Defining Airspace Complexity

Machine Learning for Airspace Complexity Management

Presented by: Daniel Weckler

Bryan Matthews, Nikunj Oza, Shayan Monadjemi, Shawn Wolfe, Hossein Eskandaripour

Motivation

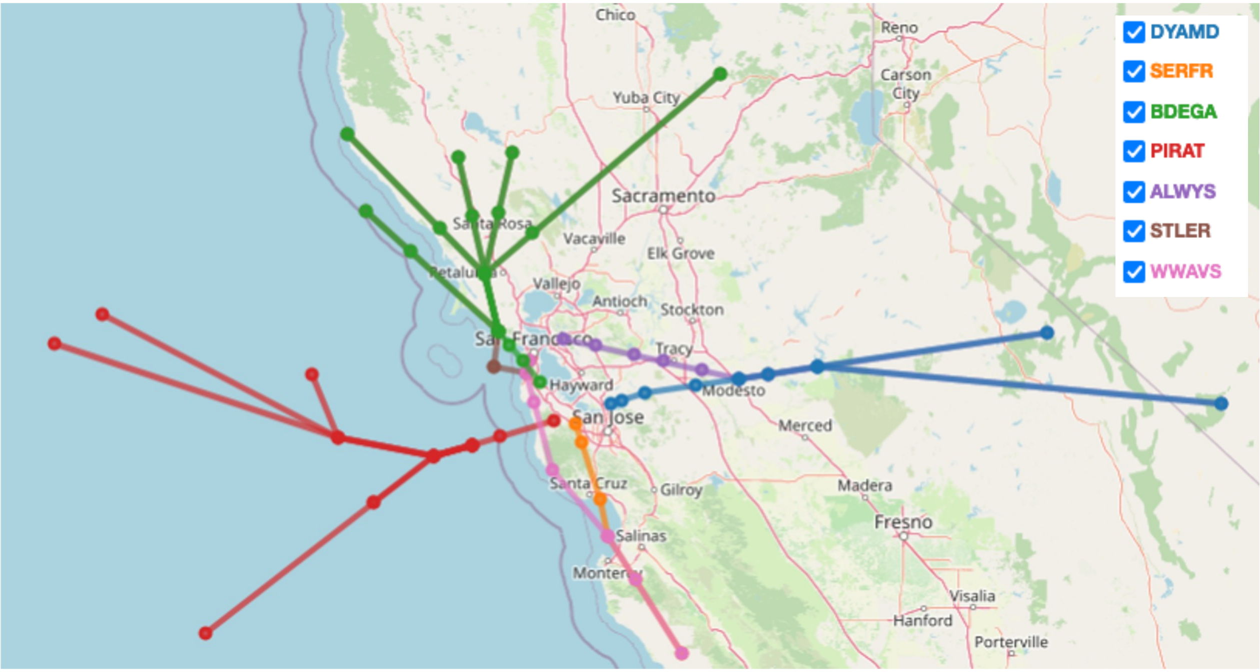
- Design a measure to capture high airspace complexity.
- Identify factors that lead to this high airspace complexity
- Makes things easier for air traffic controllers by being able to anticipate how complex the airspace will be and prepare appropriate response.
  - Complexity is analogous to controller workload
  - More communications = higher complexity

Data Available for the Project

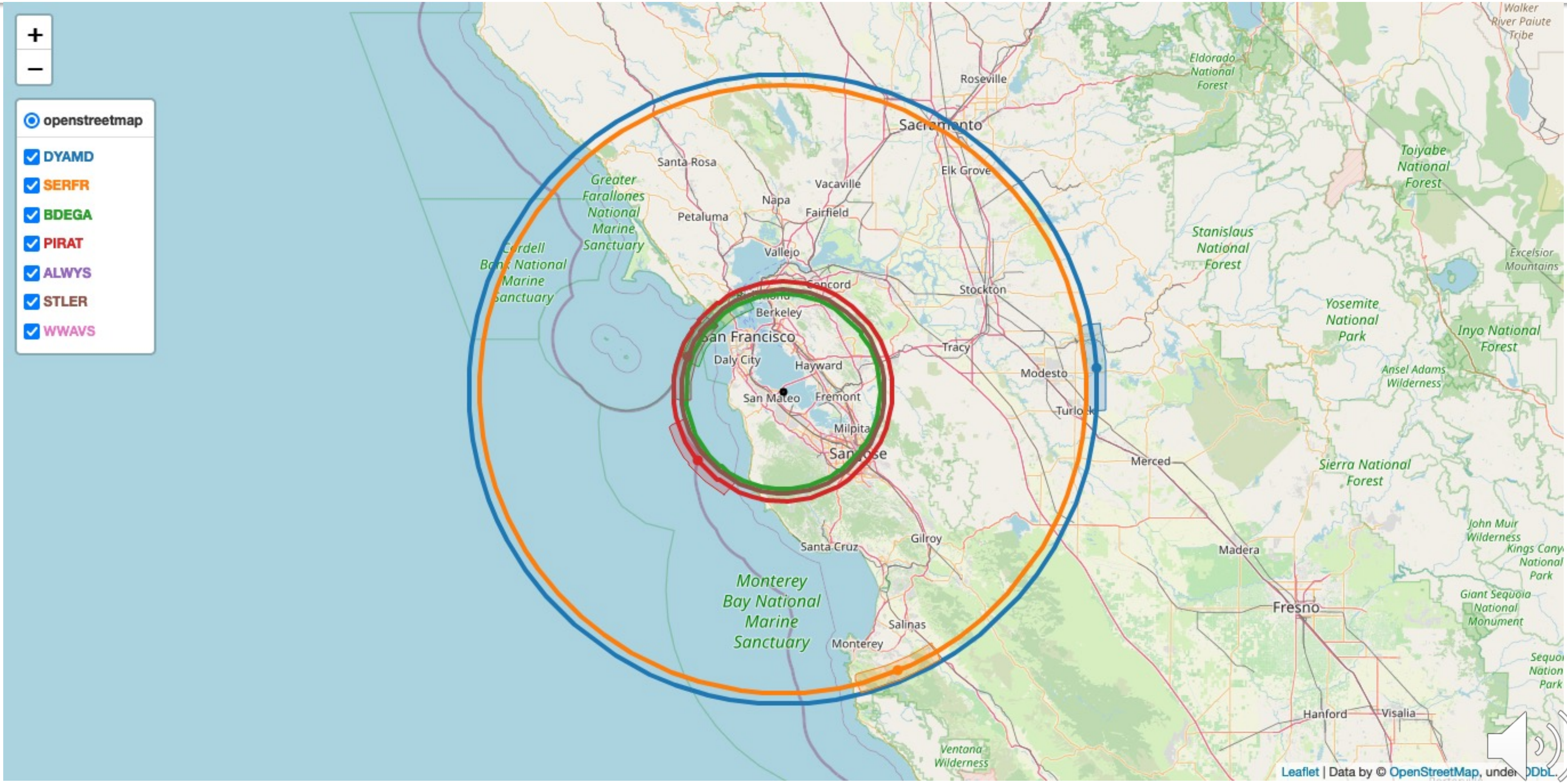
- Sherlock Data
  - Integrated Flight Format (IFF) Files
    - Real time data
      - Latitude/Longitude coordinates
      - Unequal time steps
    - Very large
  - Reduced Data (RD) Files
    - Flight parameters
      - Top of Climb/Descent
      - Departure/Arrival times
      - Flight plans
      - Origin/Destination Airport
      - Destination/runway
    - One entry per flight (smaller filesize)
- Sherlock Data (cont.)
  - Event Data (EV) Files
    - Events that occur throughout the flight (from takeoff to landing)
    - Standard event markers (touchdown, top of descent, etc.)
    - Safety events (go arounds, etc.)
- Other Files:
  - METAR
    - Ground-based weather data
  - ASPM
    - Aviation system performance metrics
    - Arrival Rate
    - Efficiency score
    - Computed after the fact (data would not be available in real time)

STAR Routes (Standard Terminal Arrival Routes)

- Each aircraft (usually) files for one RNAV STAR route in their flight plan (Field10 in the RD data).
- They are the procedures that controllers expect pilots to follow
- Guides the flights to a point where the TRACON controllers can then sequence the arrivals for their appropriate runway
- These star routes combined with the final approach fix also determine what we define as the maneuvering area for the flight.
  - Flights no longer on the STAR, but are not yet lined up on final approach.

