Google Research

Machine-learned contrail forecasting

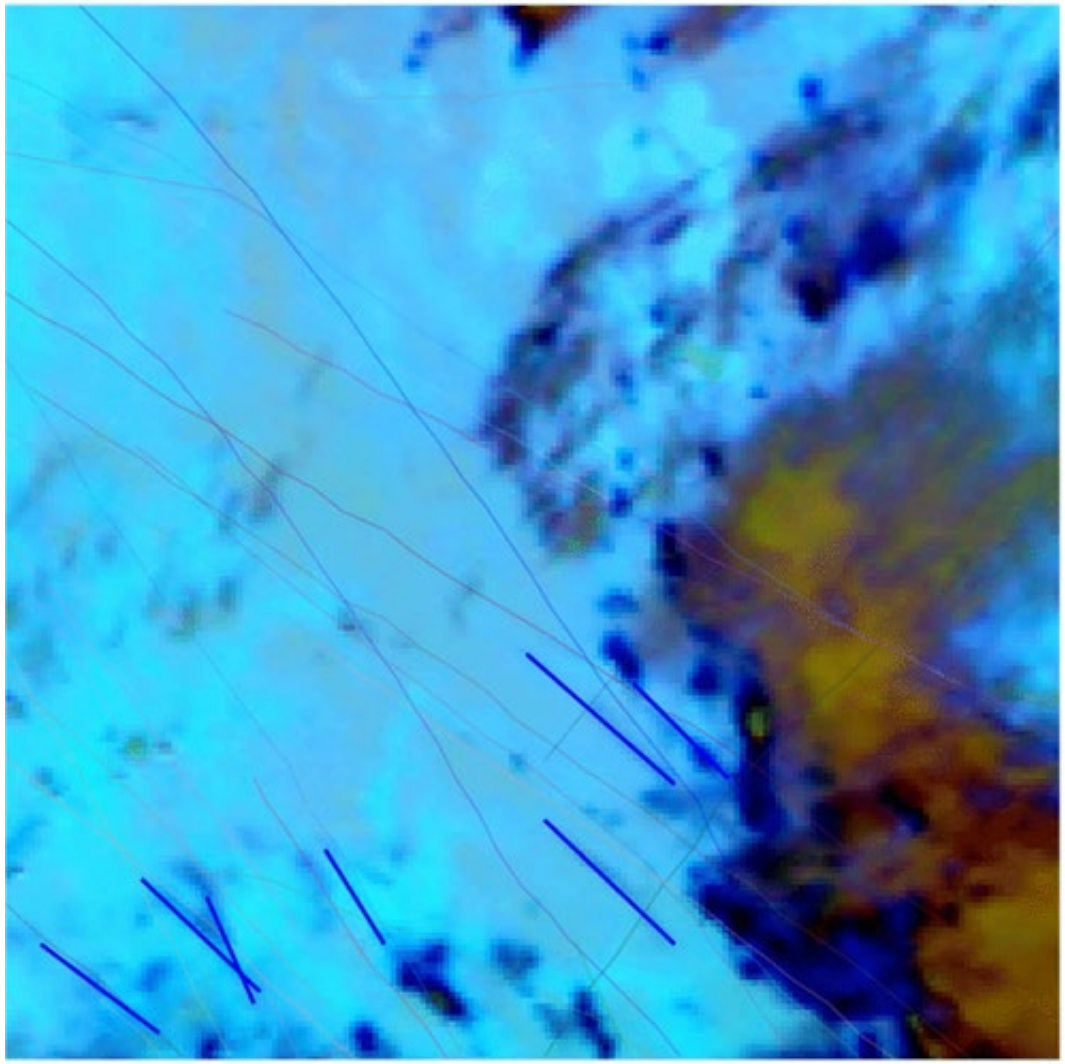
mccloskey@google.com, on behalf of our team

