GEORGE MASON UNIVERSITY

Application of AI to Collision Risk Safety Analysis for the National Airspace System

John Shortle, Shahab Aref, Lance Sherry

September 17, 2024

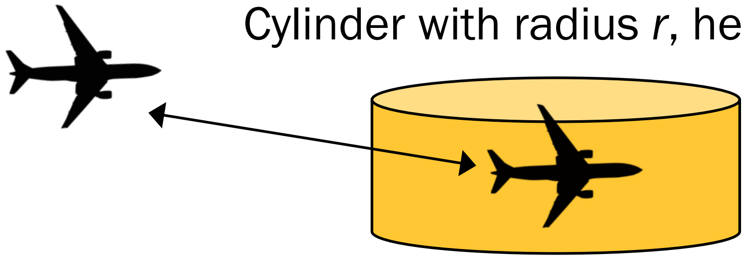
Acknowledgments

- This work is supported by the Federal Aviation Administration, Somil Shah.
- This talk solely represents the opinions of the authors.

Collision Risk Modeling

- Collision Risk Models (CRM) can be used to demonstrate that proposed procedures meet airspace safety standards.
- Is collision risk less than 1E-9?

Proximity volume:
Cylinder with radius *r*, height *h*



Flight Track Data for Collision Risk Models

- Historic flight tracks
  - Limited number of flight tracks to evaluate 1E-9 event
  - May not be relevant for evaluating new procedures, new technology, different airports, etc.
- Simulated flight tracks
  - Computer model may not capture nuances of real operations
  - Computationally prohibitive to generate enough tracks to obtain reasonable confidence interval for 1E-9 event
- Synthetic tracks
  - Potential to generate realistic new (never been seen) tracks, based on historic data (Krauth et al. 2023)

Krauth, Timothé, Adrien Lafage, Jérôme Morio, Xavier Olive, and Manuel Waltert. 2023. Deep generative modelling of aircraft trajectories in terminal maneuvering areas. *Machine Learning with Applications*, 11, 10046.

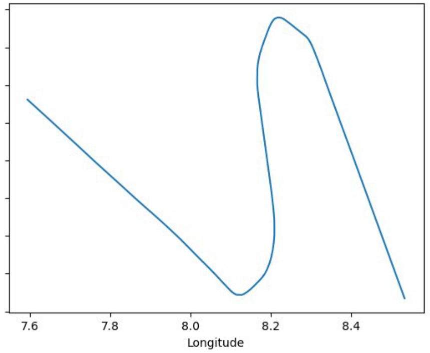
Generation of Synthetic Flight Tracks

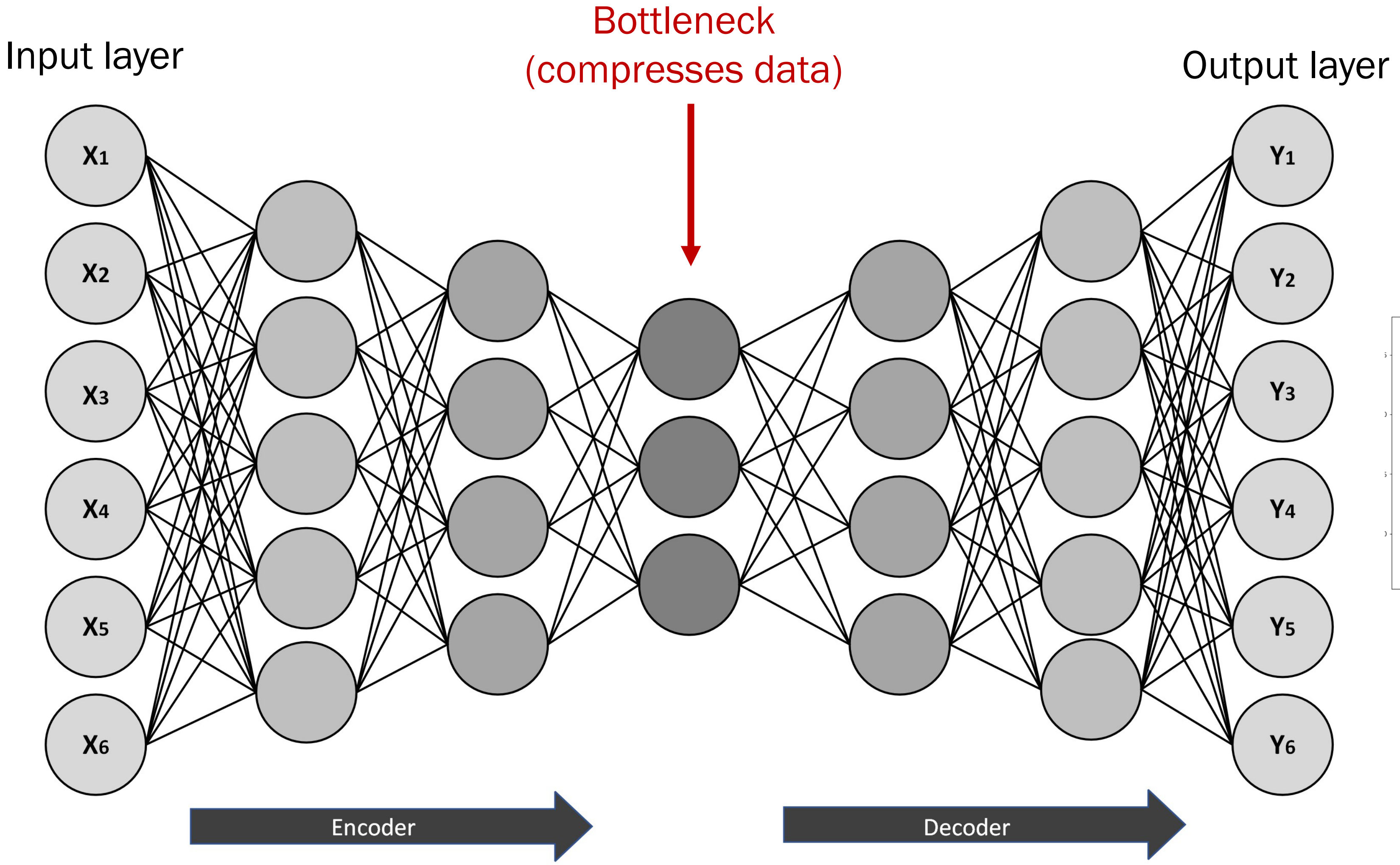
- **Objective:** Can AI/ML methods be used to generate new "realistic" synthetic flight trajectories for use in collision risk models?
- AI/ML methods: Variational autoencoder, Gaussian mixture models, generative adversarial network, …
- How to measure "realism"
  - Tracks conform to laws of physics and aircraft performance limits
  - Adherence to flight procedures
  - Statistical comparison to real flight trajectories

[1] T. Krauth, A. Lafage, J. Morio, X. Olive, and M. Waltert, "Deep generative modelling of aircraft trajectories in terminal maneuvering areas," Machine Learning with Applications, vol. 11, p. 100446, 2023.
[2] S. Jung and M. J. Kochenderfer, "Learning terminal airspace traffic models from flight tracks and procedures," in 2019 IEEE/AIAA 38th Digital Avionics Systems Conference (DASC). IEEE, 2019, pp. 1–8.
[3] P. Lukes and P. Kulmon, "Generating realistic aircraft trajectories using generative adversarial networks," in 2023 24th International Radar Symposium (IRS). IEEE, 2023, pp. 1–10.