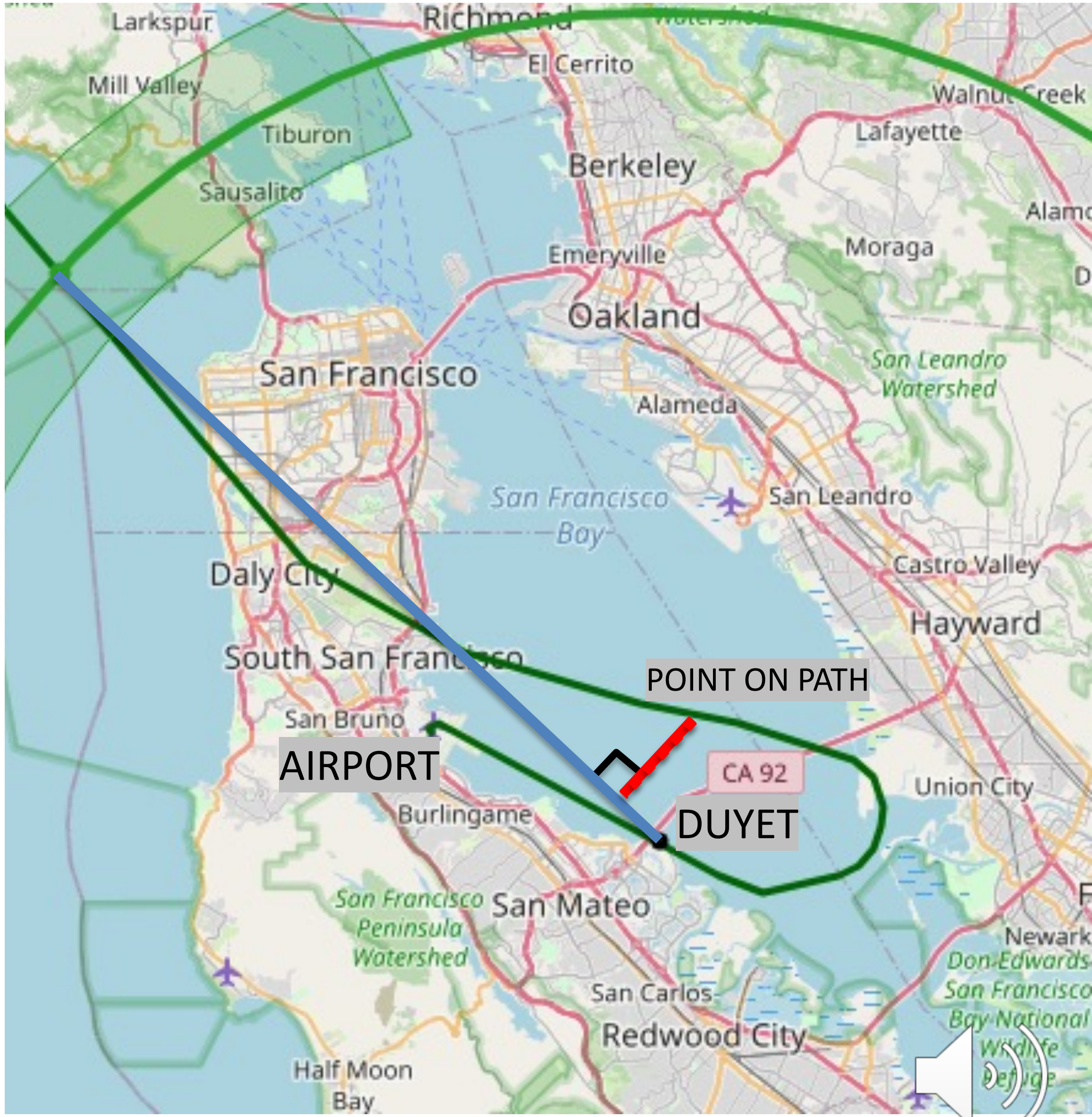


Maneuvering Areas (Runway 28L)