Google Research

1. Matching of flights to detected contrails
2. Forecasting

Google Research

Matching - attributing a contrail to the flight that made it



Yellow - Flight path

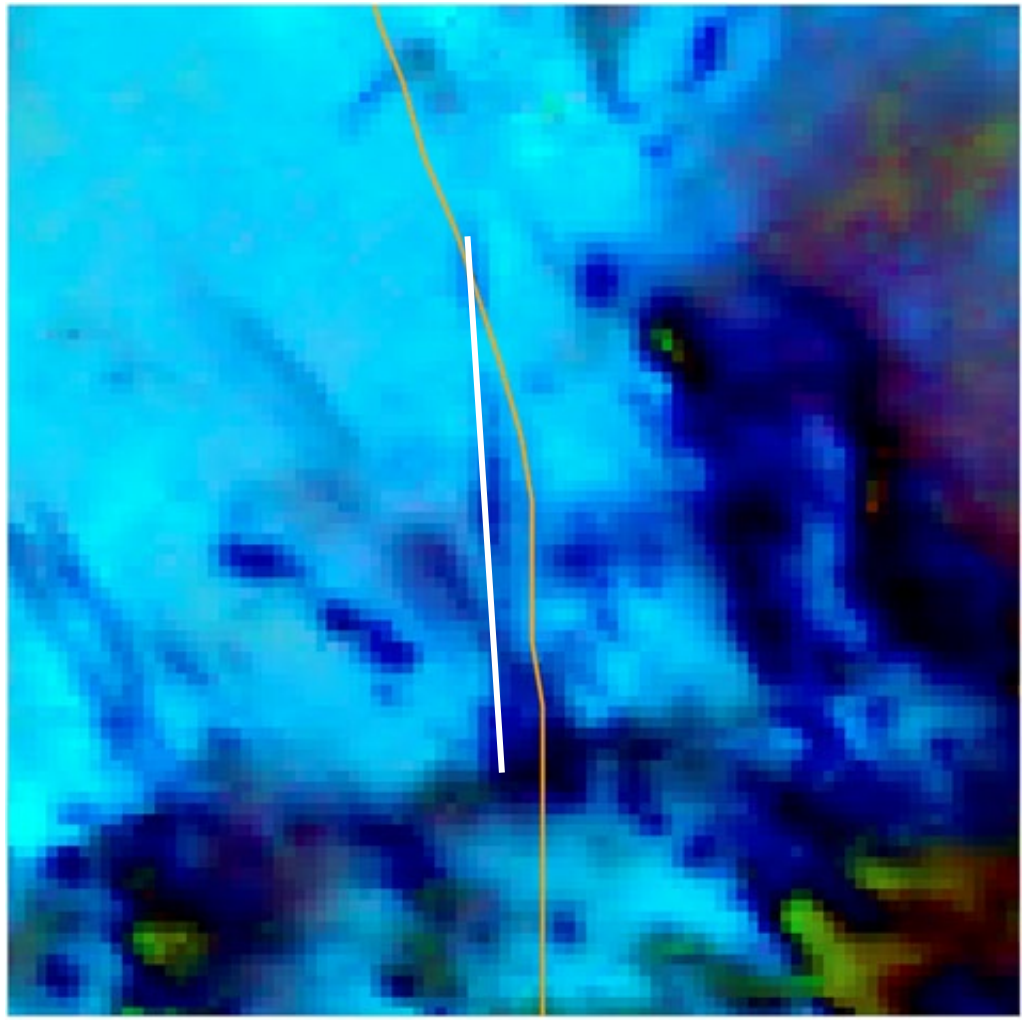
Orange - Advected Flight Path

Dark - Other Advected Flights