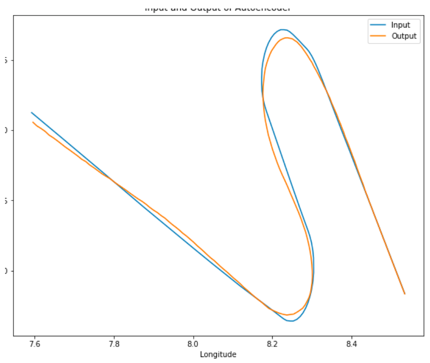
Autoencoder

Input trajectory



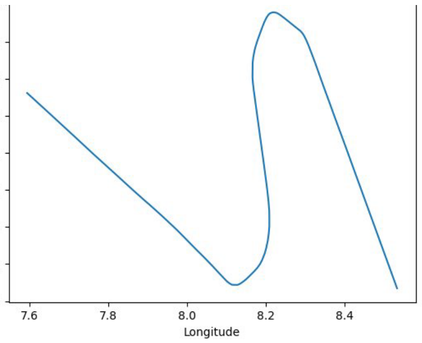


Output trajectory



Variational Autoencoder

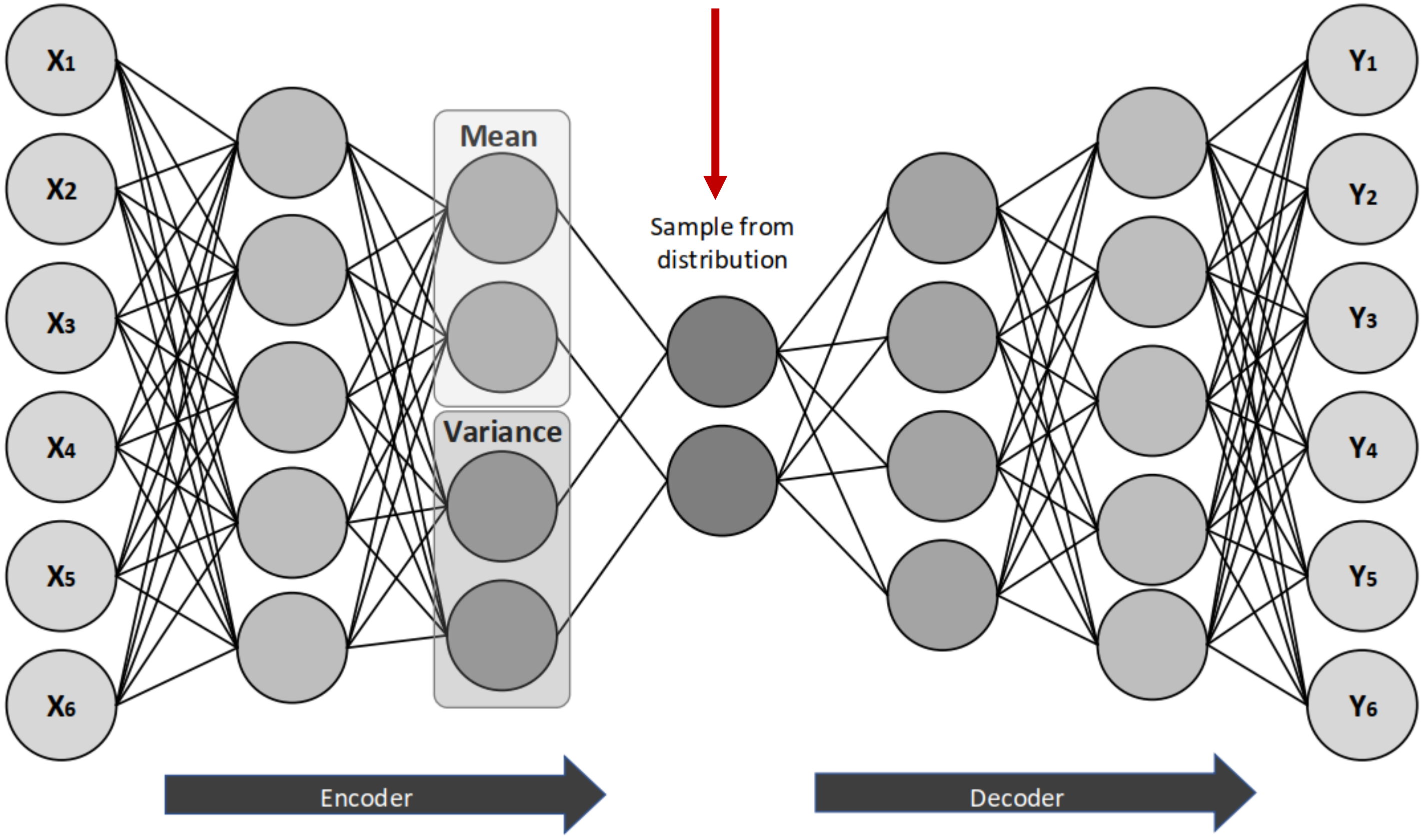
Input = seed trajectory



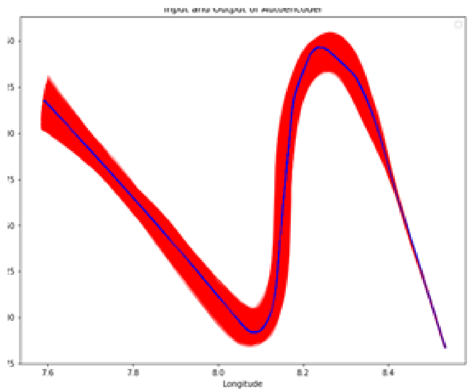
Input layer

Bottleneck compresses data and stores mean and variance

Output layer



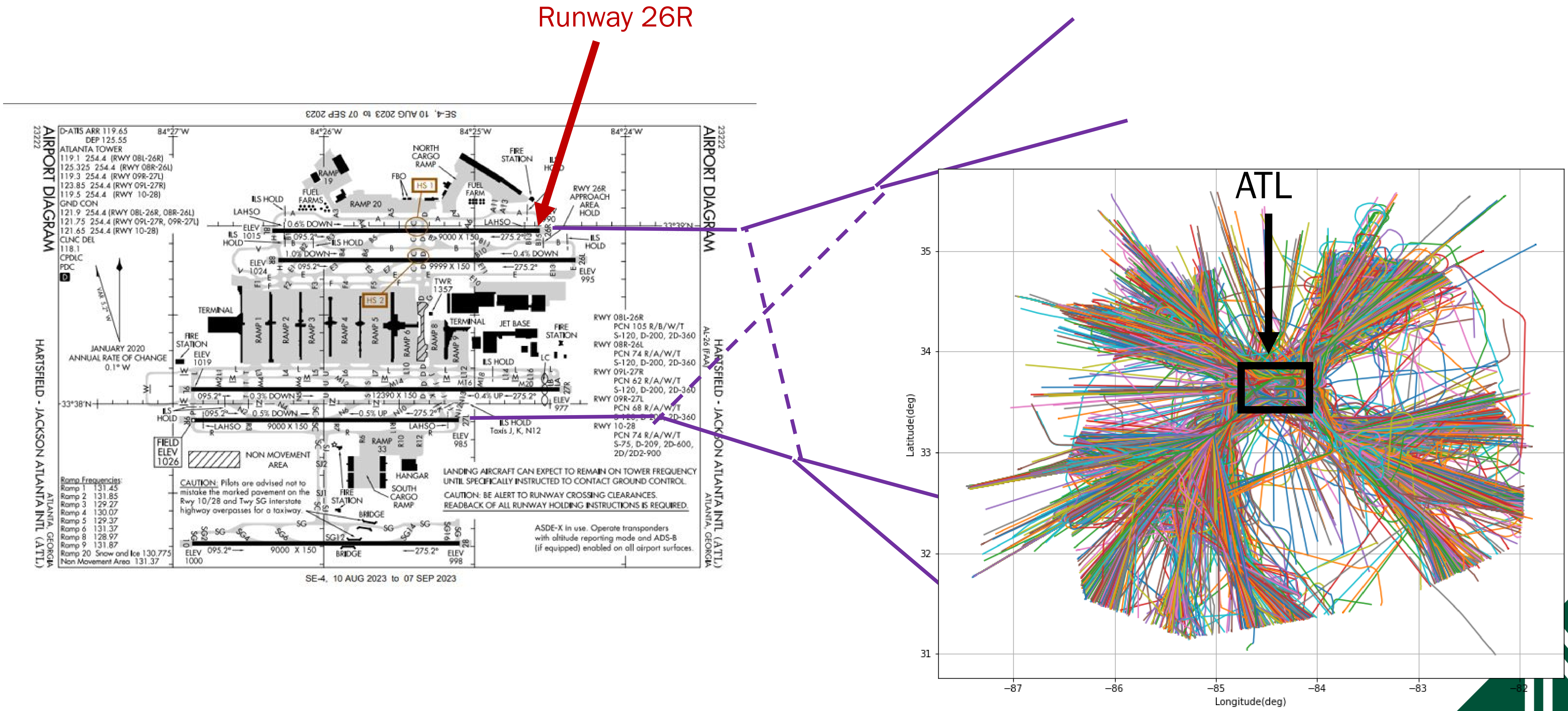
Output = synthetic trajectory



Can generate **<u>multiple synthetic tracks</u>** from single seed track

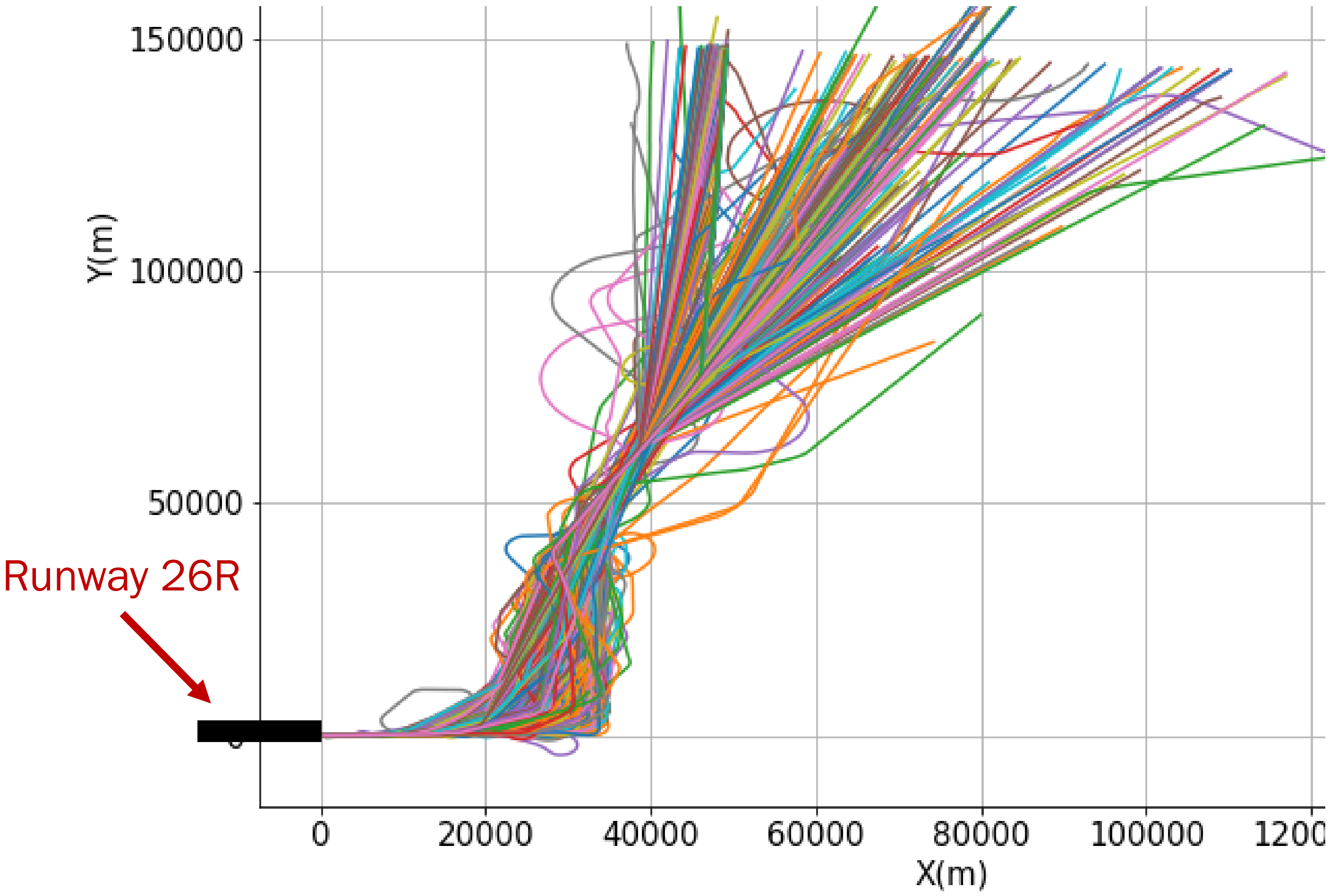
Kingma, Diederik P., and Max Welling. "An introduction to variational autoencoders." Foundations and Trends® in Machine Learning 12.4 (2019): 307-392.

Atlanta Airport



Data Set: Arrivals to ATL, Runway 26R