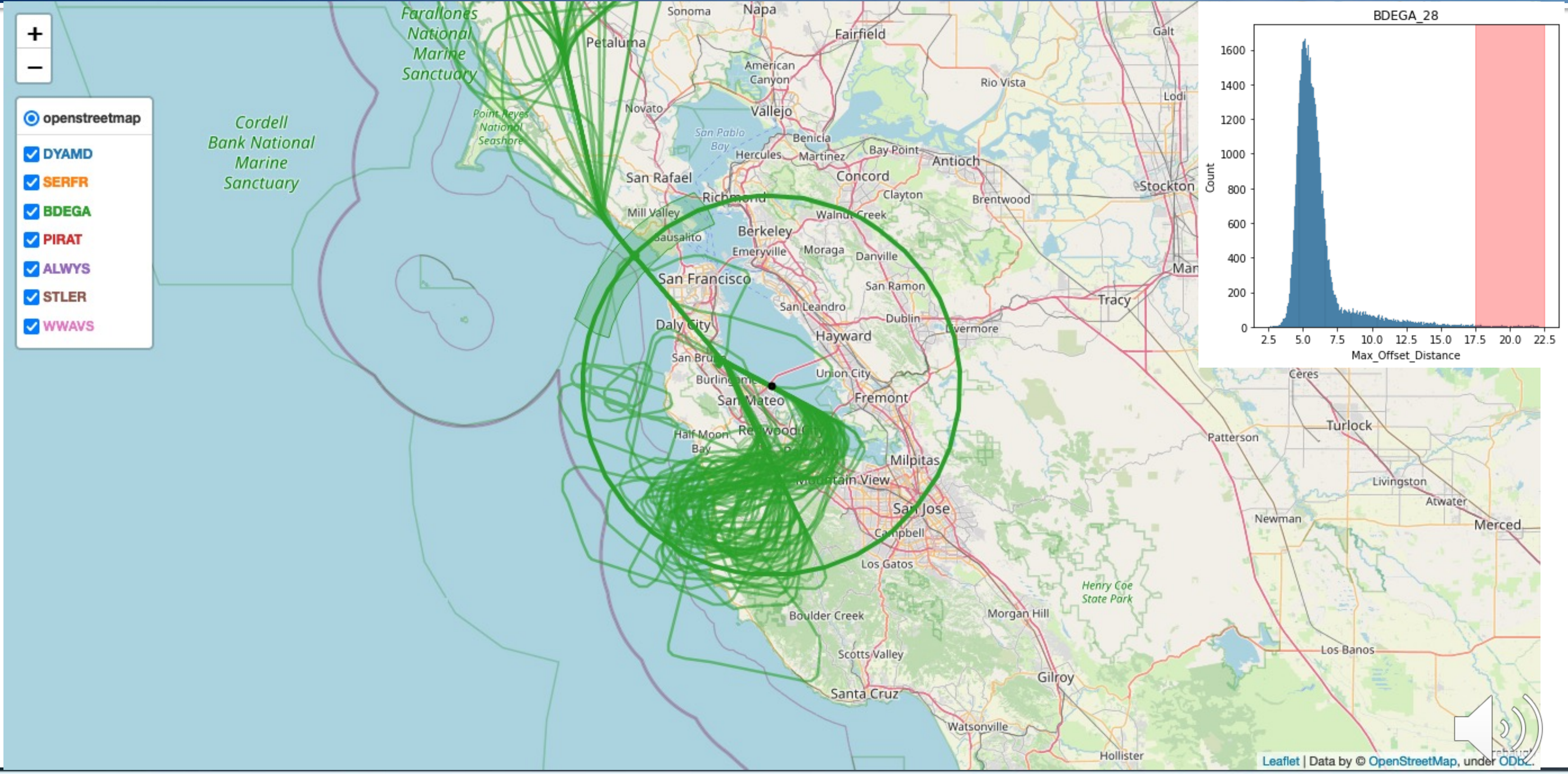


Lateral Offset Distance

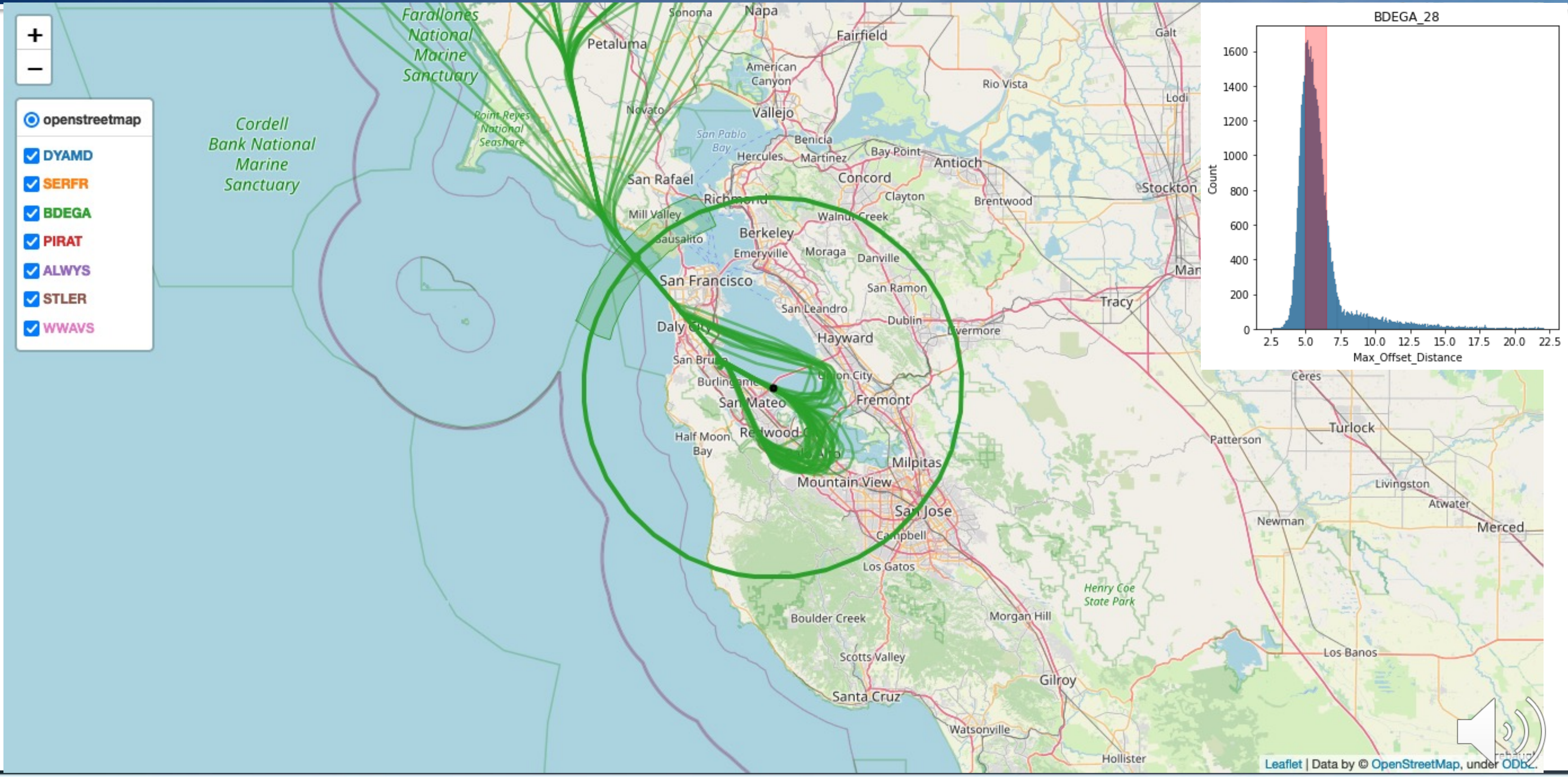
- Direct path theory:
  - Straight line between the star fix and the final runway fix (in this case, it's DUYET)
  - Deviation from this path = more complex
- All STAR-runway pairs have different deviation distributions, so we characterize the offsets based on their position in the distribution rather than the offset itself
- We initially considered using this for our complexity metric



