Machine Learning Based Terminal Descent Predictions

James "Zach" Wells

Mentors: Krishna Kalyanam[1], Tejas Girish Puranik, Manish Kumar[2]

Code: A

Spring 2022

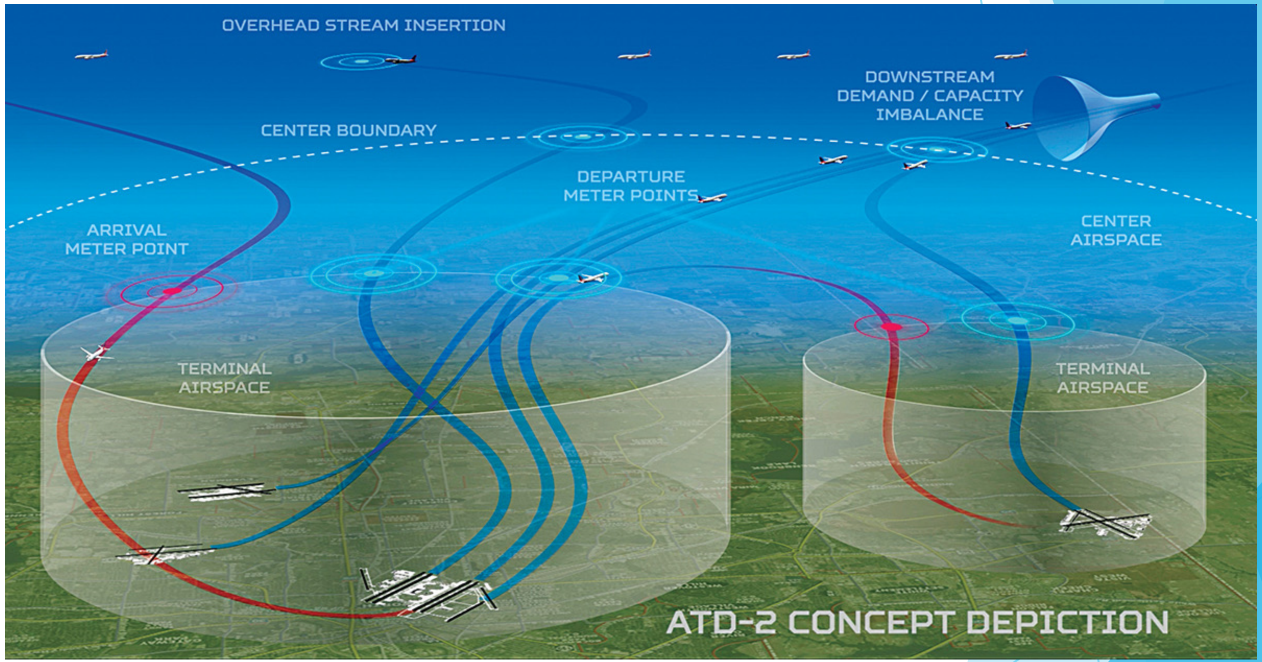
[1]Nasa Ames Research Center [2]University of Cincinnati

About Me

- FIRST LEGO League and FIRST Robotics inspired me to get into robotics
- Received my BS-ME at University of Cincinnati in 2018
- Joined Cooperative Distributed Systems lab at The University of Cincinnati in Summer of 2018
- Received MS-ME at University of Cincinnati in 2021
- Pursuing PhD in ME at University of Cincinnati
- Areas of Research include:
  - Path Planning/Traffic Management through path prediction
  - Advanced Platform Design for Unmanned Aerial Systems
  - GPS Denied Navigation and State Estimation and Control

Next Generation Air Transportation System: “NextGen”

- Collaboration with the Federal Aviation Administration (FAA) and NASA
- Goal to increase autonomy and develop new tools for pilots, air traffic controllers and other airspace users
- Will result in an increase in safety and efficiency of airspace usage
- Programs within NextGen that Aviation Systems Division contributes to include:
  - Modeling and Simulation of airspace
  - Tactical Air Traffic Management (ATM)
  - Strategic ATM
  - Surface Automation
  - Historical Foundations for ATM Research