- **2,356 flight tracks**
- January to April, 2020
- Track points include longitude, latitude, altitude, and timestamp.
- Flight tracks have varying numbers of data points, ranging from 167 to 1,050 points.
- Converted to Universal Transverse Mercator (UTM) projection



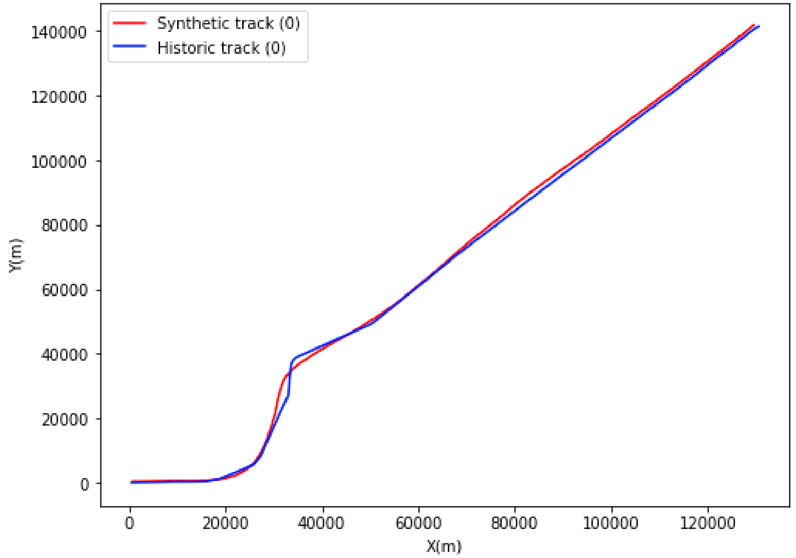
Methodology to Train VAE

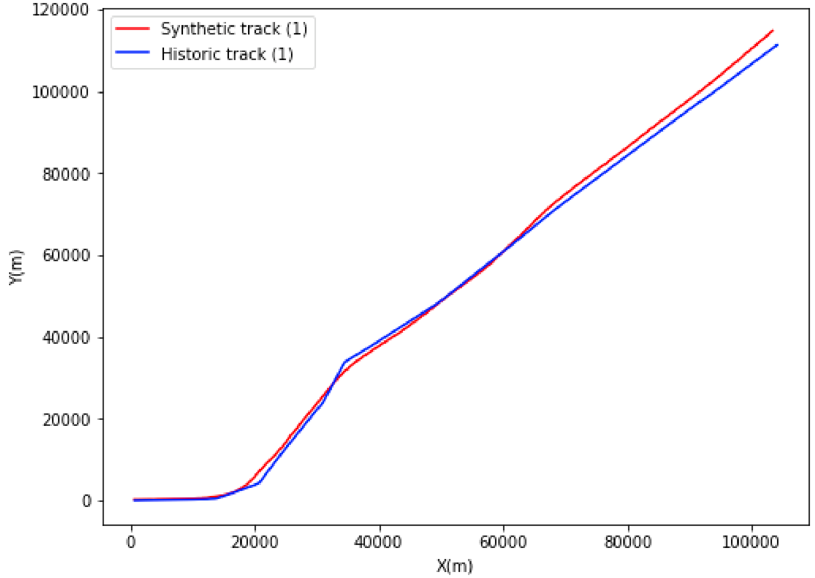
1. Data preparation
   A. Interpolate to achieve common number of points for each track (400)
   B. Calculate velocity (ground speed) at each track point
   C. Normalize data
2. Train variational autoencoder
   - Input/output layers are 1,600 x 1 vector
   - Encoder: 1600, 256, 128, 64, bottleneck (30)
   - Decoder: bottleneck (30), 64, 128, 256, 1600
3. Generate synthetic flight tracks