Fringe Sampled Flights

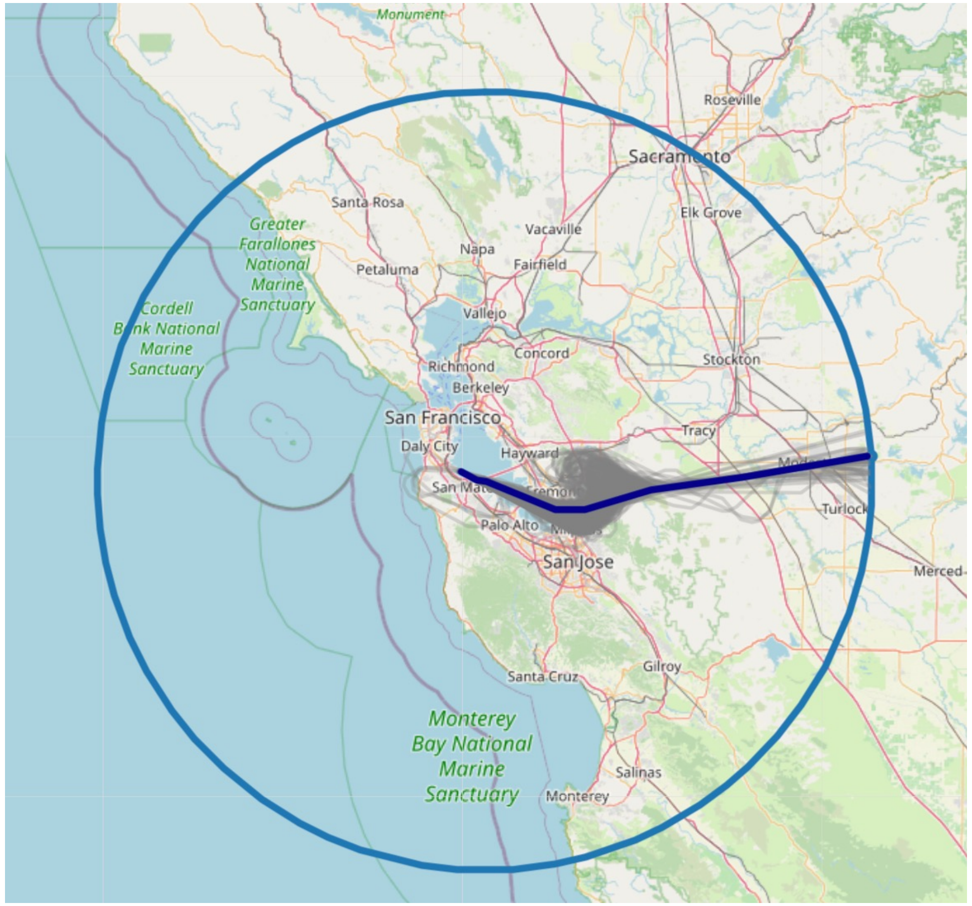


Median Sampled Flights



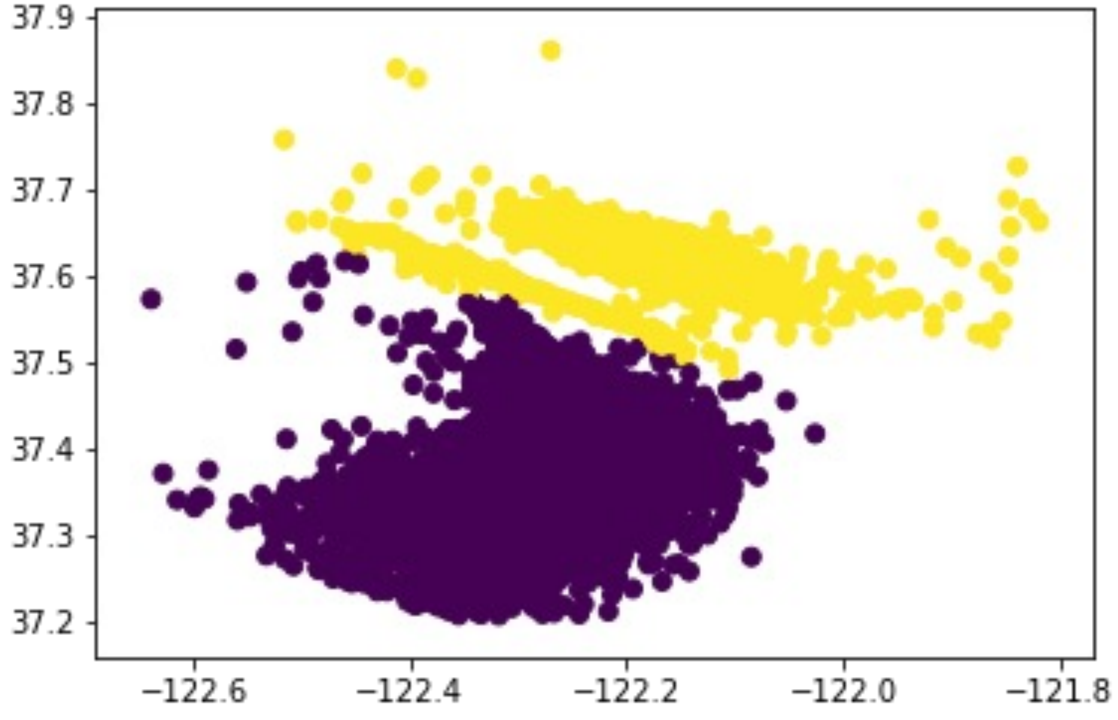
Refining Our Distributions: Representative Path

- Each flight/runway configuration has a flight trajectory that most flights will more or less follow in the absence of other traffic
- Sample flights from the median of the offset distance distribution
- From here we can collect offset distance distributions much like we did for the prior metric.
- Utilizing this, we can also predict which flights will have conflicting landing schedules

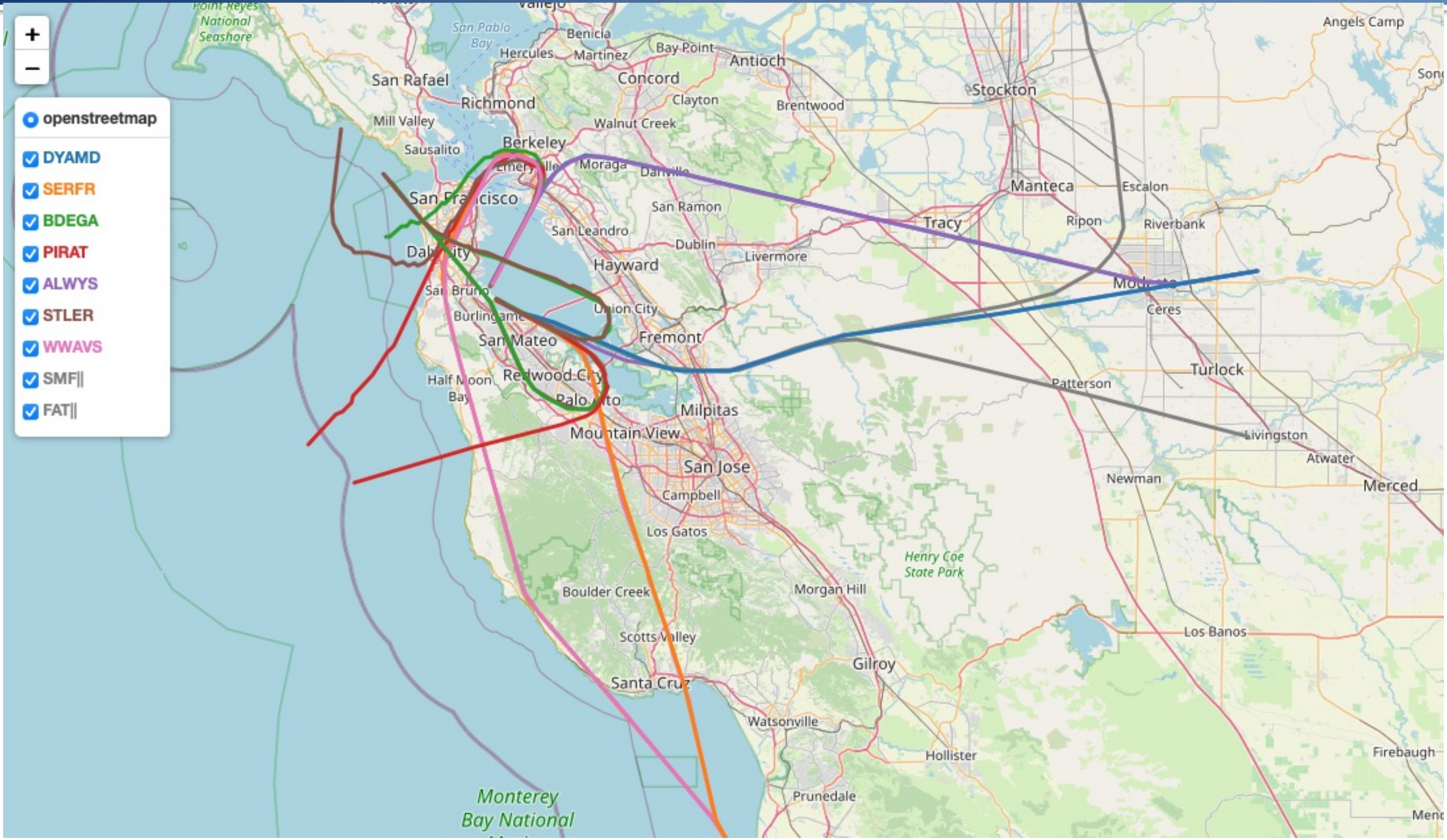


Split Paths

- Some routes (BDEGA and STLER) take two different paths to get to the airport
- This isn't documented anywhere (to our knowledge)
- Fortunately we're able to use K-means clustering to separate the two paths and create a representative path for each one