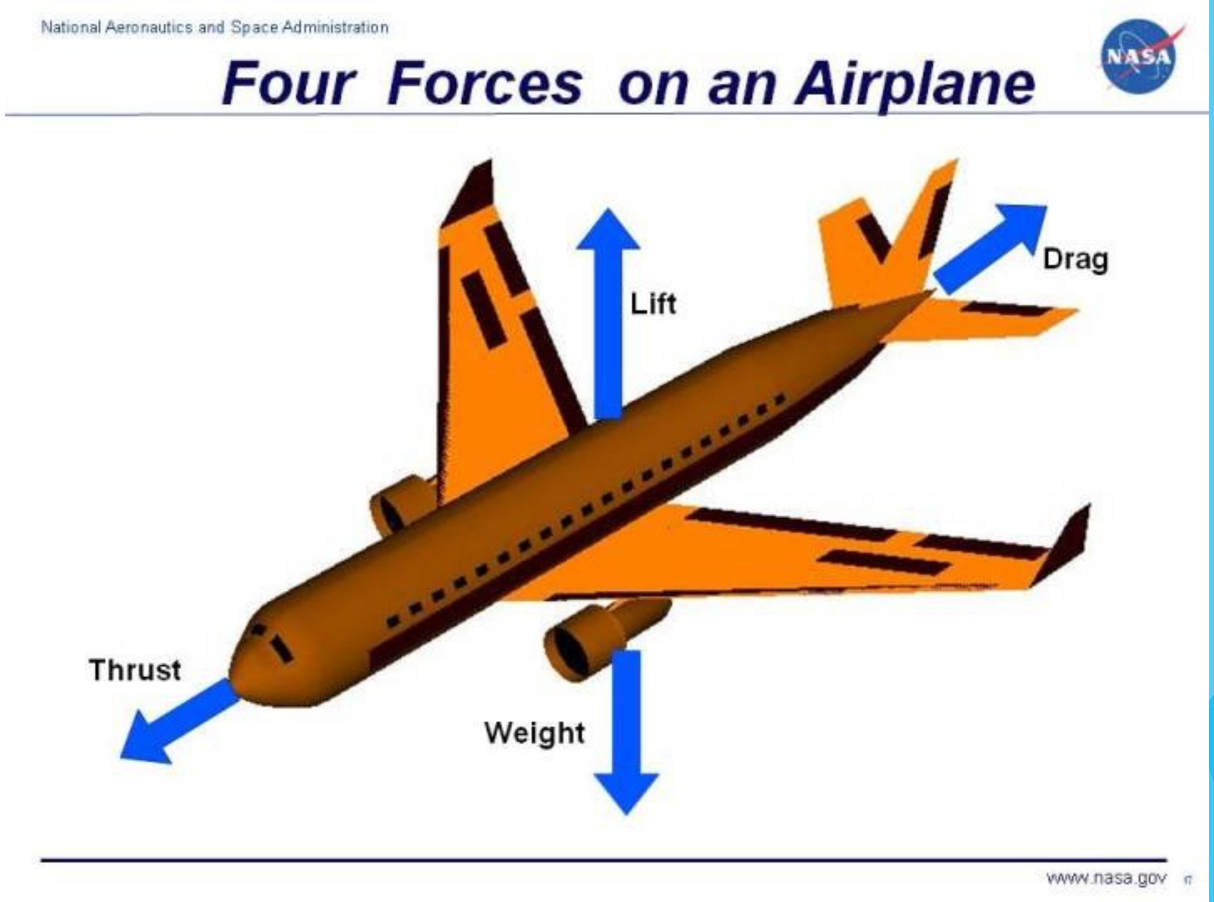
https://aviationsystems.arc.nasa.gov/research/atd2/index.shtml

This graphic shows NASA’s Integrated Arrival, Departure and Surfaces (IADS) concept

https://aviationsystems.arc.nasa.gov/research/nextgen_atm_research.shtml#:~:text=%E2%80%9CNextGen%2C%E2%80%9D%20short%20for%20the,of%20aviation%20on%20the%20environment.

Problem Statement

- In order to increase efficiency and safety, we need to find a way to better predict aircraft trajectories and time of arrivals
- Current physics-based models used are based on aircraft profiles
- Not all aircraft have corresponding profiles
- Physics based estimates rely on many variables (drag coefficients, weight, etc.) which are either not easily measurable or are not provided by aircraft manufactures
- Physics based parameters and forces seen in the diagram are oversimplified and are dynamic and coupled



https://www.grc.nasa.gov/www/k-12/airplane/forces.html

Desired Outcome

- This work focuses on the airport and not the aircraft itself
- Want to study the airspace and traffic patterns around a given airport using Machine Learning Techniques
- Want a neural network to capture parameters physics model can't incorporate
  - Airline procedures
  - Airport dependent procedures
  - Mitigates pilots influenced commands
- From the patterns identified using clustering algorithm, want to fuse IFF data with weather data and other meta data with a neural network to predict time of arrival and trajectory

Literature Survey

Trajectory Clustering

- Clustering algorithms have been used for ship, aircraft and automobile trajectory grouping
- Two prominent clustering algorithms used for Trajectory clustering include Density Based Spatial Clustering (DBScan) and Hierarchical DBSCAN (HDBSCAN)
- Trajectory clustering has been used to solve traffic management, conflict detection and optimization-based problems