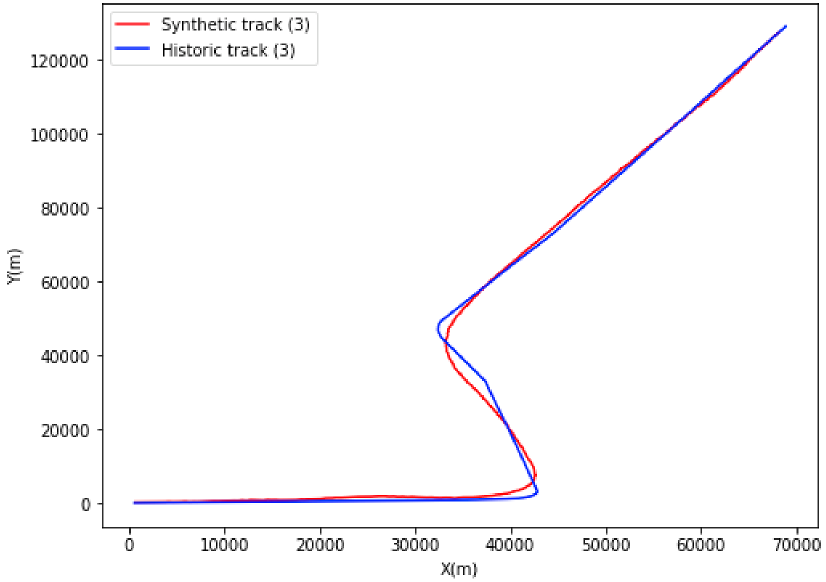
VAE Training Details

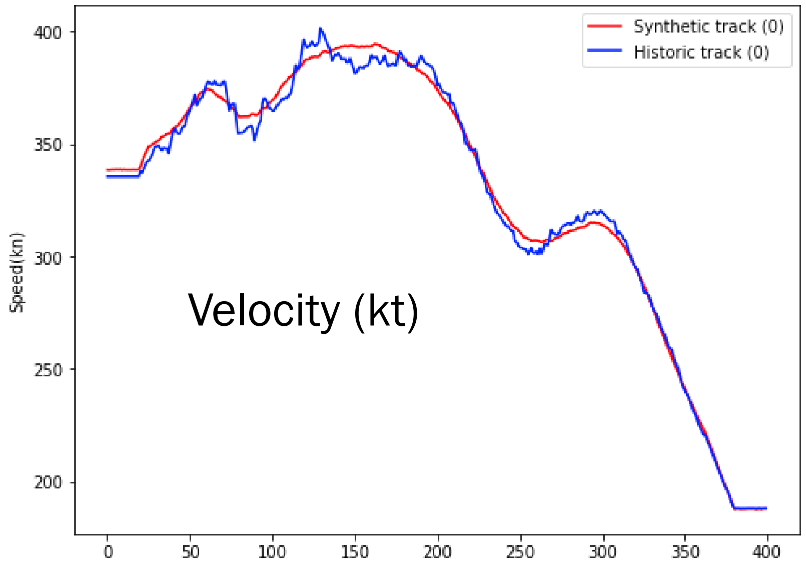
- Four linear layers
  - Encoder: 1600, 256, 128, 64, bottleneck
  - Decoder: bottleneck, 64, 128, 256, 1600
- Bottleneck size: 30
- ReLU activation function
- Learning rate: 0.001
- Batch size: 16
- Epochs: 1,500
- Scheduler is used to reduce learning rate if the loss value does not decrease for 40 consecutive iterations
- Loss = reconstruction loss + KL coef. × KL divergence loss
- Training time is around 10 min
- System: CPU: AMD EPYC 7543 32-Core, GPU: NVIDIA A100-SXM4-10GB