Blue lines - Detected contrails

Google Research

Automated matching



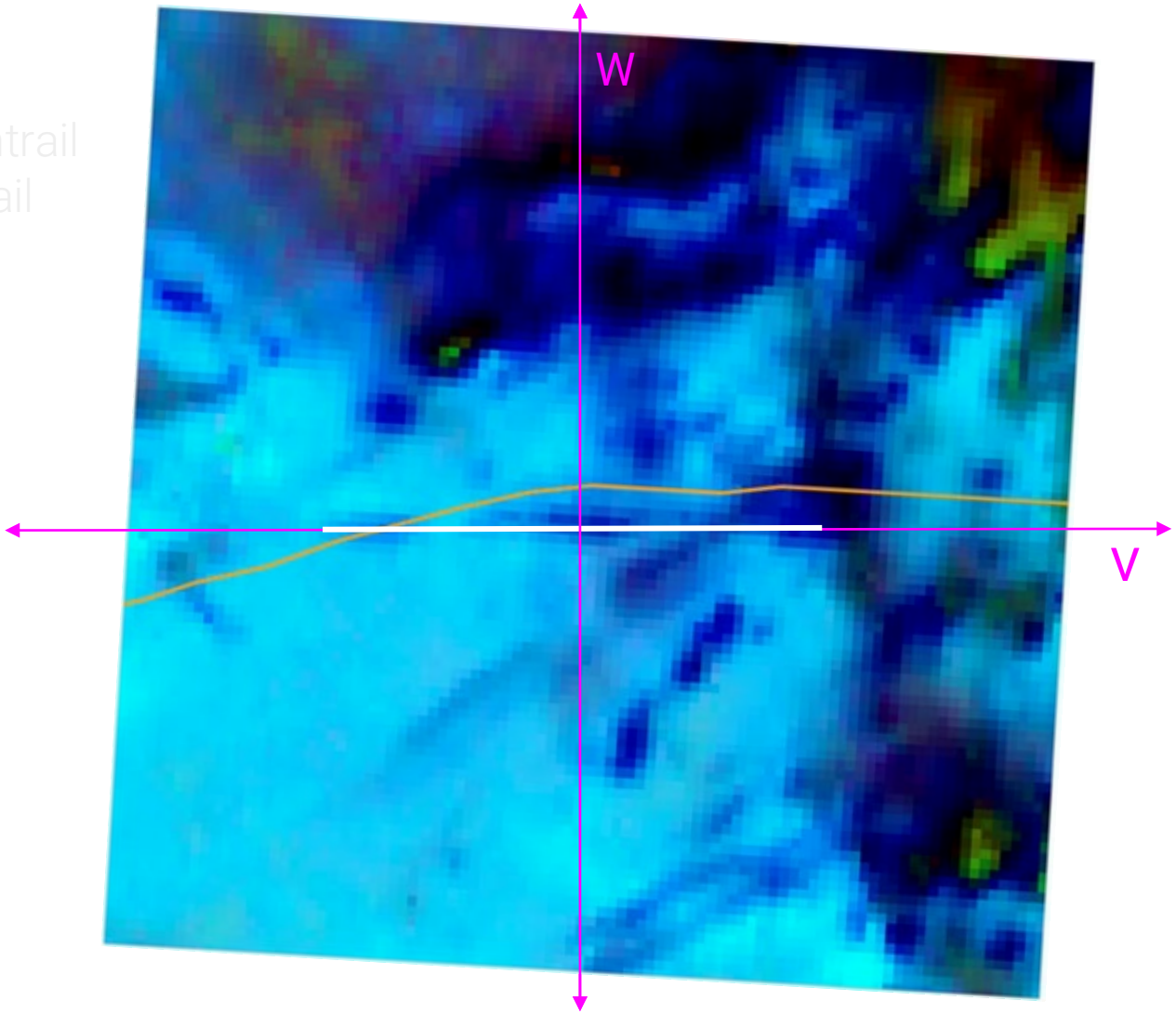
Orange - Advected Flight Path

White line - Detected contrail

Google Research

Automated matching steps

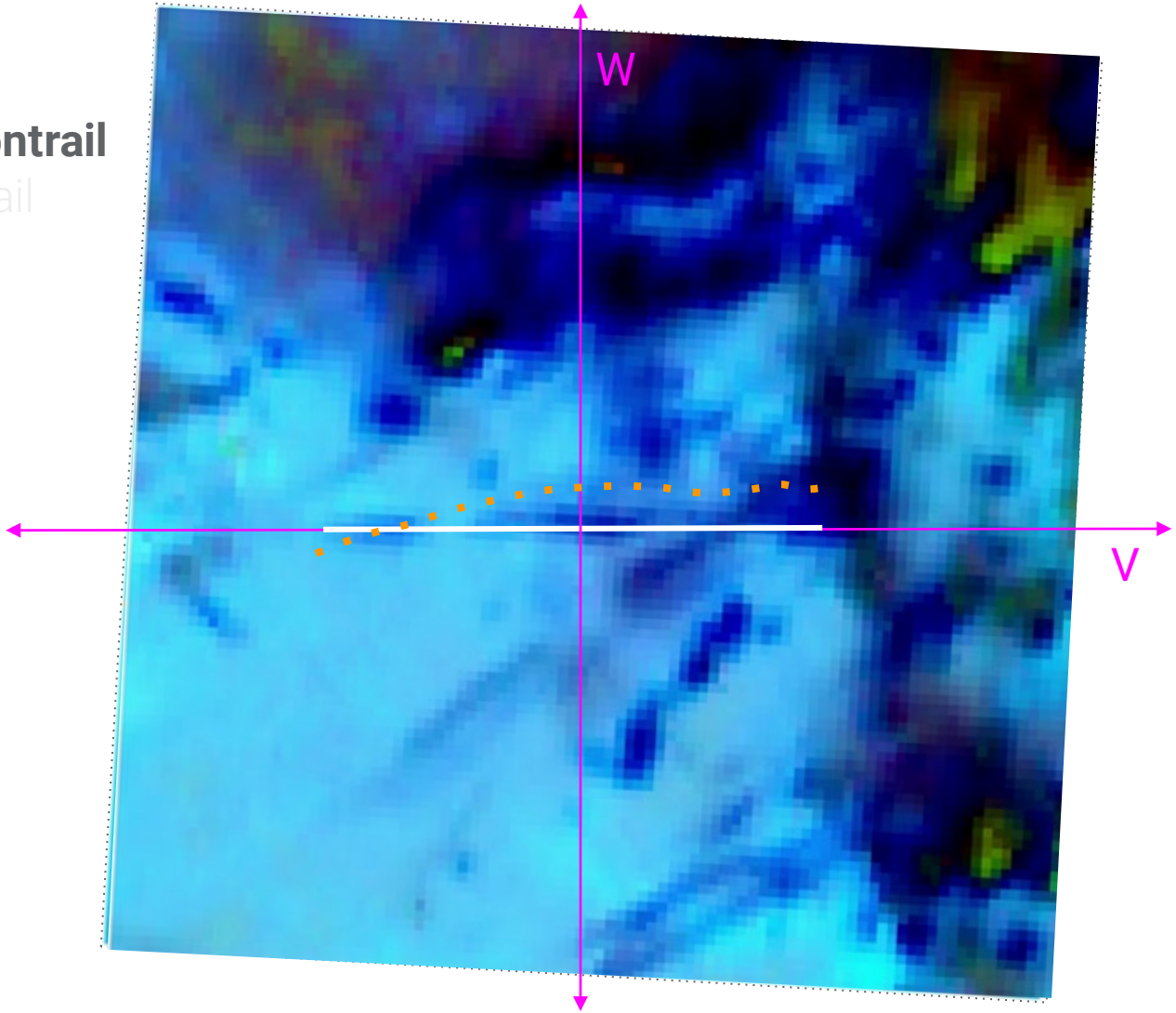
1. **Consider the contrail frame of reference**
2. Crop advected waypoints to length of contrail
3. Fit affine transform of points to the contrail



Google Research

Automated matching steps

1. Consider the contrail frame of reference
2. **Crop advected waypoints to length of contrail**
3. Fit affine transform of points to the contrail