Neural Network Prediction

- Long-Short Term (LSTM) Neural Networks have been used to predict aircraft states and behavior during final approach
- Prediction work for ETA has been done using a Recurrent Neural Network for entire flight
- A time/path prediction LSTM neural network has been completed, however it did not incorporate weather data

Planned Steps

Step 1: Flight Data Collection → Step 2: Flight Area Reduction → Step 3: Data Cleaning → Step 4: Data Resampling → Step 5: Data Clustering and Preliminary Analysis → Step 6: Weather Data Collection → Step 7: Neural Network Training

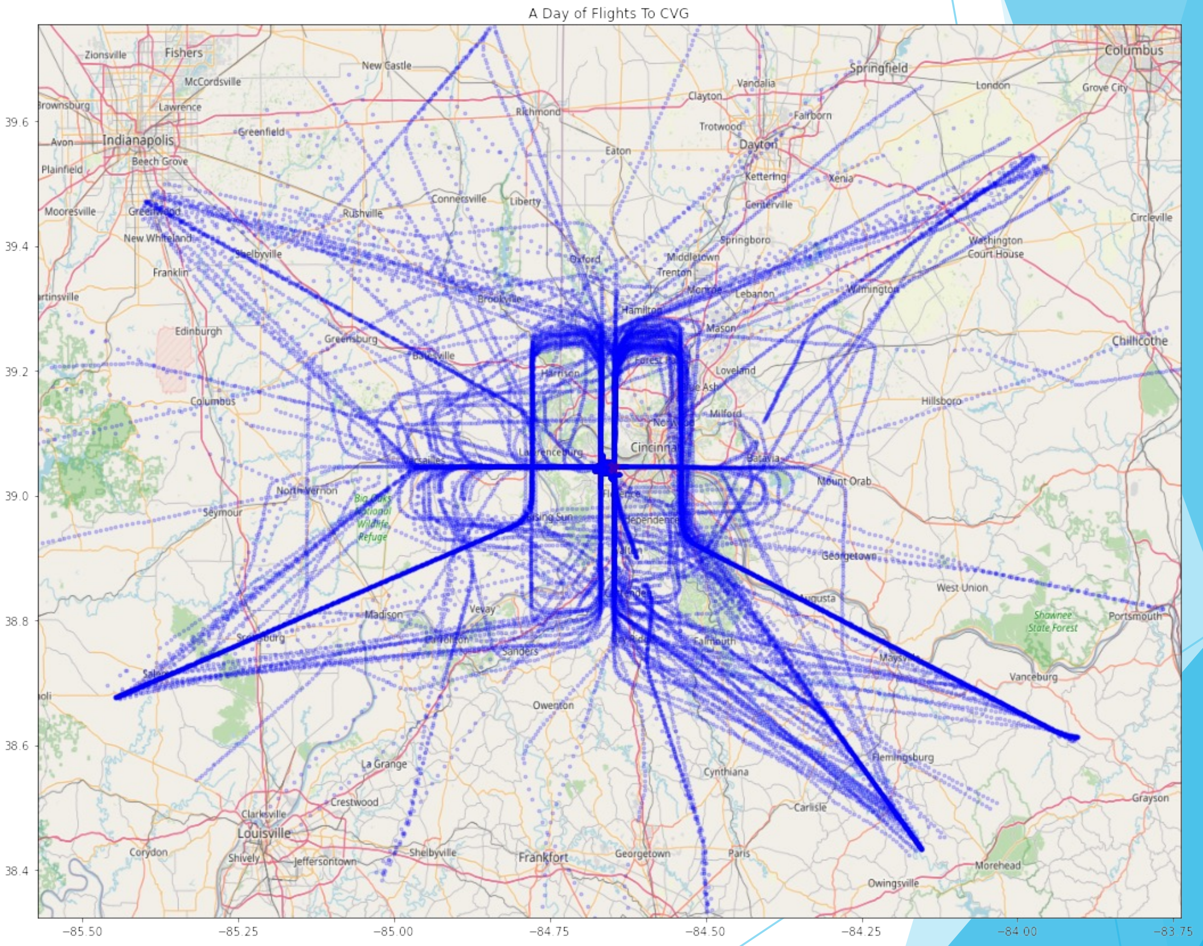
Data Source

- Focused on Cincinnati-Kentucky International Airport (CVG) with flight data from January 1, 2020, through January 14, 2020
- Data accessed through Sherlock Data Warehouse
- Using Track Point (IFF) Data
- To get the fused set of track point data, you must download all flights for a given day
- Results in very large files (~6 GB csv files) which cannot be opened in Excel