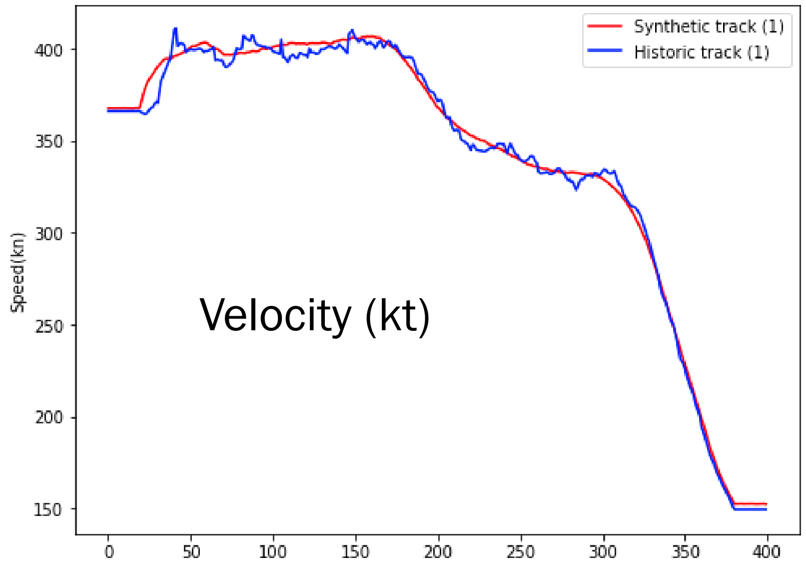
VAE Can Reproduce Historic Tracks

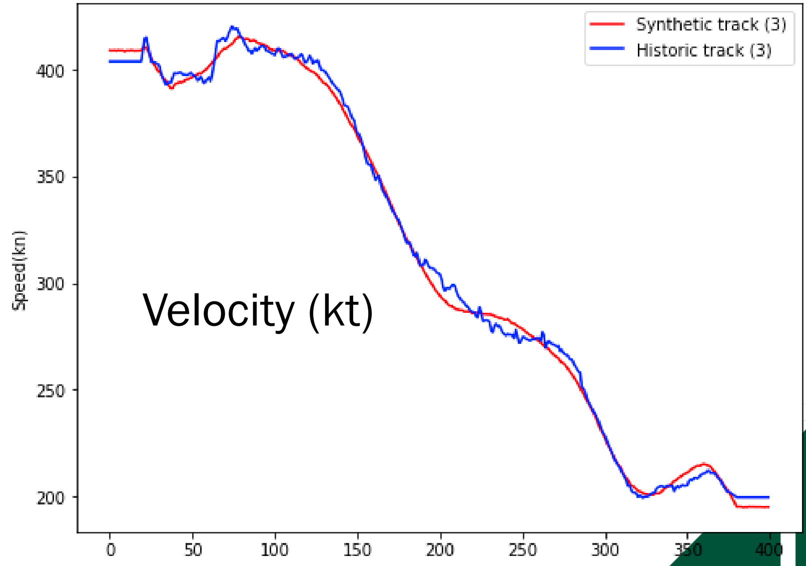




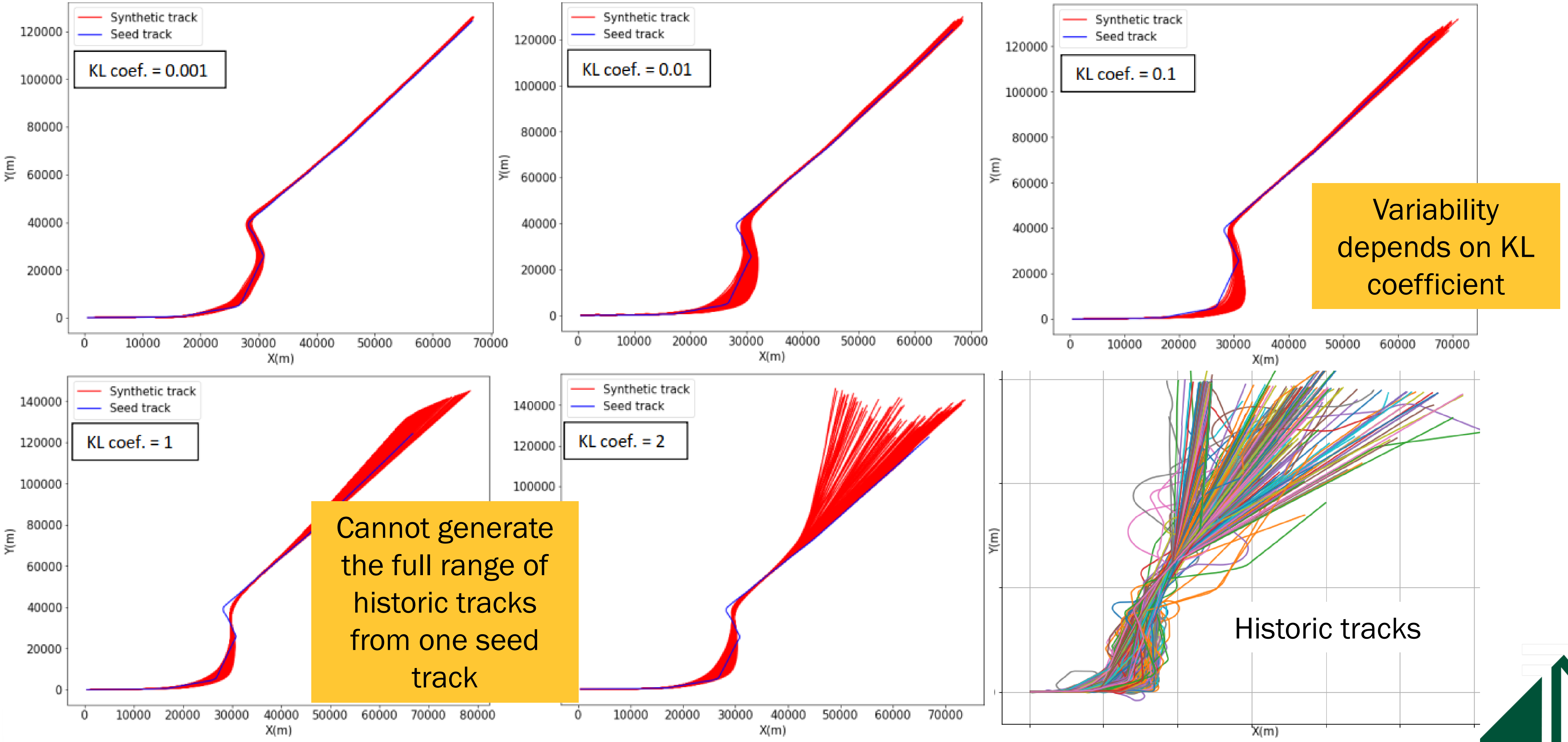








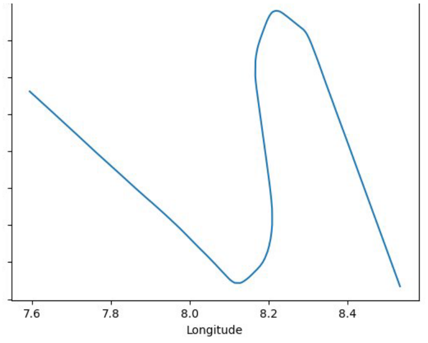
Generating Multiple Tracks from One Seed Track

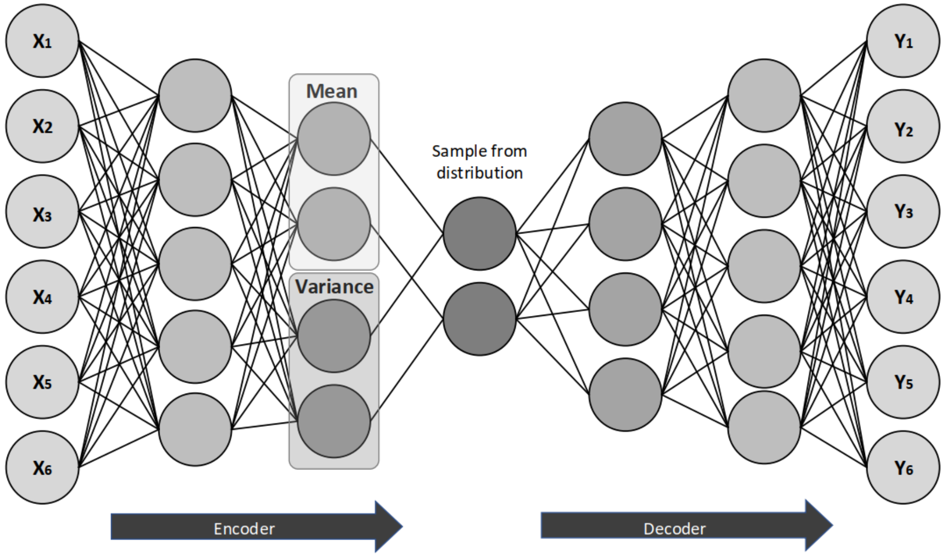


Question

- Cannot generate full range of historic tracks from one seed track
- How many seed tracks are needed to reproduce the statistical distributions of historic tracks?
  - Provides a measure how many “new” tracks can be realistically generated from existing data set

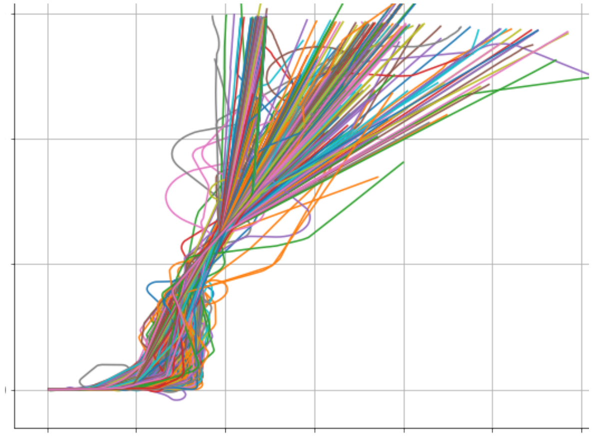
N seed tracks





?

