image

3) Competitor generates a submission file that gets returned for a score

2) Competitor uses their algorithm to determine the information of interest

```
branch: Sidereal
competitor_name: Competitor A
display_true_name: true
file: 000.fits
sats:
- dec0: 0.9824085674177758
  dec1: 0.9827694532849132
  flux: 7457.090027770151
  mag: 13.494536657817662
  ra0: 0.7990369541568129
  ra1: 0.7961051351350783
  x0: 915.2029741109612
  x1: 1089.8194690755304
  y0: 724.7517276023891
  y1: 582.6143955011476
stars:
- dec: 0.984135803379026
  flux: 1189804.0268676954
  mag: 8.050434365599326
  ra: 0.7917760447120902
  x: 1436.2030497375654
  y: 441.1471290131511
- dec: 0.9828443340978492
  flux: 777127.4472634476
  mag: 8.567482602510315
  ra: 0.787298190854206
  x: 1507.5133555757925
  y: 67.04054447333851
- dec: 0.9814851323107805
  flux: 245450.66978989646
  mag: 9.755139911280292
  ra: 0.789036863363843
  x: 1281.0602737235993
  y: 54.51005867202332
...
---
file: 001.fits
...
...
```

The "Target" and "Sidereal" branches test the competitor on detection and calibration tasks

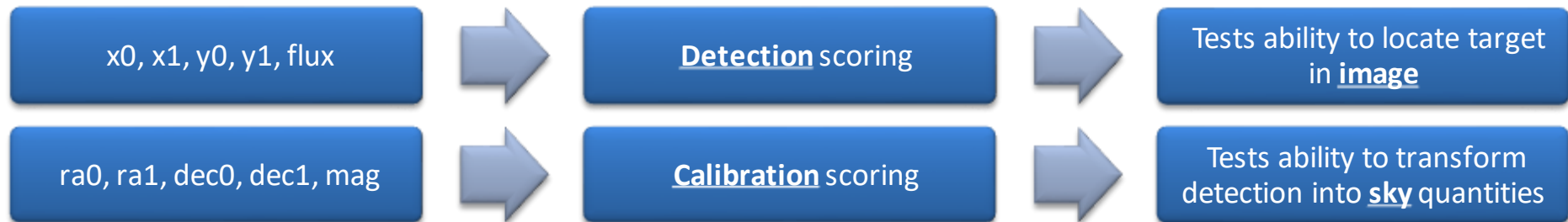
Target/Sidereal Branch Information

Tests ability to detect targets within ground-based optical images

■ Branch Simulation

- Images include optical distortion, variable surface brightness background, and vignetting, and cover a range of exposure times.
- Star positions and fluxes are derived from Gaia DR2.

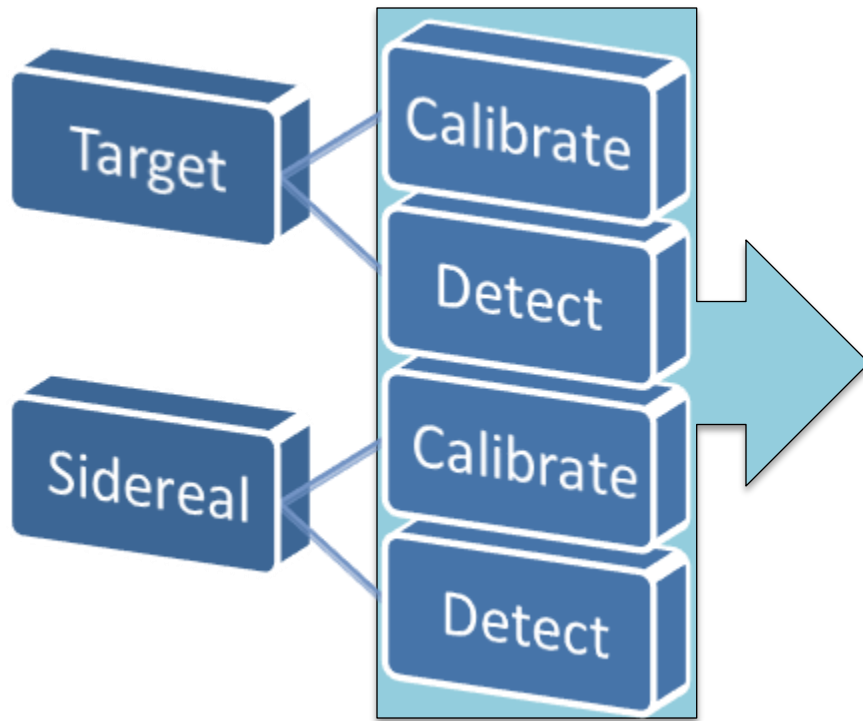
■ Scores (based on type of submission information)