https://sherlock.datawarehouse.arc.nasa.gov/data_access/

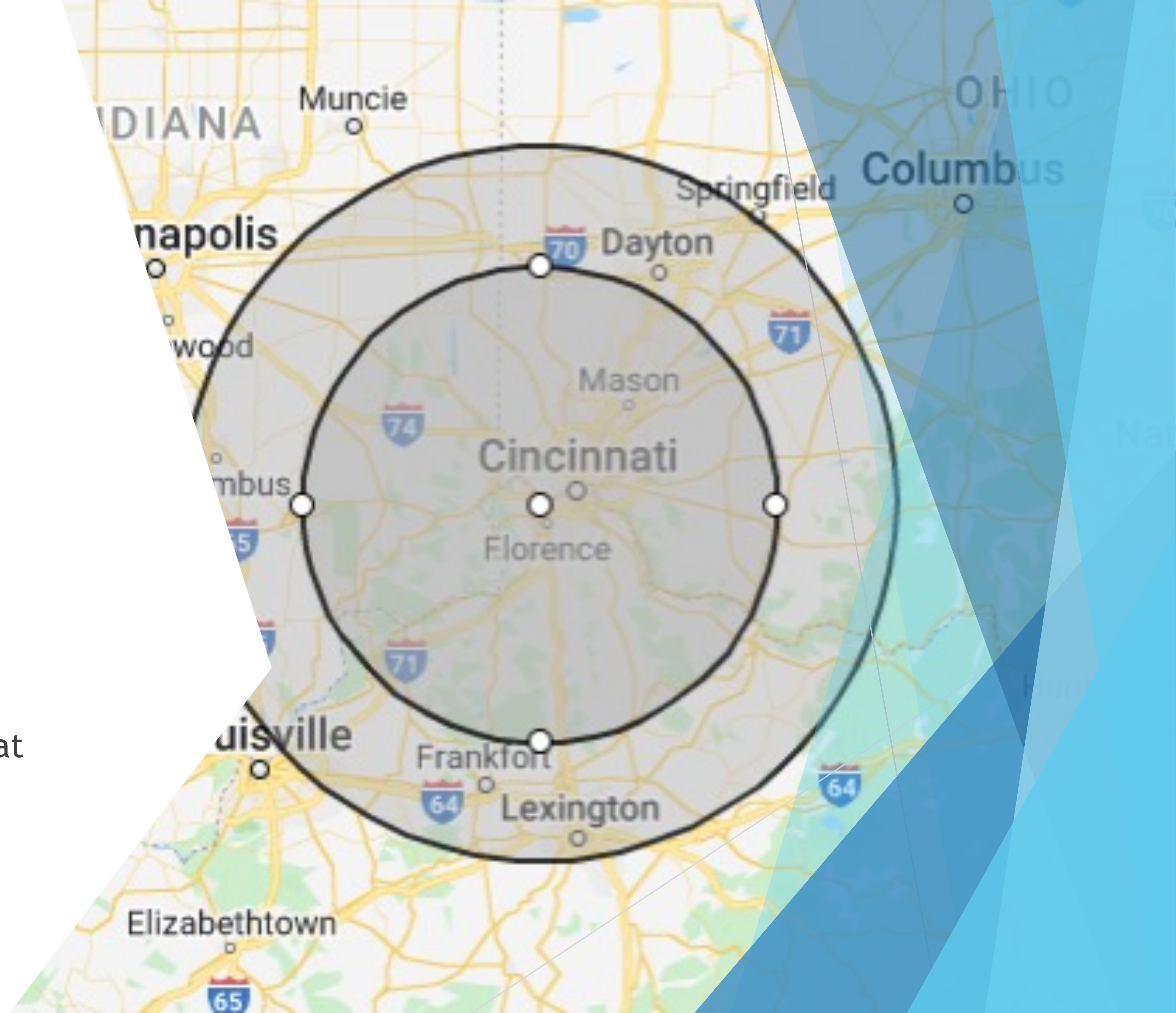
Step 1: Data Collection

- Only wanted flights landing at CVG
- IFF type 4 records were used to identify aircraft taking-off and landing at CVG
- IFF type 3 records were the track record points
- Unique Aircraft ID (UAID) was used to link the two records together
- Wrote code to filter the large CSV files into daily logs for the CVG airport
- Data collected included: UAID, Time, Latitude, Longitude, Altitude, Ground Speed, Course Rate Of Climb
- This code isn't specific to landing at CVG. Can be used for take-off/landings at any airport



Step 2: Reducing Flight to 50 miles

- Full flight needs to be shortened to fifty-mile radius
- Did this by applying haversine algorithm to flight data
- Used coordinates for CVG and incoming flight data.
- Saved data while flight is less than 50 miles
- This isn't specific to landing- can be used for data filtering at any point in the world/any portion of flight