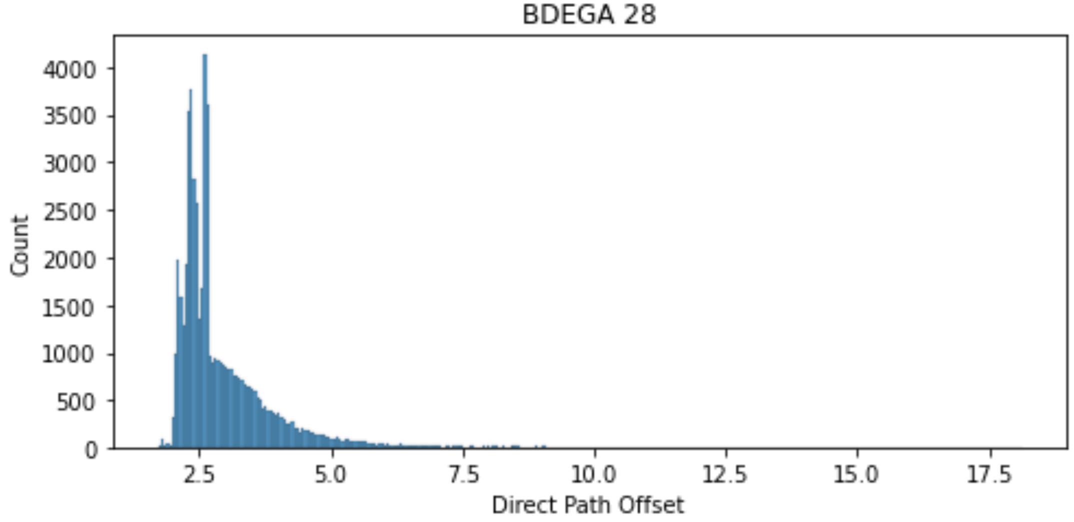


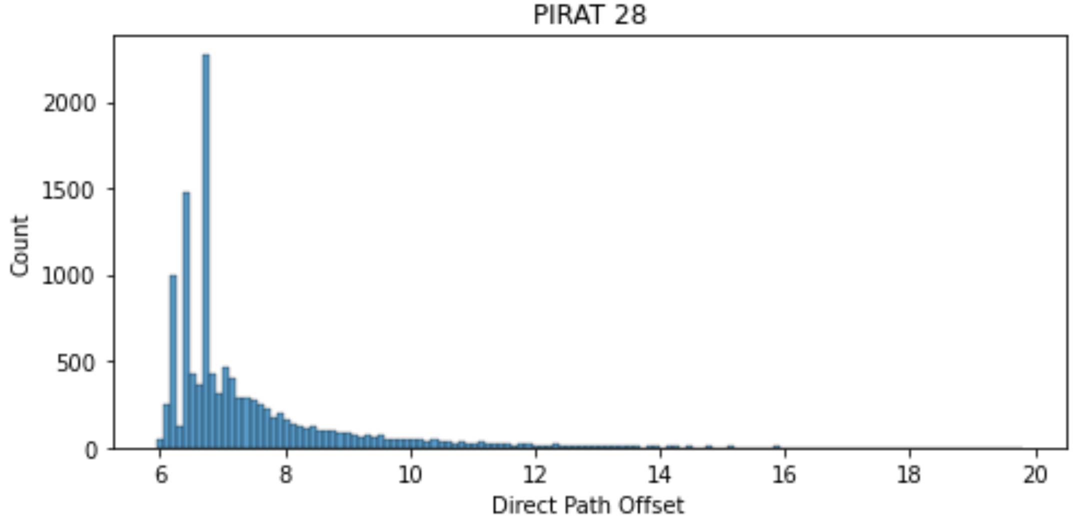
All Representative Paths



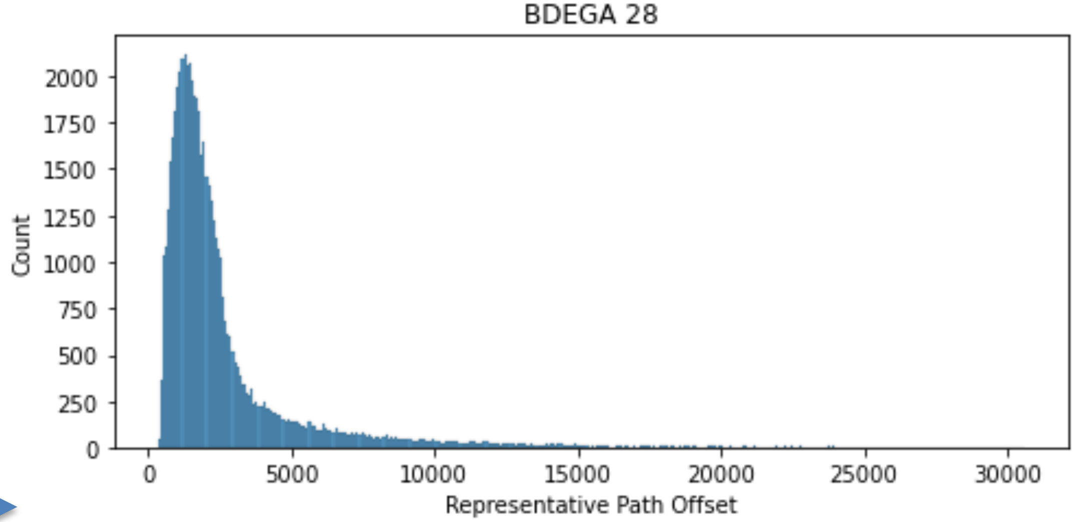
Distribution Comparison

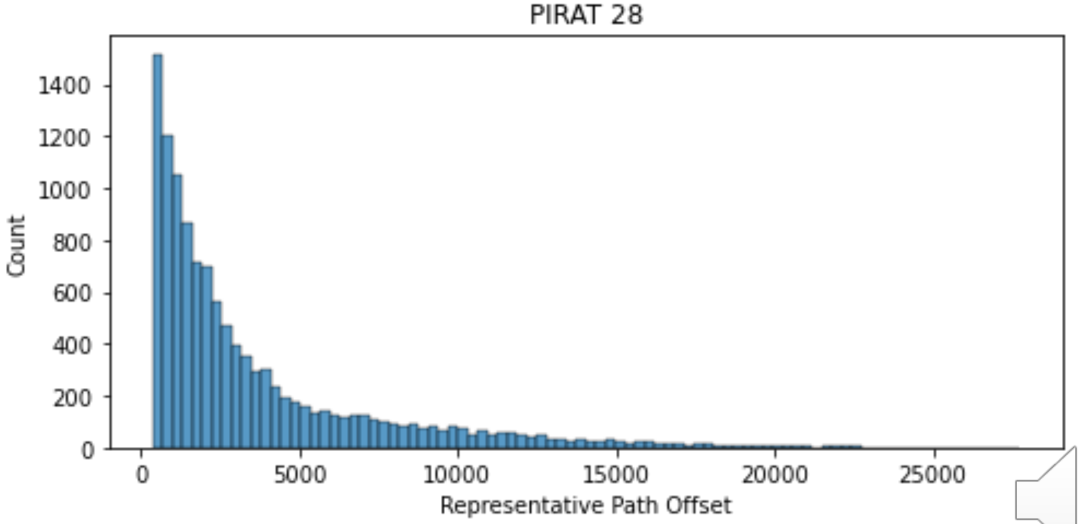
Direct Path





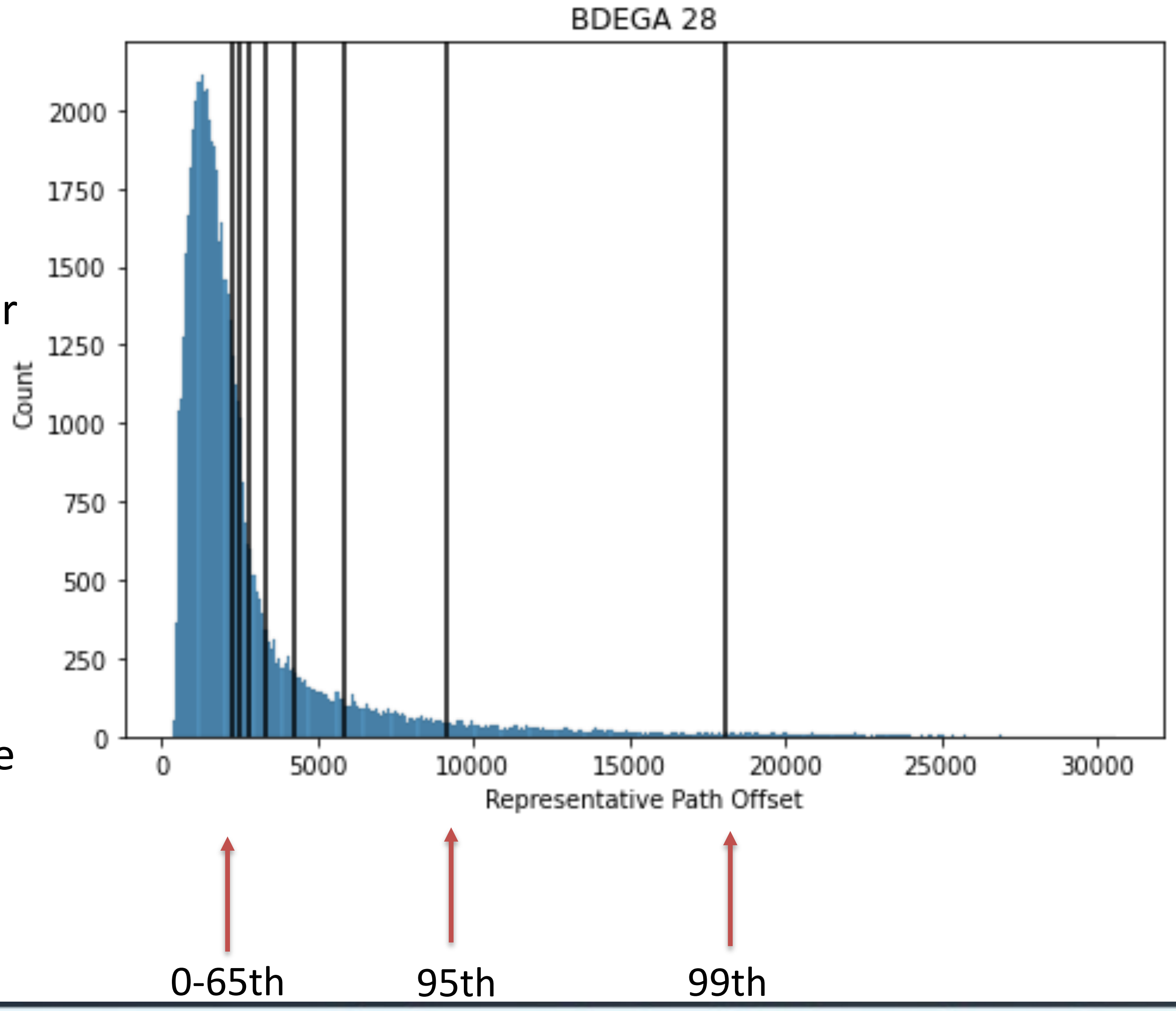
Representative Path





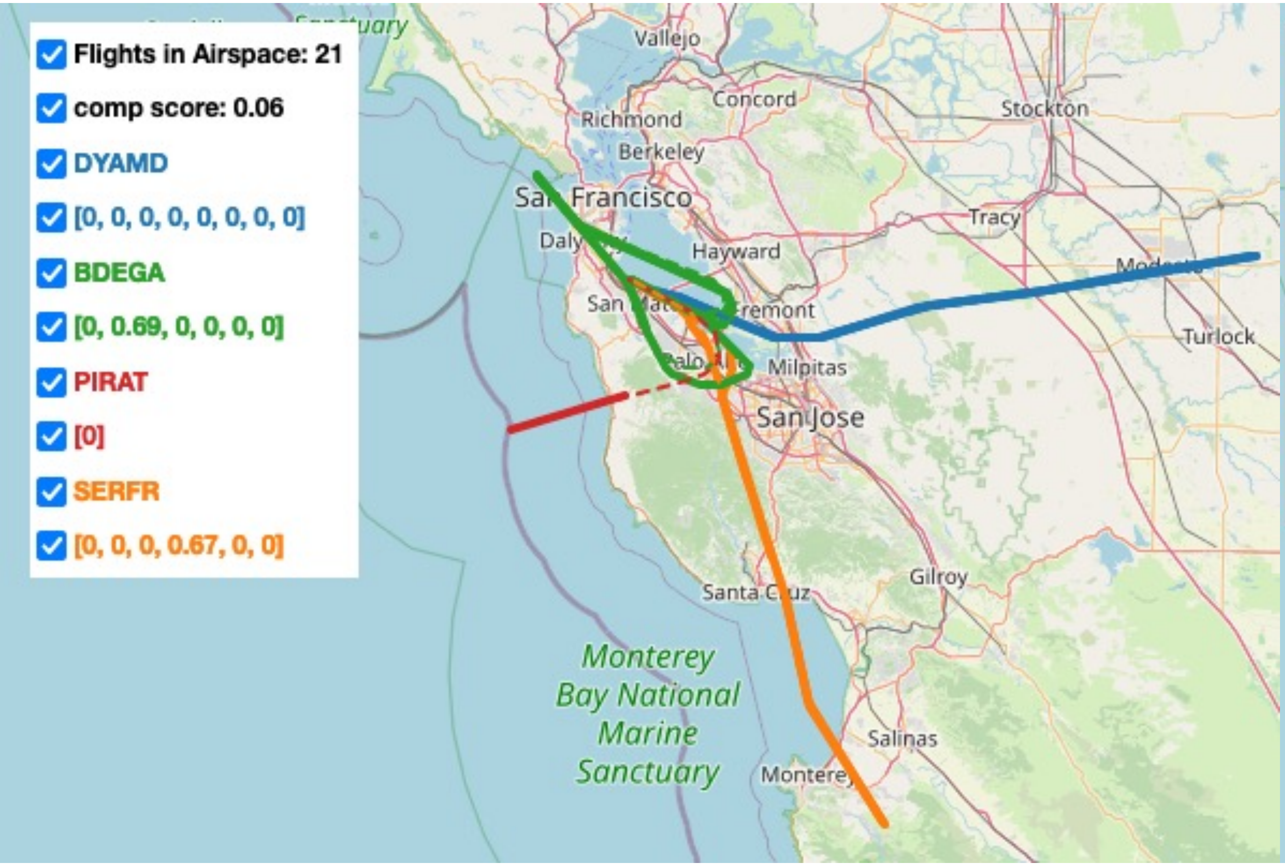
Complexity Score Methodology

- Use a statistical measure of the offset distance from the representative path
- For each 15 minute sample
  - Get all the flights that are currently in their respective maneuvering area during each window
  - Map each flight's offset into percentiles (*65th-95th)* for each runway/star combination