2,356 synthetic tracks

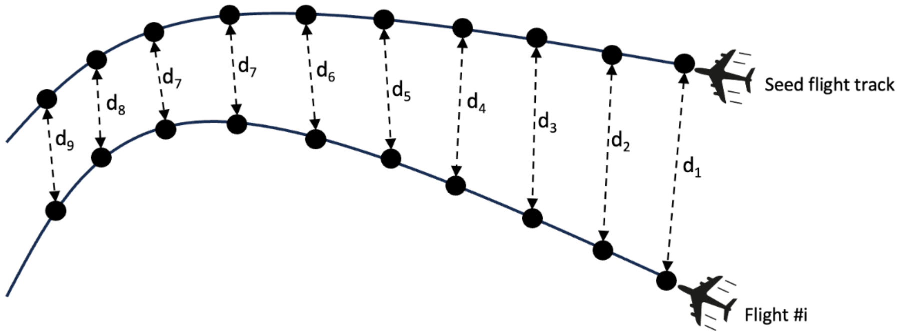


Metrics to Evaluating Synthetic Tracks

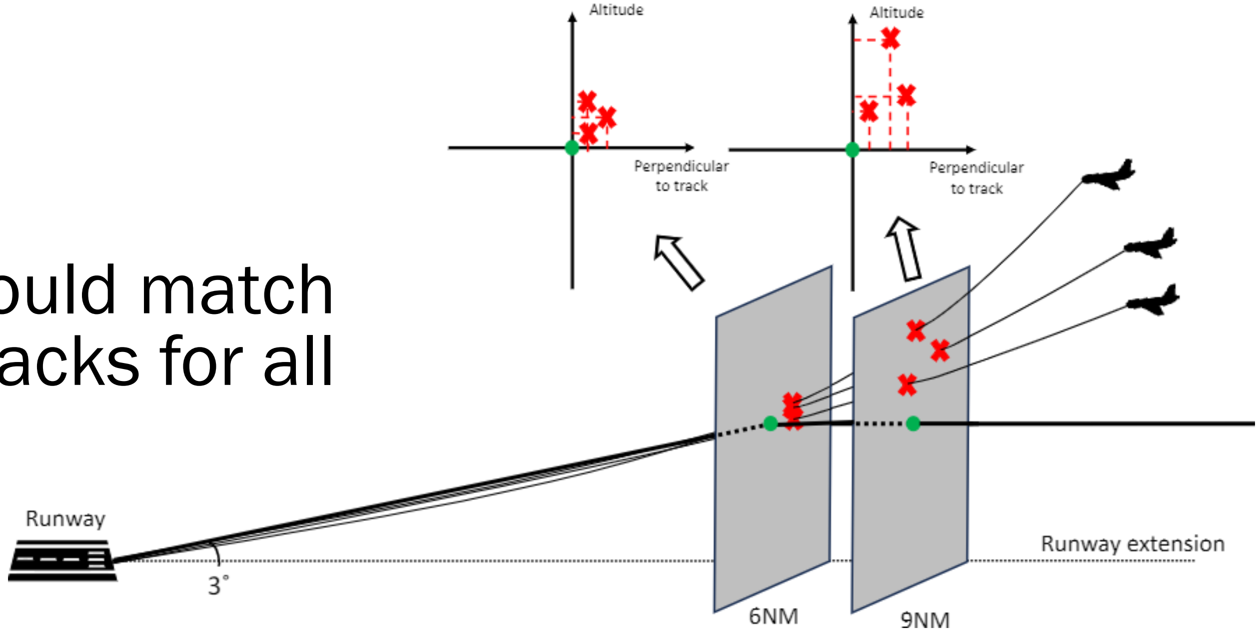
- Along-track distance (ATD)
- Average cross-track distance (ACTD)
- 6nm and 9nm from the runway threshold:
  - Lateral dispersion (LD)
  - Vertical dispersion (VD)
  - Velocity dispersion (SD)

- Goal: Synthetic tracks should match (in distribution) historic tracks for all metrics

Average cross-track distance



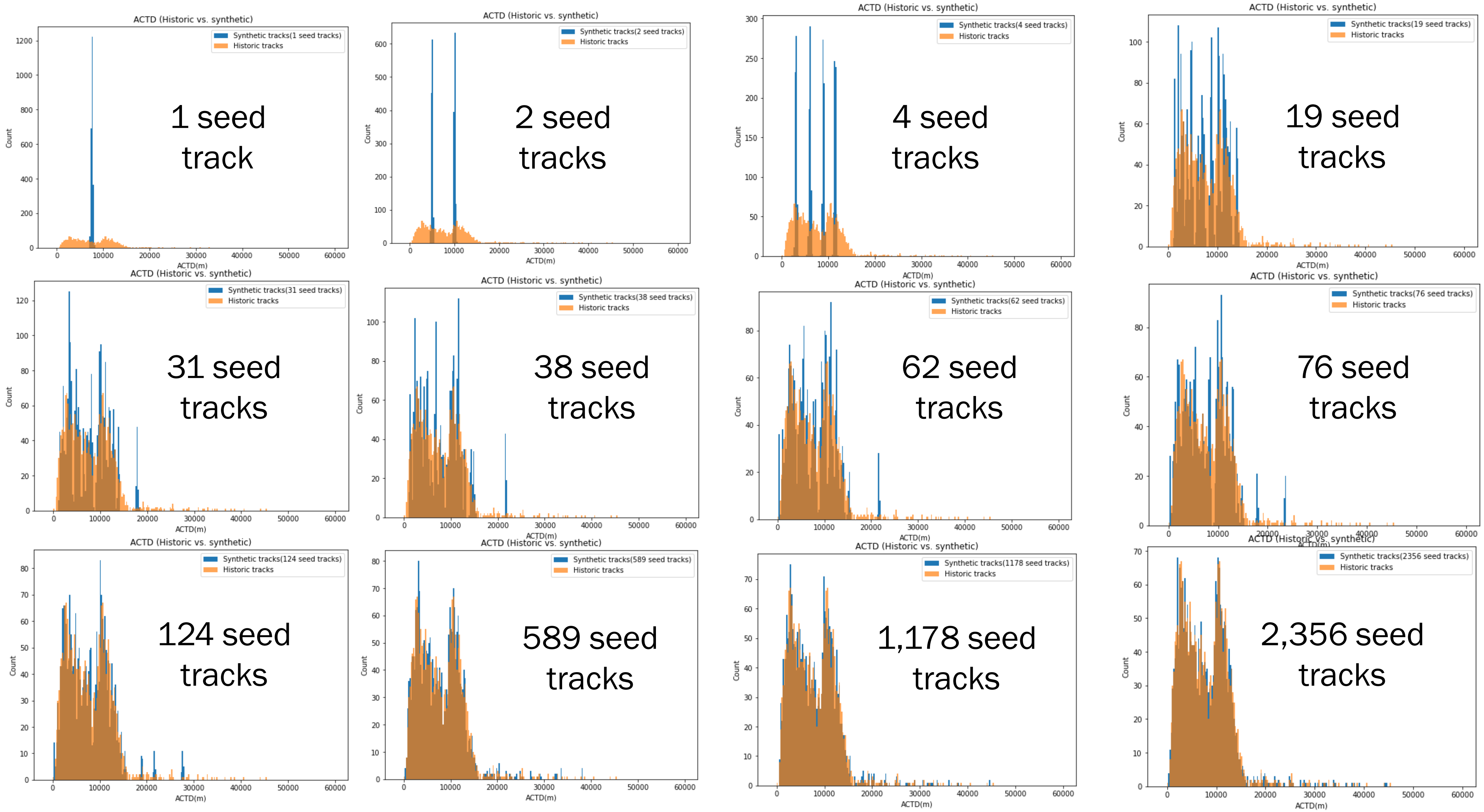
Dispersion metrics at 6nm and 9nm



Methodology

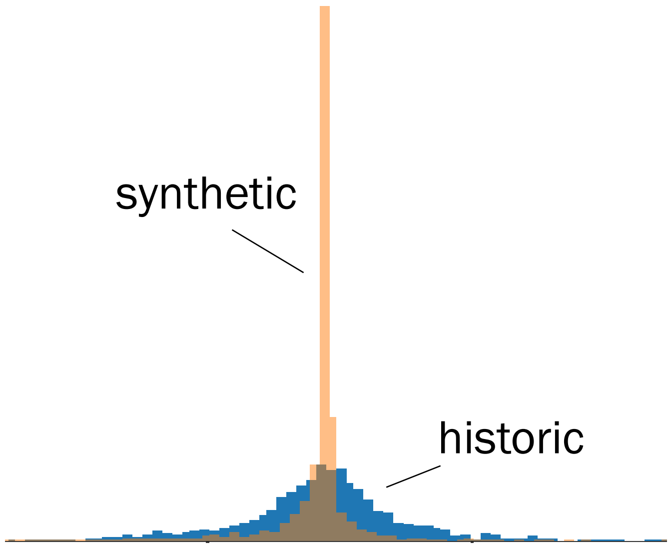
- Train variational autoencoder with a given KL coefficient
- Select *n* seed tracks (e.g., *n* = 1, 2, 4, 19, …, 1178, 2356)
  - Sort historic flight tracks according to track length
  - Choose seed tracks evenly distributed among sorted historic tracks
- From each seed track, generate (2,356 / *n*) synthetic tracks
  - I.e., 2,356 total synthetic tracks
  - Example: 19 seed tracks, each track used to generate 124 synthetic tracks (2,356 synthetic tracks)
- Calculate metrics (ATD, ACTD, LD, VD, and SD) for synthetic tracks
- Compare metric distributions, synthetic versus historical