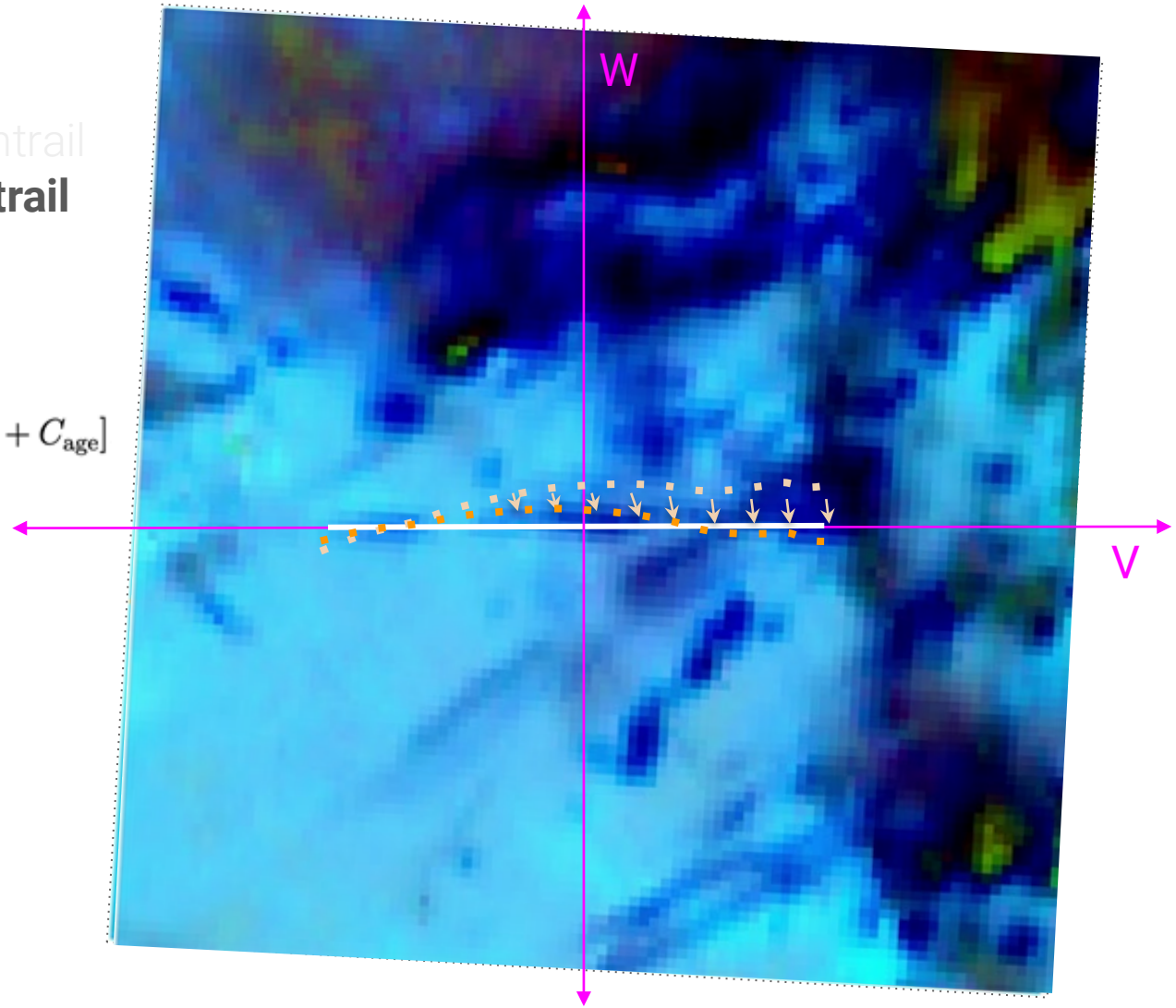


Google Research

Automated matching steps

1. Consider the contrail frame of reference
2. Crop advected waypoints to length of contrail
3. **Fit affine transform of points to the contrail**

match error = $[C_{\text{fit}} \frac{1}{N} \sum_{i=1}^{N} w_i^2 + C_{\text{shift}}(V^2 + W^2) + C_{\text{angle}}(1 - cos(\theta)) + C_{\text{age}}]$



Low match error indicates a good match.

Geraedts et al 2023, A scalable system to measure contrail formation on a per-flight basis (in review)

Automated matching - important details

- Geostationary imagery scans are not instantaneous.
- Use parallax correction, on the advected flight segment.
  - Flight segment altitude is better known than the contrail altitude.