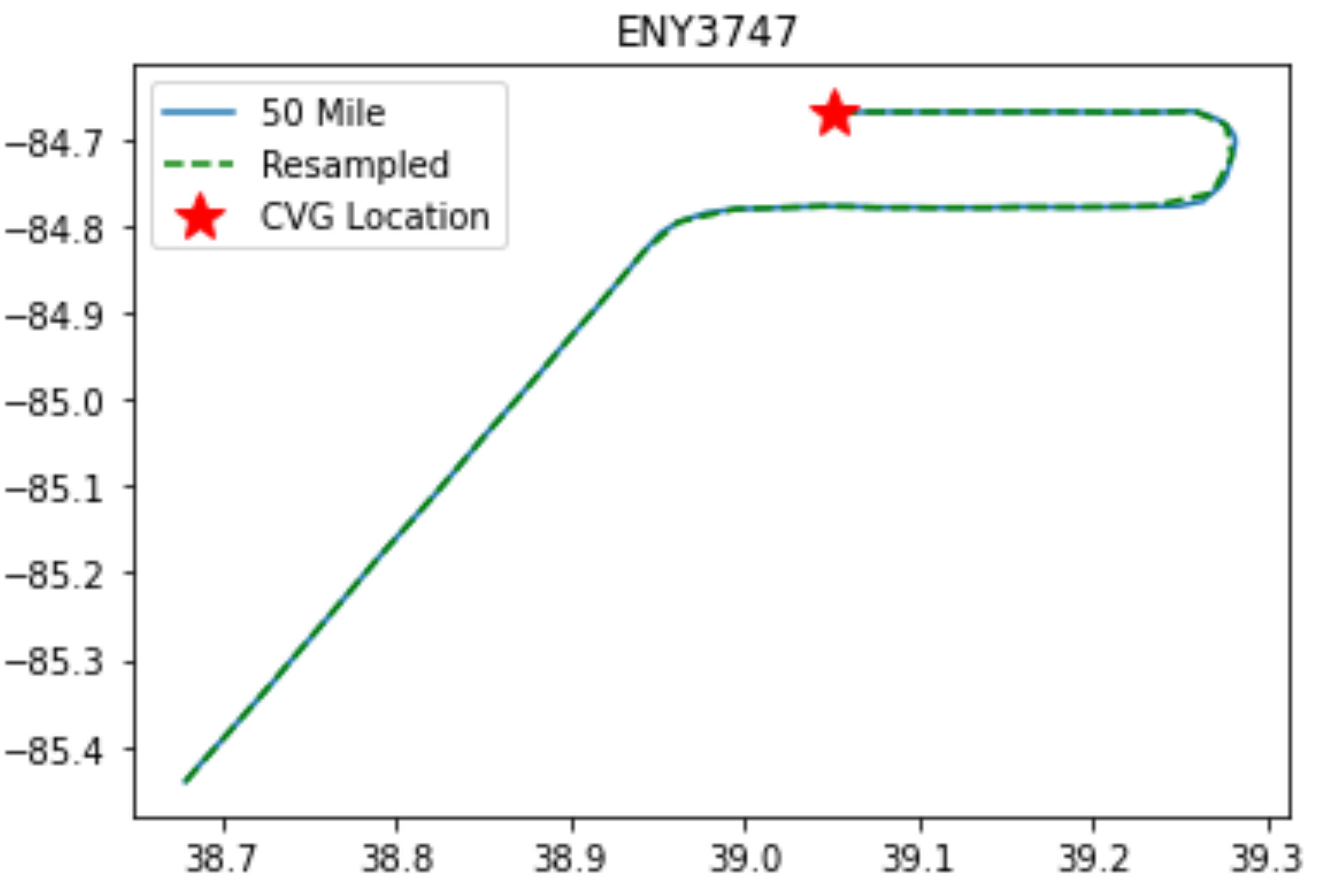


Step 3: Data Cleaning

- Cleaning the data is crucial step
- Data was recorded for planes even while landed
- Developed code to stop recording once an aircraft was below a specific altitude
- Used sectional chart altitude as wide range of altitudes across airport
- Some manual cleaning was performed to remove the edge cases