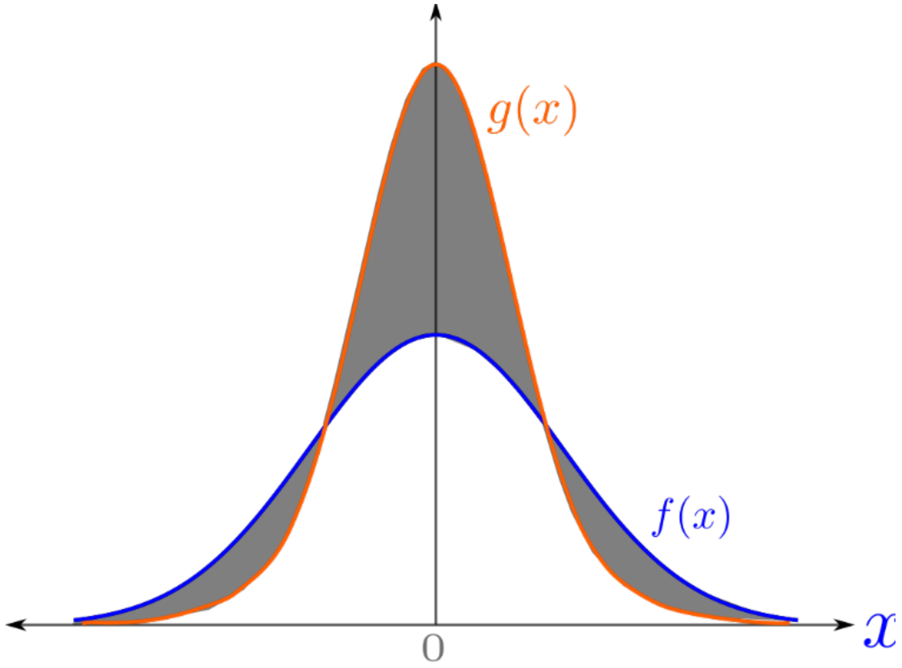
Track Length: Synthetic versus Historic



How Do Distributions Compare Between Synthetic and Historic Tracks

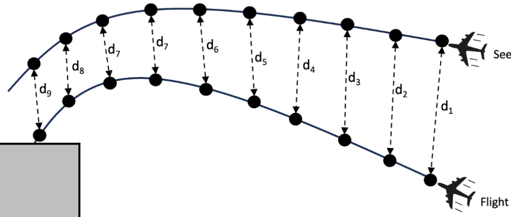
- Total Variation Distance (TVD): Absolute area between the two distributions (historic and synthetic)





Example: Average Cross-Track Distance

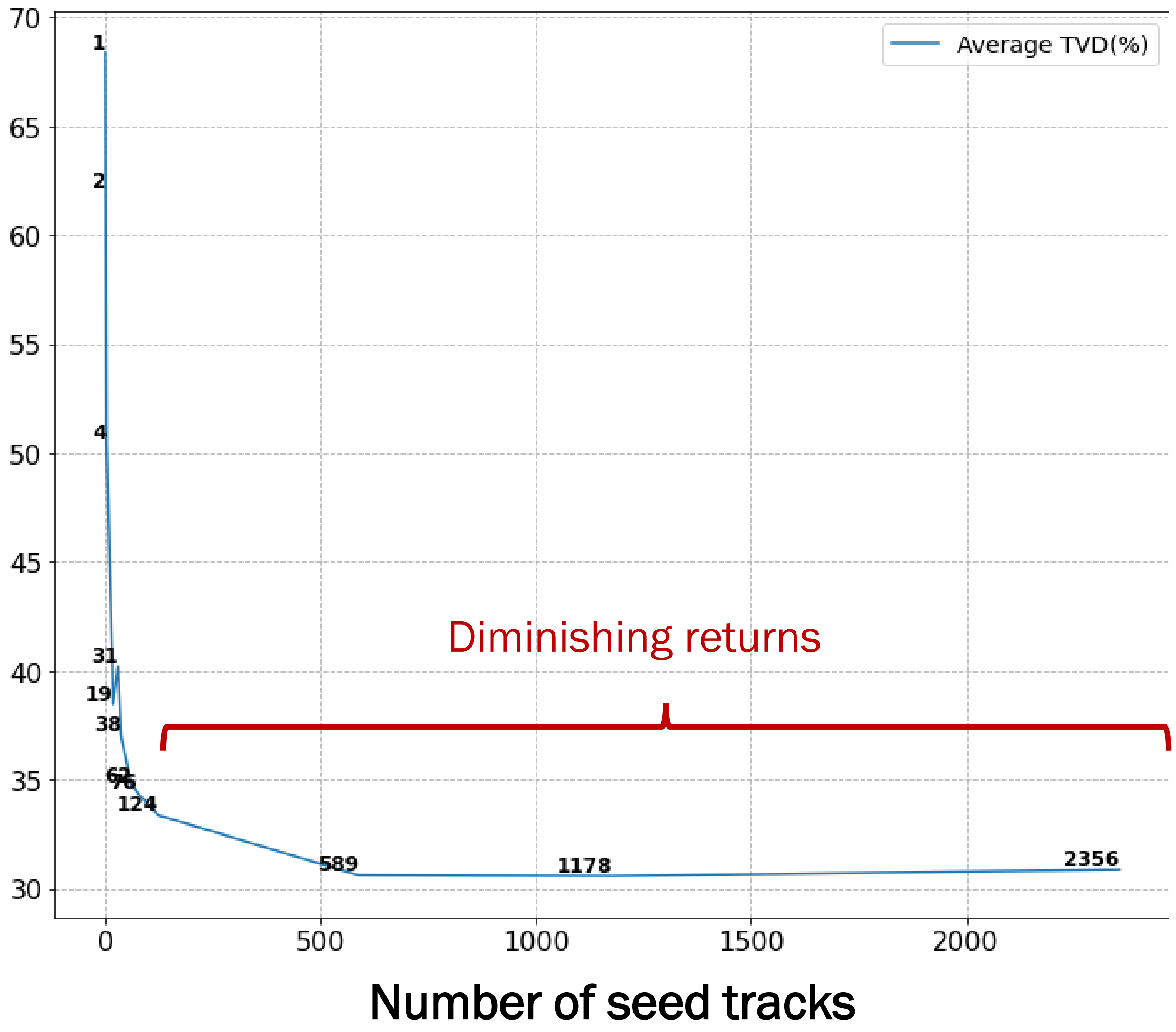
- Total variation distance (TVD in percent), historic vs. synthetic



KL coef. \ #ST	1	2	4	19	31	38	62	76	124	589	1178	2356
0.001	93.25	84.72	73.51	34.84	28.6	23.08	22.24	22.41	14.21	10.48	9.76	8.4
0.01	86.12	77.67	76.4	33.06	26.91	29.58	22.87	18.63	21.56	11.54	9.72	10.86
0.1	62.14	58.4	55.6	20.88	19.31	17.19	18.8	20.54	18.04	10.23	10.86	10.69
1	49.4	40.87	37.01	25.72	25.12	23.26	24.57	23.89	24.57	23.93	23.0	23.81
2	54.37	50.5	36.41	42.31	43.12	39.05	41.25	39.77	40.62	42.31	41.97	41.55

Aggregate Results Combining All Metrics

Average total variational distance, <u>for all metrics</u>, between synthetic and historic tracks