■ Notes

- Regardless of branch, each target is treated as a **streak**
- Supplying star information is optional
- Minimum submission required to receive a score:
 - x0, x1, y0, y1, and flux for each target in an image, for ALL images

Target/Sidereal Branch Scoring Metrics



- For each branch and task combination, submissions will receive the following scores:
 - **Completeness:**
 - The number of true satellites found in the submissions, divided by the number of true satellites
 - $\text{completeness} = (n_{\text{true_found}} / n_{\text{true}})$
 - **False positive rate:**
 - The number of false positive detections found in submissions (that could not be associated with a true satellite), over the number of true satellites
 - $\text{fpr} = 1 - (n_{\text{associated}} / n_{\text{true}})$
 - **Position error:**
 - Root mean square error (RMSE) of submitted position
 - **Magnitude error:**
 - Root mean square error (RMSE) of submitted magnitude
- Detection scores use x/y/flux and calibration scores use ra/dec/mag

IOD Branch Structure

1) Competitor is supplied set of catalogs.

| | time | ra | dec | px | py | pz |
|----|------------|-----------|-----------|------------|------------|-------------|
| 1 | 1.261915E9 | 328.25559 | -22.15021 | 4.482517E6 | 3.077780E6 | -3.323293E6 |
| 2 | 1.261915E9 | 329.04517 | -19.62639 | 4.480271E6 | 3.081053E6 | -3.323289E6 |
| 3 | 1.261915E9 | 329.7824 | -17.1358 | 4.478023E6 | 3.084324E6 | -3.323284E6 |
| 4 | 1.261915E9 | 330.46313 | -14.69234 | 4.475773E6 | 3.087593E6 | -3.323280E6 |
| 5 | 1.261915E9 | 331.08696 | -12.29511 | 4.473520E6 | 3.090861E6 | -3.323276E6 |
| 6 | 1.261915E9 | 331.67144 | -9.95214 | 4.471265E6 | 3.094126E6 | -3.323271E6 |
| 7 | 1.261915E9 | 332.22074 | -7.67313 | 4.469008E6 | 3.097391E6 | -3.323267E6 |
| 8 | 1.261915E9 | 332.72822 | -5.44577 | 4.466748E6 | 3.100653E6 | -3.323263E6 |
| 9 | 1.261915E9 | 333.19972 | -3.29777 | 4.464486E6 | 3.103914E6 | -3.323258E6 |
| 10 | 1.261915E9 | 333.63152 | -1.20693 | 4.462221E6 | 3.107174E6 | -3.323254E6 |
| 11 | 1.261915E9 | 334.03362 | 0.81513 | 4.459954E6 | 3.110431E6 | -3.323249E6 |
| 12 | 1.261915E9 | 334.40729 | 2.77093 | 4.457685E6 | 3.113688E6 | -3.323245E6 |
| 13 | 1.261915E9 | 334.7542 | 4.65925 | 4.455413E6 | 3.116942E6 | -3.323241E6 |
| 14 | 1.261915E9 | 335.07145 | 6.48579 | 4.453139E6 | 3.120195E6 | -3.323236E6 |
| 15 | 1.261915E9 | 335.36612 | 8.23166 | 4.450863E6 | 3.123446E6 | -3.323232E6 |
| 16 | 1.261915E9 | 335.63907 | 9.92325 | 4.448584E6 | 3.126695E6 | -3.323227E6 |
| 17 | 1.261915E9 | 335.8853 | 11.55396 | 4.446303E6 | 3.129943E6 | -3.323223E6 |
| 18 | 1.261915E9 | 336.11521 | 13.11908 | 4.444019E6 | 3.133189E6 | -3.323219E6 |
| 19 | 1.261915E9 | 336.32759 | 14.63162 | 4.441733E6 | 3.136434E6 | -3.323214E6 |
| 20 | 1.261915E9 | 336.51369 | 16.09348 | 4.439445E6 | 3.139676E6 | -3.323210E6 |