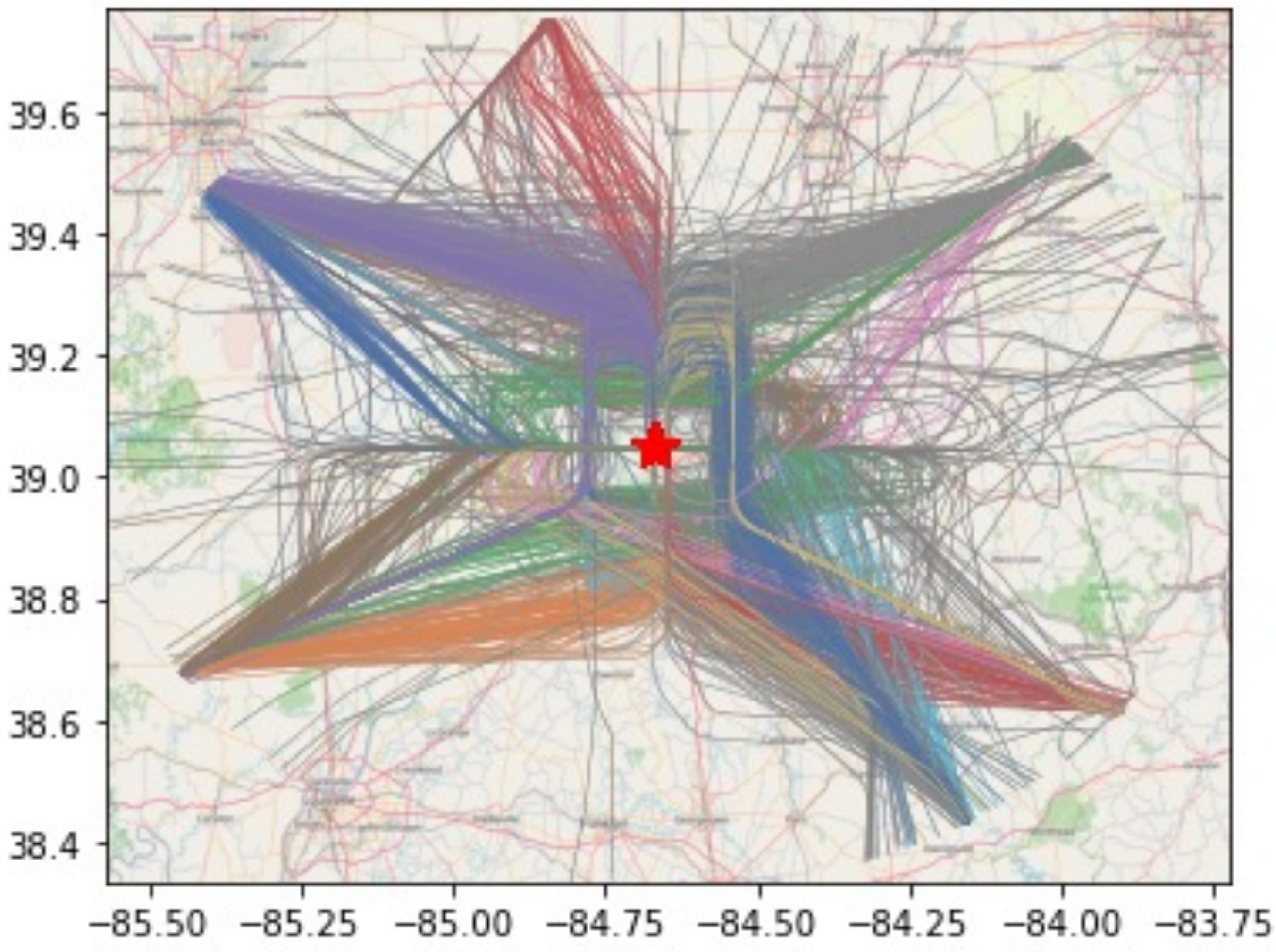
Step 4: Data Resampling

- For clustering, all trajectories needed the same number of points
- Initially tried to use a python library, however data type issues prohibited this
- Developed custom resampling code
- Code resamples based on linearly spaced points between UNIX entry time and final "landing" time
- 50 sample points were used for resampling
- Time points are then compared to flight data and the closest is used for the resampling point



Step 5: Data Clustering

- Clustering was done to gain insight into patterns in flights and to generate possible outputs for future NN
- Clustering was completed based on Latitude and Longitude
- Hierarchical Density Based Scan (HDB-Scan) clustering algorithm was selected
- HDB-Scan was selected over DBScan and other clustering algorithms because its ability to identify clusters of varying density[1]
- 21 Unique Clusters + set of 1 Un-clustered data were generated



[1]Corrado, S. J., Puranik, T. G., Fischer, O. P., & Mavris, D. N. (2021). A clustering-based quantitative analysis of the interdependent relationship between spatial and energy anomalies in ADS-B trajectory data. *Transportation Research Part C: Emerging Technologies*, *131*, 103331. https://doi.org/10.1016/j.trc.2021.103331