  - The percentile that the flight falls into is the complexity score for that flight
  - The overall complexity score is the average of all the flights.

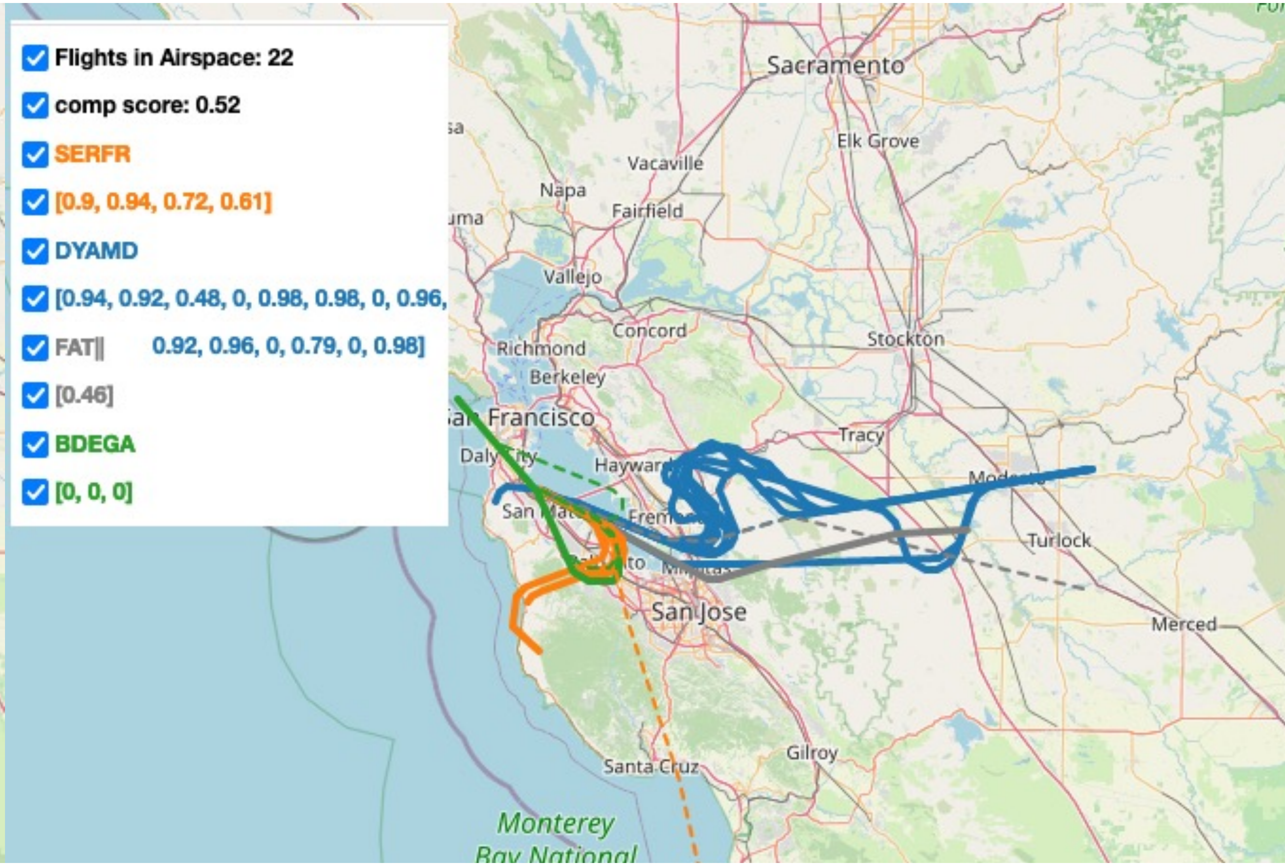


Flight Track Maps

Nominal Snapshot



Anomalous Snapshot



Constructing the Dataset for ADOPT

- Dataset Construction
  - 1 hour windows
  - Event within 15 minutes after the window
  - Nominal Threshold: 30th percentile
  - Anom Threshold: 85th percentile
- Features (Measured)
  - Available runways
  - Runway configuration
  - Visibility
  - Go arounds
  - Temperature
  - Flights in airspace
- Alternate complexity measures
  - Measure α
    - Clustering Maneuvering Points
  - Measure β
    - Equating complexity with fixed transition likelihood
  - Star Entropy
    - Measure of the heterogeneity of the arrivals
- Derived features
  - Ratio of *flights on representative path* vs *total flights in airspace*
  - Recent runway switch
  - Number of conflicts

Model Results

```
Training Samples: 832
 - Nominal: 416
 - Adverse: 416
Validation Samples: 556
 - Nominal: 278
 - Adverse: 278
Test Samples: 156
 - Nominal: 78
 - Adverse: 78

Epochs: 200
Lambda: 0.01
Dropout Rate: 0
Batch Size: 32
```

```
AUC Train: 0.9179816292802032
AUC Test: 0.9132149901380671
Precision: 0.8701298701298701
Recall: 0.8589743589743589
F1 Score: 0.8645161290322582

Confusion Matrix:
[[68 10]
 [11 67]]

Guide:
|TN,FP|
|FN,TP|
Positive = Anomalous
Negative = Nominal
```