Example IOD branch catalog

2) Competitor uses their algorithm to determine the information of interest

Example of minimum IOD submission file

```
branch: IOD
competitor_name: Competitor A
display_true_name: true
---
IOD:
- rx: -1537532.386353053
  ry: -6096250.110596744
  rz: -2932138.755877168
  t: 1261905338.001
  vx: -311.47745042782947
  vy: 3343.468207729812
  vz: -6799.62176813366
file: SAT_0000.csv
---
IOD:
...
...
```

3) Competitor generates a submission file that gets returned for a score

The "IOD" branch tests the competitors' ability to make accurate initial orbit determinations

IOD Branch Information

Tests accuracies of initial orbit determinations, inferring positions and velocities from observations over a fraction of an orbital period

- **Branch Simulation**
 - Optical, angle-only observations
 - Varying track lengths
 - Noise consistent with typical target-tracking images
- **Submission:** For each target, the competitor is given a catalog with multiple observations, and asked to return:
 - The state ($r_x, r_y, r_z, v_x, v_y, v_z$) at time t
 - t = the time associated with the last observation (chronologically) in the .csv file
 - Supplying this for ALL targets is the minimum submission criteria to receive a score
 - Downrange locations (optional)
 - Ra and dec of target 5, 15, 30, and 60 minutes downrange
 - Predicted location covariance matrix (optional)
 - State vector covariance matrix (optional)
- If downrange locations are not supplied, we use the SGP4 propagator to propagate forward to 5, 15, 30, and 60 minutes

IOD Branch Scoring Metrics

- **Root mean square error (RMSE) scores**

- rx, ry, rz, vx, vy, vz
- Determined from the supplied state

- **Propagated scores**

- 5, 15, 30, and 60 minutes downrange
- For SGP4 propagator and/or your submitted downrange locations
- Score determined by how many targets are within FOV at propagated time
 - Determined using angular distance from truth position at that time

- **Statistic scores**

- Cramer Von-Mises probabilities:
 - Downrange positions and state estimate
 - Describe how well errors match uncertainties
- Z-scores:
 - One score for each of 5 quantiles
 - Compares empirical quantiles to those expected for a Gaussian or normal distribution