Cluster Analysis – Full Data Set

Cluster	Mean (Seconds)	Mean (Minutes)	Standard Dev (Seconds)	Standard Dev (Minutes)
-1	1081.636364	18.02727273	304.6880987	5.078134978
0	1081.157895	18.01929825	72.33785	1.205630833
1	756.5421687	12.60903614	61.63794427	1.027299071
2	983.5952381	16.39325397	130.0623586	2.167705976
3	829.4625	13.824375	66.25234784	1.104205797
4	917.5892857	15.29315476	96.22773894	1.603795649
5	674.3076923	11.23846154	64.91578523	1.081929754
6	995.4705882	16.59117647	75.32917382	1.25548623
7	762.3595506	12.70599251	62.63937219	1.043989536
8	936.5588235	15.60931373	102.5481784	1.709136307
9	929.8703704	15.49783951	92.6213011	1.543688352
10	701.7931034	11.69655172	53.30904	0.888484
11	1027.4	17.12333333	217.1432707	3.619054511
12	1064.033333	17.73388889	107.2431764	1.787386274
13	724.1923077	12.06987179	93.35994003	1.555999
14	748.5488959	12.47581493	78.24515571	1.304085928
15	937.2051282	15.62008547	72.23898354	1.203983059
16	846.75	14.1125	69.91021027	1.165170171
17	866.8791946	14.44798658	153.8325121	2.563875202
18	990.0144509	16.50024085	101.7065217	1.695108695
19	824.0487805	13.73414634	65.19275406	1.086545901
20	948.2979351	15.80496559	90.07143852	1.501190642

Cluster Analysis Statistics

- Minimum Standard Deviation was 0.88 minutes
- Maximum* Standard Deviation was 3.619 minutes
- Average* Standard Deviation was 1.529

*Does Not include the statistics about un-clustered data

Results Discussion

- For clustering to be a stand-alone viable solution, "we want accuracy of the prediction to be within ±5% of flight time until Meter fix or 20 seconds whichever is larger" [1]
- Not the case with clustering alone
  - Results from clustering are statistical in nature and only look at positional information from 2 weeks of flight data – no additional meta data
  - The clusters included go around flights
  - A few flights were clustered with discrepancies in their landings
- These predictions and time estimates will be further improved by the development of the Neural Network

[1] Roland M. Sgorcea and Lesley A. Weitz. "Role of Aircraft Descent Speed Schedule Prediction in Achieving Trajectory Based Operations," AIAA 2021-2374. *AIAA AVIATION 2021 FORUM*. August 2021.

Future Work- Step 6 and Step 7

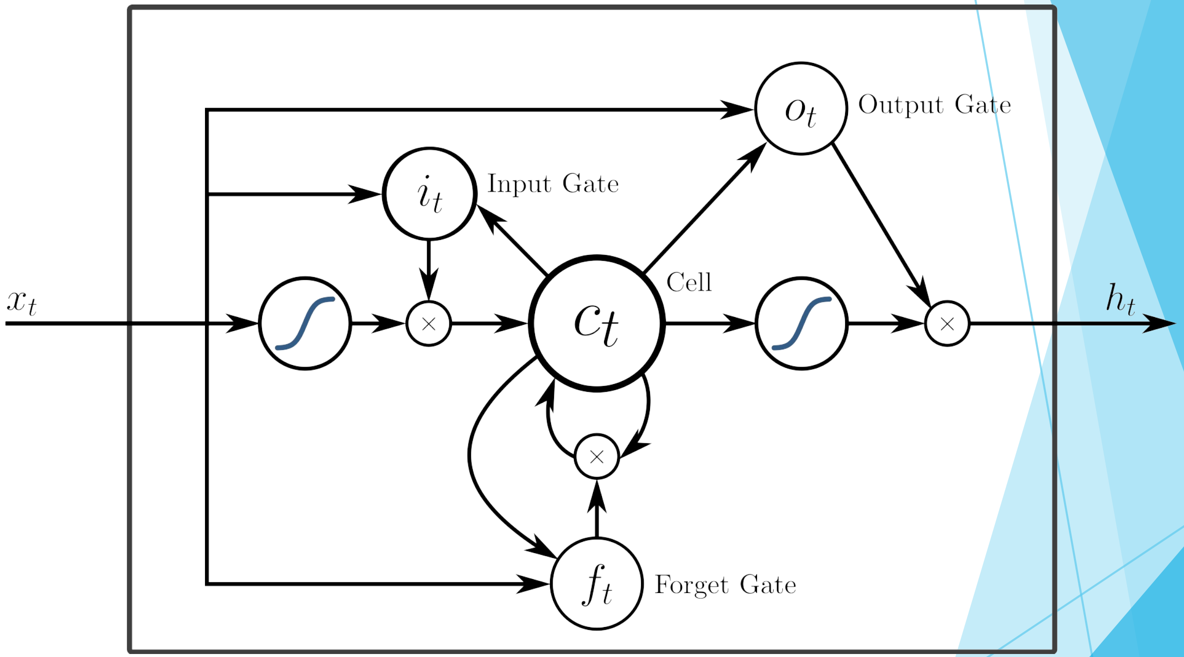
Step 6: Collecting and Processing Weather Data