Summary and Next Steps

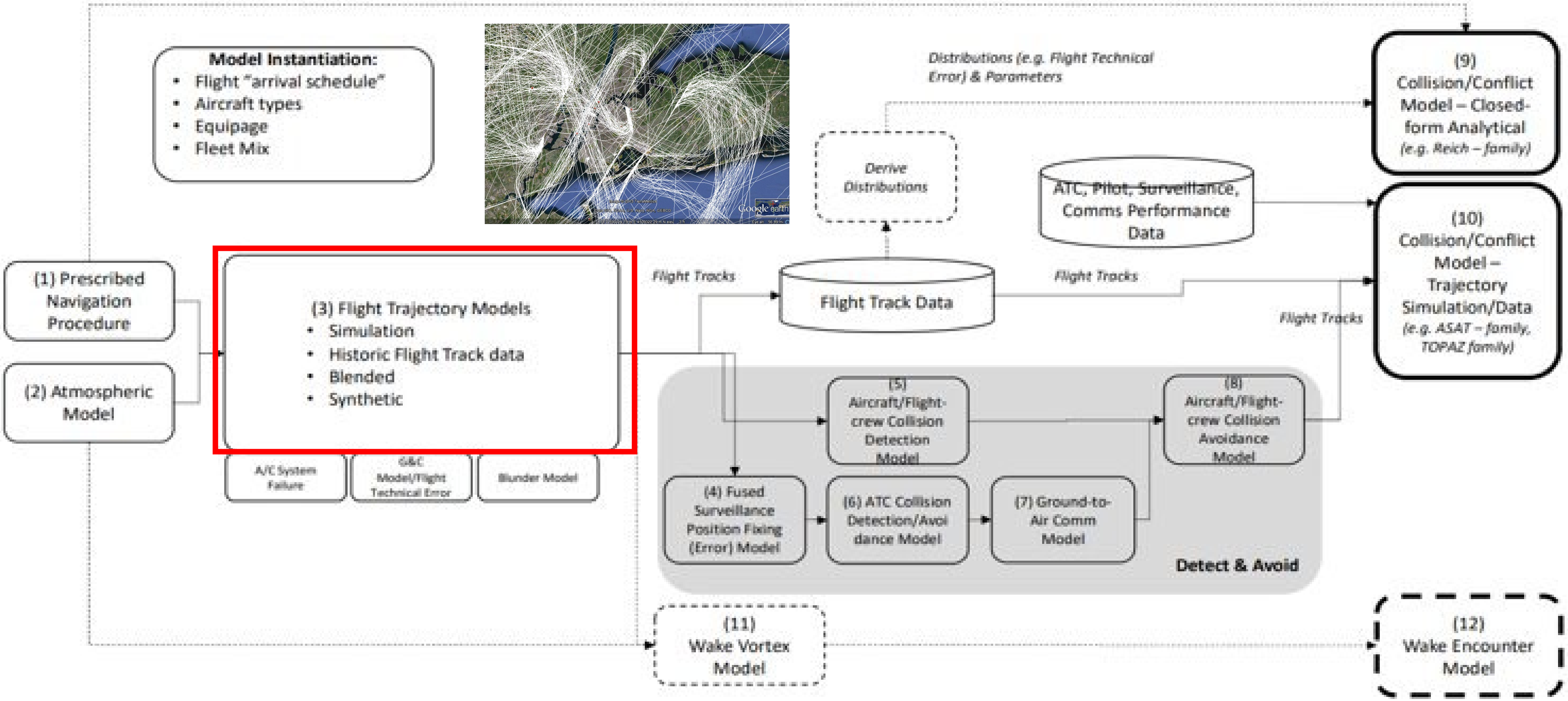
- Trained VAE on flight track data from ATL (2,536 tracks)
- Evaluated synthetic tracks on a range of metrics, comparing synthetic and historic tracks, varied number of seed tracks and KL coefficient
- Synthetic flight tracks matched distributions of historic flight tracks using 1/20 of tracks as seed tracks

Future work

- Use alternate data sets (different arrival patterns, airports; off-nominal events), do results still hold?
- Evaluate risk-based metrics associated with synthetic tracks (requires generating pairs of synthetic tracks)
- Evaluate physics-based metrics
- Use case: Train VAE on existing procedure (many historic flights), tune VAE using data from a new procedure (small number of tracks from flight simulator)

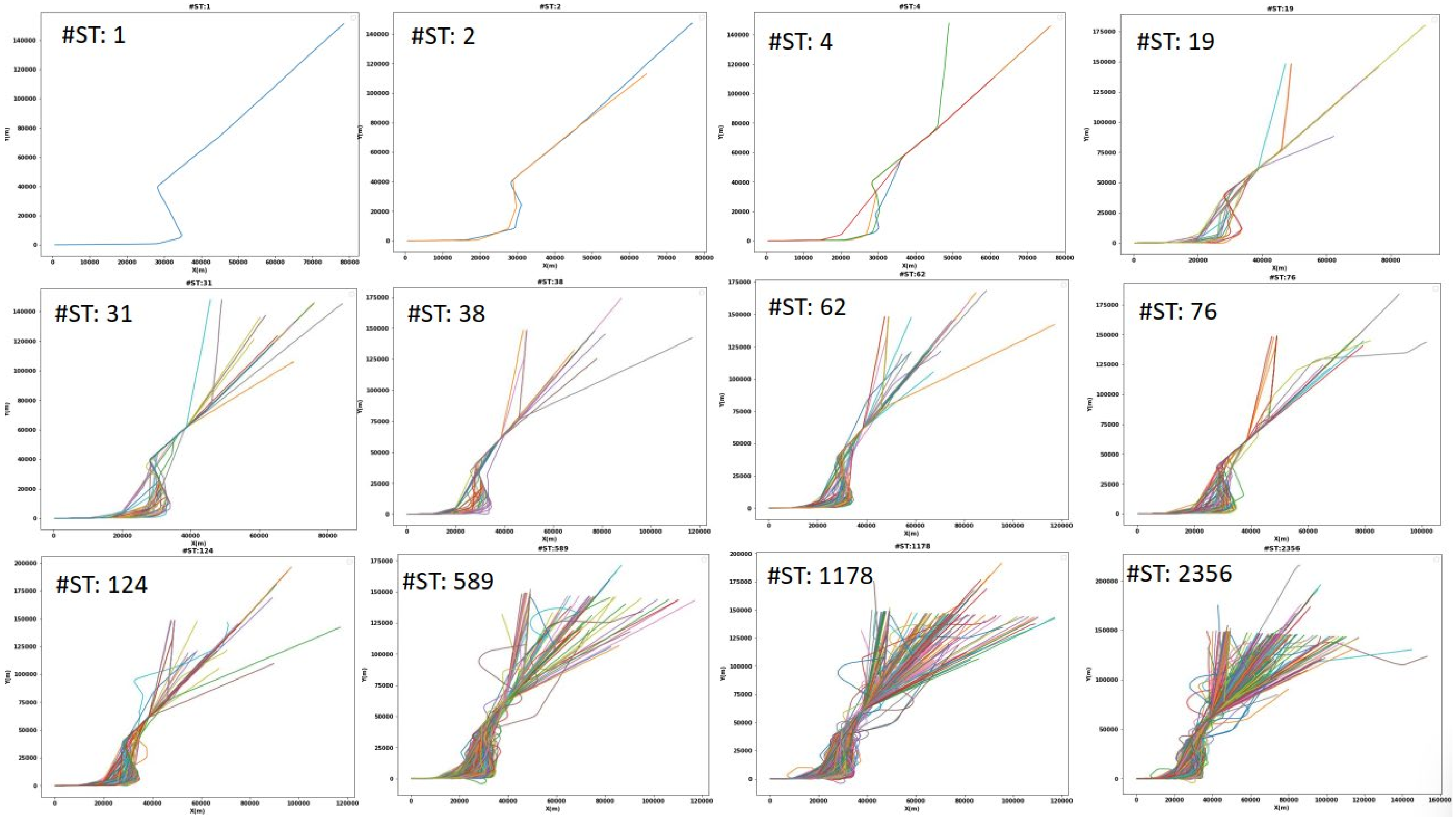
Questions?

Functions for Collision Risk Modeling



[1]Thapa, Ashim Kumar, John Shortle, and Lance Sherry. "Air-to-Air Collision Risk Models (CRM) in the Terminal Airspace." 2023 Integrated Communication, Navigation and Surveillance Conference (ICNS). IEEE, 2023.

Multi-seed Track Generation



Synthetic Track Data