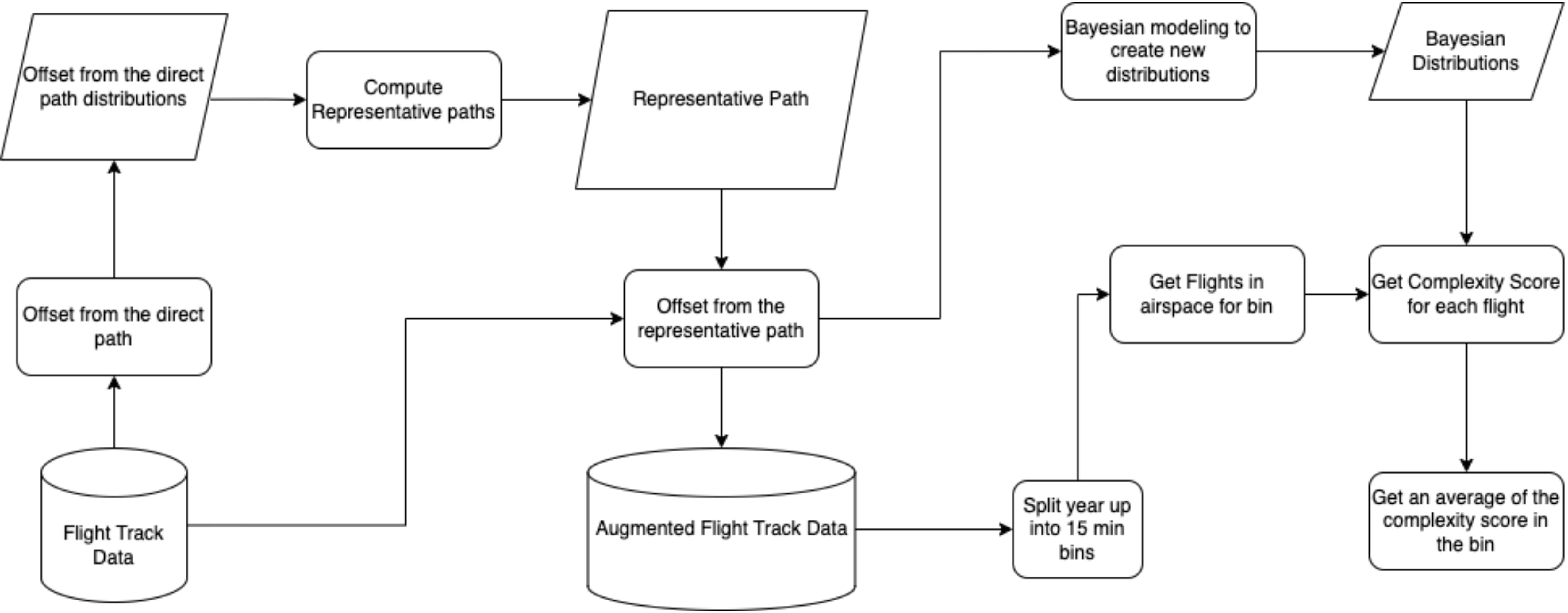
Clustering Table

Cluster ID	Cluster Size	Off Path Ratio	STAR Entropy	# of Conflicts	Measure α	Measure β	Go-Around	Runway Config	Visibility
0	53	✓	✓	✓	✓	✓			
1	92	✓	✓	✓	✓		✓		
2	91	✓	✓		✓				
3	114	✓	✓	✓	✓		✓		
4	125	✓	✓		✓				
5	125	✓	✓	✓	✓				
6	81	✓	✓	✓	✓		✓		
7	6	✓	✓		✓		✓	✓	
8	8	✓	✓	✓	✓				✓
Overall % impact	All	55.3%	12.7%	6.8%	7.7%	4.1%	4.5%	0.2%	0.4%

Conclusion/Future work

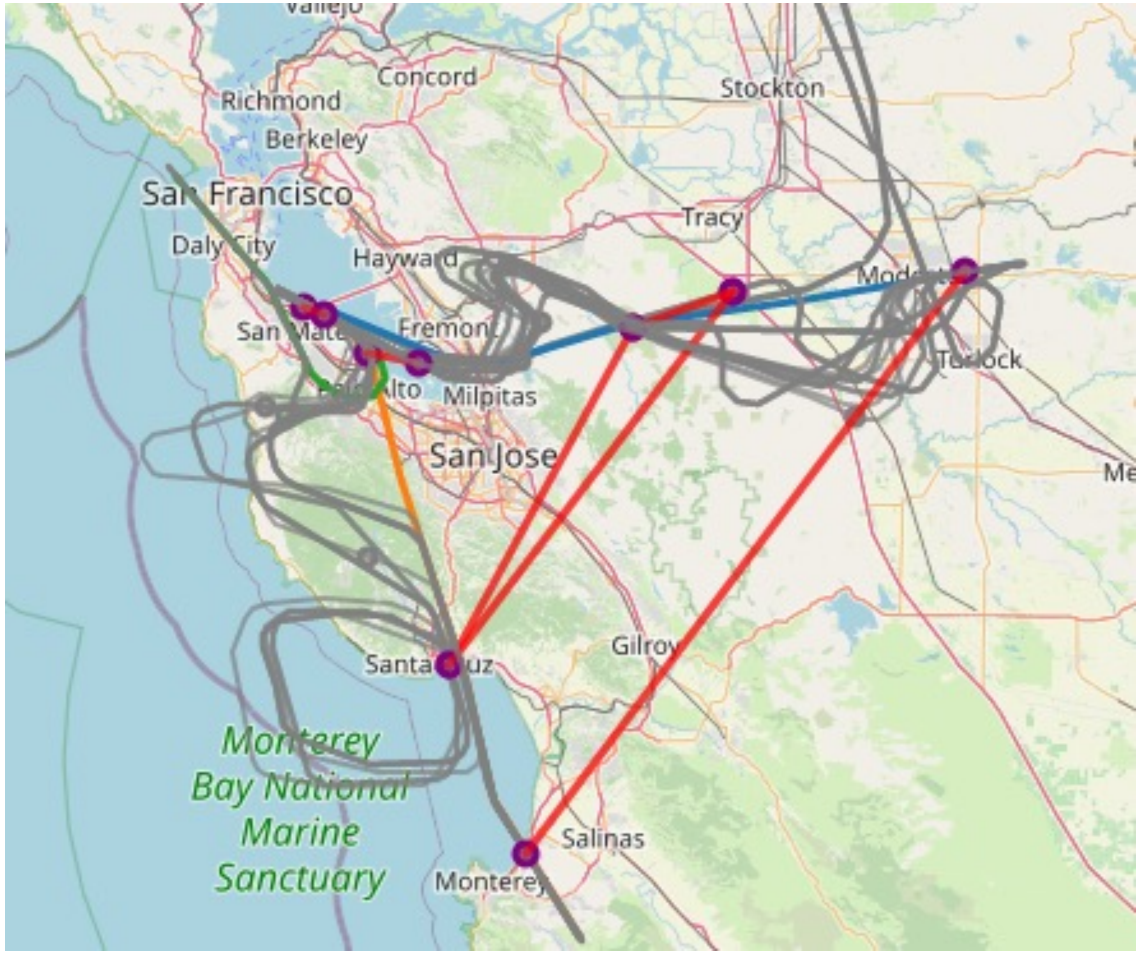
- Created a novel method for computing airspace complexity
  - Using domain knowledge from STAR paths
- Found possible precursors to a high complex airspace
- We want to improve characterizing star-runway pairs with low data
  - Distributions
  - Representative path
- Enhance the complexity score to be less sensitive to small deviations
  - Many small deviations currently can mark a low complexity time as complex
  - Improve weighting average

Appendix A: Complexity Score Pipeline



Appendix B: Conflicts

- When two or more flights are scheduled to arrive at or around the same time
- Preventative measures must be taken to slow the flights down so they arrive at different times
  - S-turns
  - Holding patterns
- These result in an offset from the optimal path, thus increasing complexity



Appendix C: STAR Entropy

- Theory: The less variations of STARs being used at the current time, the less complex the airspace is for the TRACON controllers (since they have to sequence less flights)
- Formula (scaled by total star routes)
  - proportions = (star_routes_used / total_flights_in_routes)
  - S = - Σ(proportions * log(proportions)) * total_paths
- A lot of the controller workload happens after the star route, and this approach doesn't characterize that
  - This only categorizes flights on the route
  - Use as a feature for complexity rather than complexity itself.