- Going to Use Iowa State University's METAR Data Download
- Need to determine what weather data and how to fuse it
- Need to decode and clean the data

station	valid	tmpf	dwpf	relh	drct	sknt	p01i	alti	mslp	vsby	gust	skyc1	skyc2	skyc3	skyc4	skyl1	skyl2	skyl3	skyl4	wxcodes	ice_accret	ice_accret	ice_accret	peak_win(	peak_win(	peak_win(	feel	metar	snowdept
CVG	2020-01-01 00:00	M	M	M	250	13	M	29.89	M	10	18	OVC	M	M	M	2100	M	M	M	M	M	M	M	M	M	M	M	KCVG 010(	M
CVG	2020-01-01 00:05	M	M	M	250	14	M	29.89	M	10	M	OVC	M	M	M	2200	M	M	M	M	M	M	M	M	M	M	M	KCVG 010(	M
CVG	2020-01-01 00:10	M	M	M	250	11	M	29.89	M	10	M	OVC	M	M	M	2200	M	M	M	M	M	M	M	M	M	M	M	KCVG 010(	M
CVG	2020-01-01 00:15	M	M	M	260	13	M	29.89	M	10	M	OVC	M	M	M	2200	M	M	M	M	M	M	M	M	M	M	M	KCVG 010(	M
CVG	2020-01-01 00:20	M	M	M	260	13	M	29.89	M	10	M	OVC	M	M	M	2200	M	M	M	M	M	M	M	M	M	M	M	KCVG 010(	M

Step 7: Neural Network Training

- Going to continue this work by developing and training the Neural Network as discussed
- Will use Open Access Sherlock Data Warehouse to continue upon completion of internship
- Conducting a literature survey on Neural Networks
  - Investigating Long-Short Term Memory (LSTM) Neural Network and their applications
  - LSTM is likely going to be used as its ability to capture long term temporal dependencies



https://en.wikipedia.org/wiki/Long_short-term_memory

Step 7: Training Neural Network

Inputs to Neural Network

- Latitude
- Longitude
- Altitude
- Groundspeed
- Heading
- Aircraft type
- Carrier
- Temperature
- Precipitation
- Wind (direction and speed)
- Anticipated Runway

LSTM Neural Network Path and Time Predictor

Estimated Landing Time

Predicted Path (based on modifying clusters)

Conclusions

- This work covered an introduction to the problem, the desired outcome and a brief literature survey
- The planned flow of work was presented, and current status identified
- Retrieving, filtering, cleaning and resampling of the data obtained from Sherlock was discussed
- The application of the Hierarchical Density Based Scan and results were shown
- Future work including gathering weather data and developing a Neural Network were included

References

- Wang, Lianhui, Pengfei Chen, Linying Chen, and Junmin Mou. 2021. "Ship AIS Trajectory Clustering: An HDBSCAN-Based Approach" *Journal of Marine Science and Engineering* 9, no. 6: 566. https://doi.org/10.3390/jmse9060566
- Malzer, C., & Baum, M. (2021). Constraint-Based Hierarchical Cluster Selection in Automotive Radar Data. *Sensors (Basel, Switzerland)*, *21*(10), 3410. https://doi.org/10.3390/s21103410
- Corrado, S. J., Puranik, T. G., Fischer, O. P., & Mavris, D. N. (2021). A clustering-based quantitative analysis of the interdependent relationship between spatial and energy anomalies in ADS-B trajectory data. *Transportation Research Part C: Emerging Technologies*, *131*, 103331. https://doi.org/10.1016/j.trc.2021.103331
- Roland M. Sgorcea and Lesley A. Weitz. "Role of Aircraft Descent Speed Schedule Prediction in Achieving Trajectory Based Operations," AIAA 2021-2374. *AIAA AVIATION 2021 FORUM*. August 2021
- *Samet Ayhan, Pablo Costas, and Hanan Samet. 2018. Predicting Estimated Time of Arrival for Commercial Flights. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '18). Association for Computing Machinery, New York, NY, USA, 33–42. https://doi.org/10.1145/3219819.3219874*
- W. Zeng, Z. Quan, Z. Zhao, C. Xie and X. Lu, "A Deep Learning Approach for Aircraft Trajectory Prediction in Terminal Airspace," in *IEEE Access*, vol. 8, pp. 151250-151266, 2020, doi: 10.1109/ACCESS.2020.3016289.

Questions?