Distribution Structure

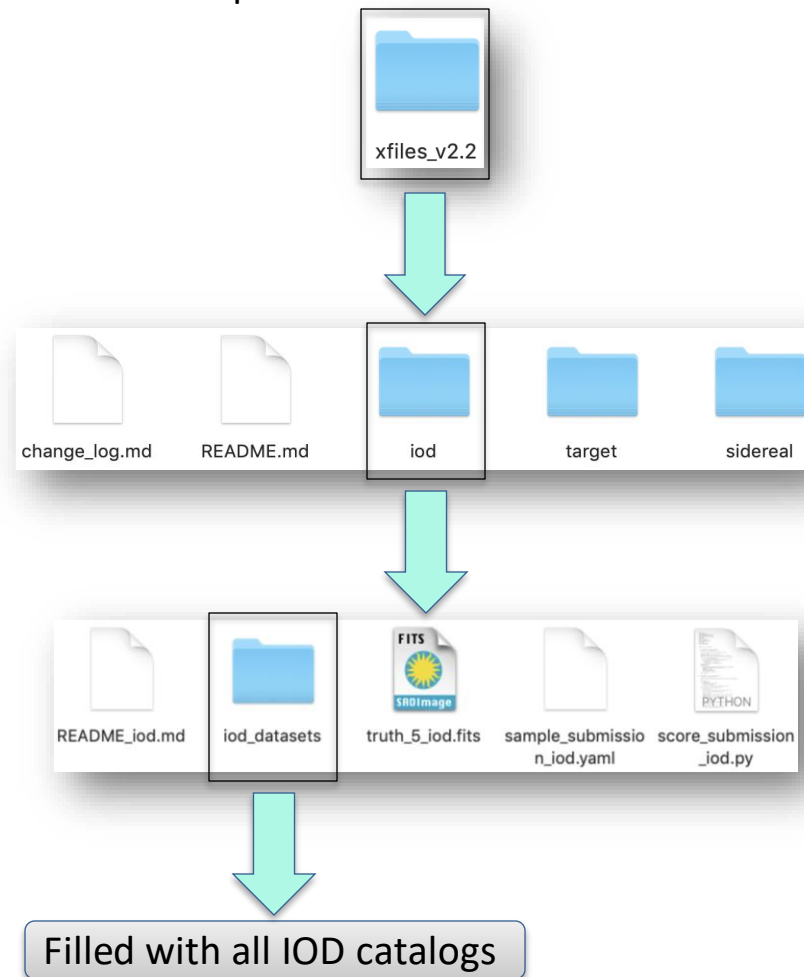
■ For the competitor

- Folder of input data
 - Target/Sidereal = FITS format images
 - IOD = .csv format catalogs
- Truth table for first few input data
 - FITS format, with extensions with star and sat info
- Sample submission file
 - YAML file demonstrating correct submission format
- Script to check and score submission
 - Python script for verifying submission format, and testing scoring on first few input data
- Documentation
 - Markdown files that describe data simulation, file structure, expected submission format, and scores.

■ For X-Files team private use

- Truth table for ALL input data
- All simulation scripts

Example File Structure – IOD Branch



X-Files Dashboard

A visual interface for understanding scores

- <https://xfiles.llnl.gov> brings the user to the info page that describes the challenge and has downloadable links for each data set.
- Upon submission, scores will be added to the dashboard.
 - Note: There is an option to remain anonymous.
- Each competitor has their own page (per branch) to visualize their scores, as well as access to all other competitors pages.
- If your score beats a current high score, you will be featured on the leaderboard.
- Interactive plots for further analysis

The screenshot shows the X-Files dashboard with a navigation bar at the top containing 'X-Files', 'Leaderboard', 'Branches', 'Submissions', and a 'Logout' link. The main heading is 'X-Files: A Data Challenge for Space Domain Awareness' with a subtitle 'Testing the robustness and accuracy of satellite detection, calibration, and characterization algorithms'. Below this, it states 'The following curated datasets ("branches") are currently available:'. There are three main sections: A. Target Tracking, B. Sidereal Tracking, and C. IOD. Each section lists the tests, scoring criteria, and download links for the respective datasets. An 'Additional Information' section at the bottom provides details about the current data version, change logs, and contact information.

X-Files Leaderboard Branches Submissions Logout

X-Files: A Data Challenge for Space Domain Awareness

Testing the robustness and accuracy of satellite detection, calibration, and characterization algorithms

The following curated datasets ("branches") are currently available:

- A. Target Tracking
 - Tests the ability to detect satellites using simulated ground-based target tracking images
 - Submissions are scored on the ability to **detect** targets using:
 - x/y position (in pixel coordinates)
 - flux (in electron counts)
 - [Target branch download](#) (1.5 GB)
- B. Sidereal Tracking
 - Tests the ability to detect satellites using simulated ground-based sidereal tracking images
 - Submissions are scored on the ability to **detect** and **calibrate** targets using:
 - x/y position (in pixel coordinates)
 - flux (in electron counts)
 - ra/dec position (in radians)
 - i-band magnitude
 - [Sidereal branch download](#) (9.6 GB)
- C. IOD
 - Tests the ability to make initial orbit determinations from observations over a fraction of an orbital period, using optical, angle-only tracks of various lengths
 - Submissions are scored on the ability to:
 - Predict satellite state at given epoch time
 - Predict downrange location of target after 5, 15, 30, and 60 minutes
 - [IOD branch download](#) (830 KB)

Additional Information:

- Current data version: **v2.2**
- A complete list of changes between data versions can be found in the [change log](#)
- A detailed description of each branch can be found in that branches **README.md** file, including
 - simulation information
 - submission formatting
 - scoring metrics
- Please send any questions, concerns, or suggestions to xfiles@llnl.gov
- Slides describing the challenge can be downloaded [here](#)

© 2021 LLNL | Prepared by LLNL under contract DE-AC52-07NA27344 | LLNL-WEB-823890

Show off your score on the X-Files leaderboard

Leaderboard

■ The leaderboard includes:

- A **Plot** that reflects the overall high score, per branch, as a function of time
- A **Table** featuring the leading score per branch

— Plot tools:

- Hovering over points reveal score, competitor, and data associated with each high score
- Hovering over a plot brings up an interactive toolbar for manipulating plots
 - Save, zoom, pan, lasso, etc.
- Clicking a branch on the legend hides those scores



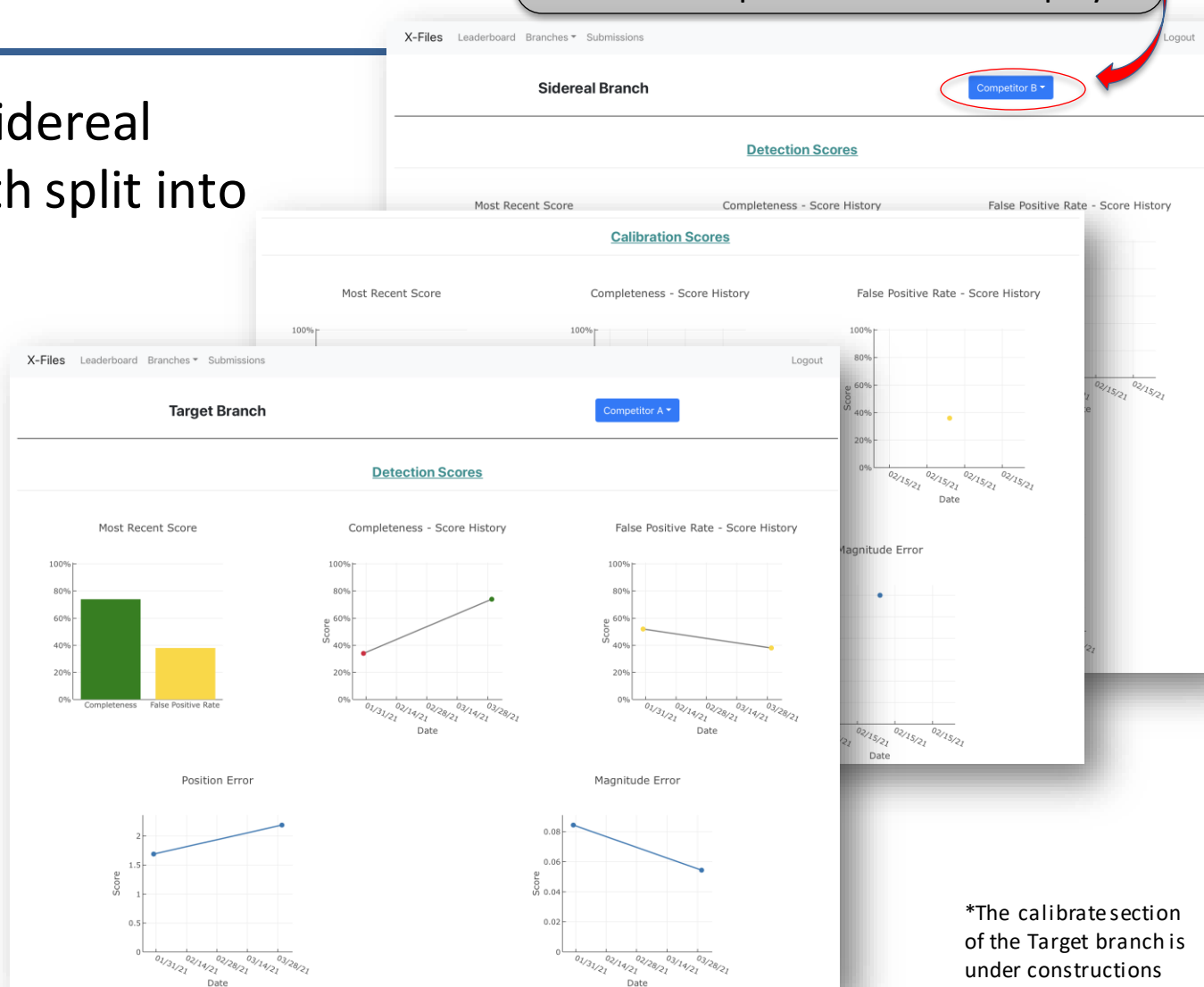
Branch Tab: Target/Sidereal

ALL branch tabs have a competitor drop-down menu in which you can select which competitor scores to display

- The Target and Sidereal branches are both split into 2 sections:

- **Detection**
 - x/y/flux
- **Calibration**
 - Ra/dec/mag

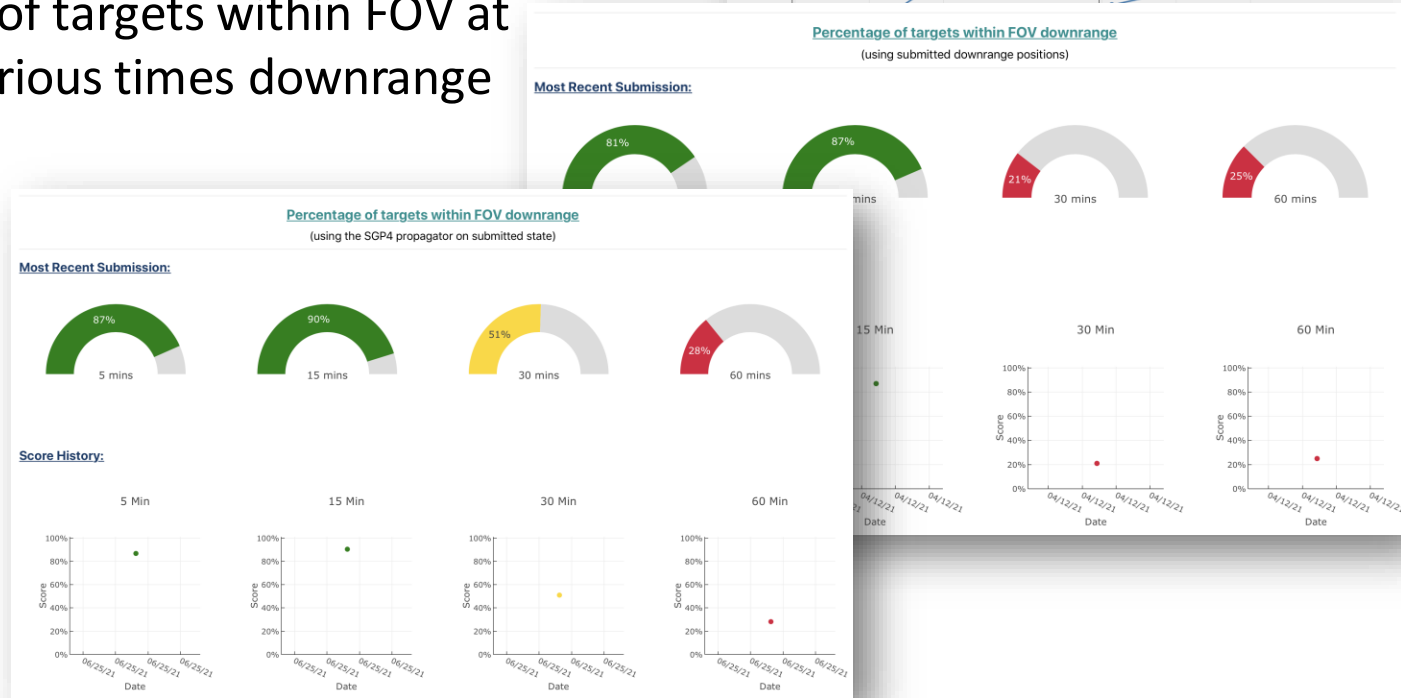
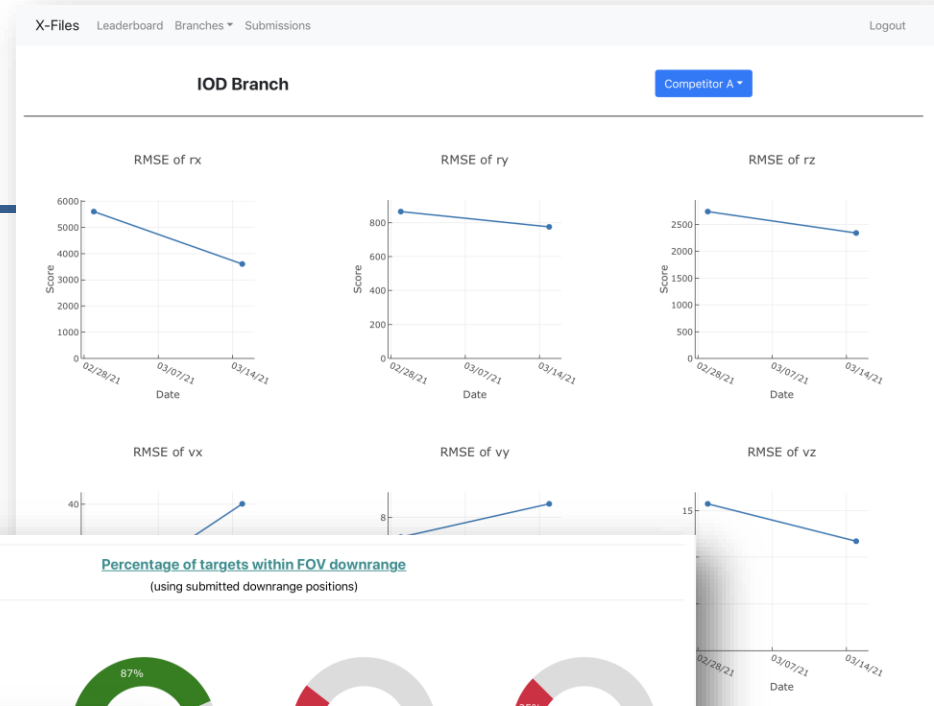
- Plots include:
 - Completeness
 - FPR
 - RMSE errors



*The calibration section of the Target branch is under constructions

Branch Tab: IOD

- The IOD tab contains plots for visualizing various IOD scores:
 - RMSE errors
 - State estimate accuracy
 - % of targets within FOV at various times downrange



Participation

- Who can participate?
 - Anyone can participate!
- How to participate?
 - Expressing interest:
 - If you are interested in competing in the X-Files data challenge, send an email to xfiles@llnl.gov
 - Accessing the challenge:
 - Once your participation is approved, you will receive login information for <https://xfiles.llnl.gov>
 - Download the data and begin testing your algorithms
 - Submit a score to populate onto the dashboard
- Why should you participate?
 - Gain attention for your SDA algorithms
 - Earn a government contract or financial reward*

*Rules apply

Coming Soon

- "Full" branch
 - Raw images --> orbits
 - Each target will have a sequence of images
 - Detect, calibrate, and perform IOD all from raw images
- Split branches by difficulty level
 - Target/Sidereal branches will have varying numbers of targets in each image
 - IOD branch will be further broken down by track length
 - "Full" branch will have maneuvers of varying difficulty to track / disambiguate
- Dashboard features
 - Auto-scoring done through dashboard
 - Info tab describing each scoring metric
 - Interactive table allowing the user full control of data



Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.