Automated matching - notably, we are currently not using…

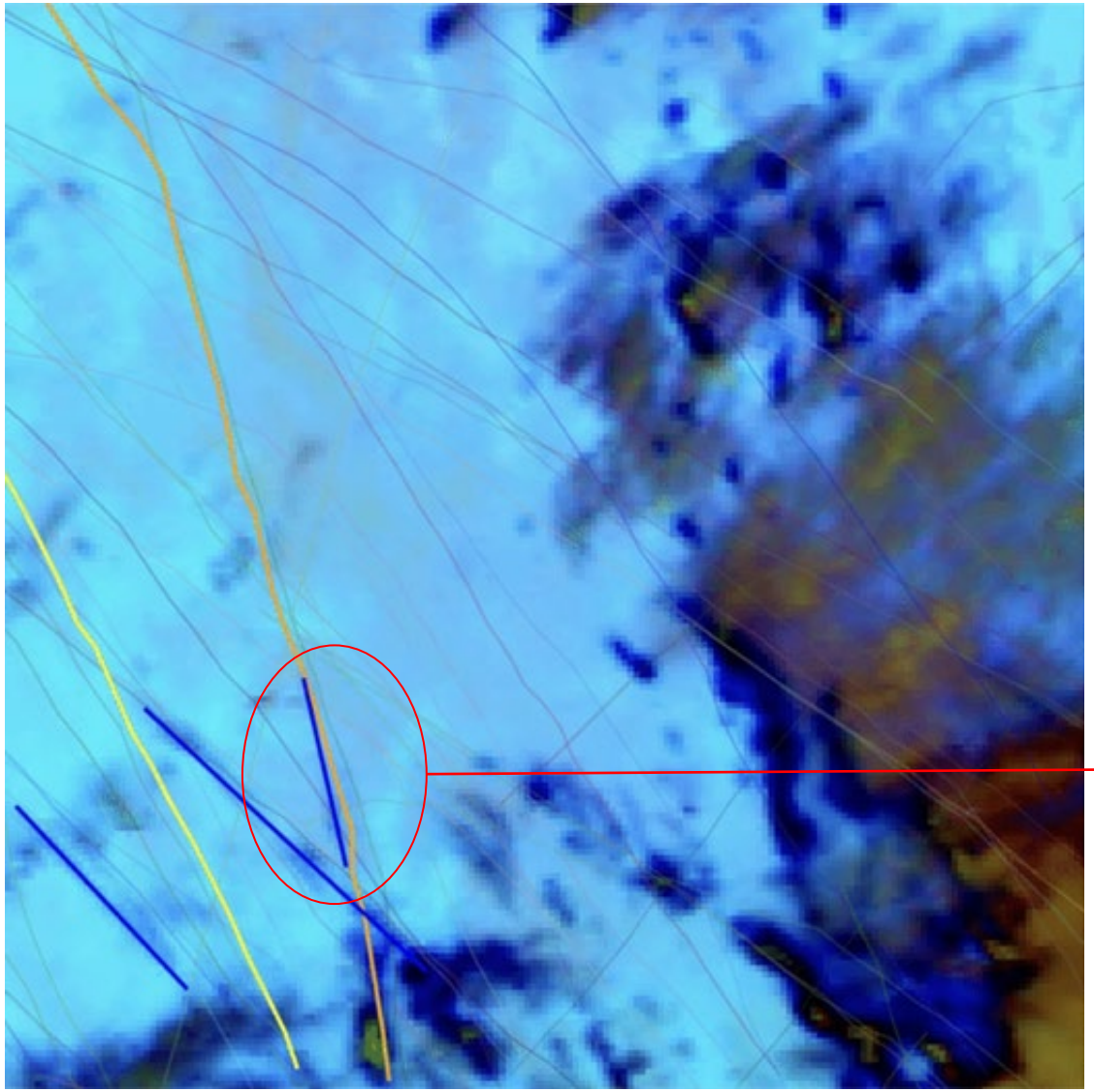
1. Estimated altitude of contrail
2. Multi-temporal sequences of inputs
3. Contrail width, shape

Google Research

1. Matching of flights to detected contrails
2. Forecasting

Google Research

Forecasting - labels for a learned model



Yellow - Flight path

Orange - Advected Flight Path

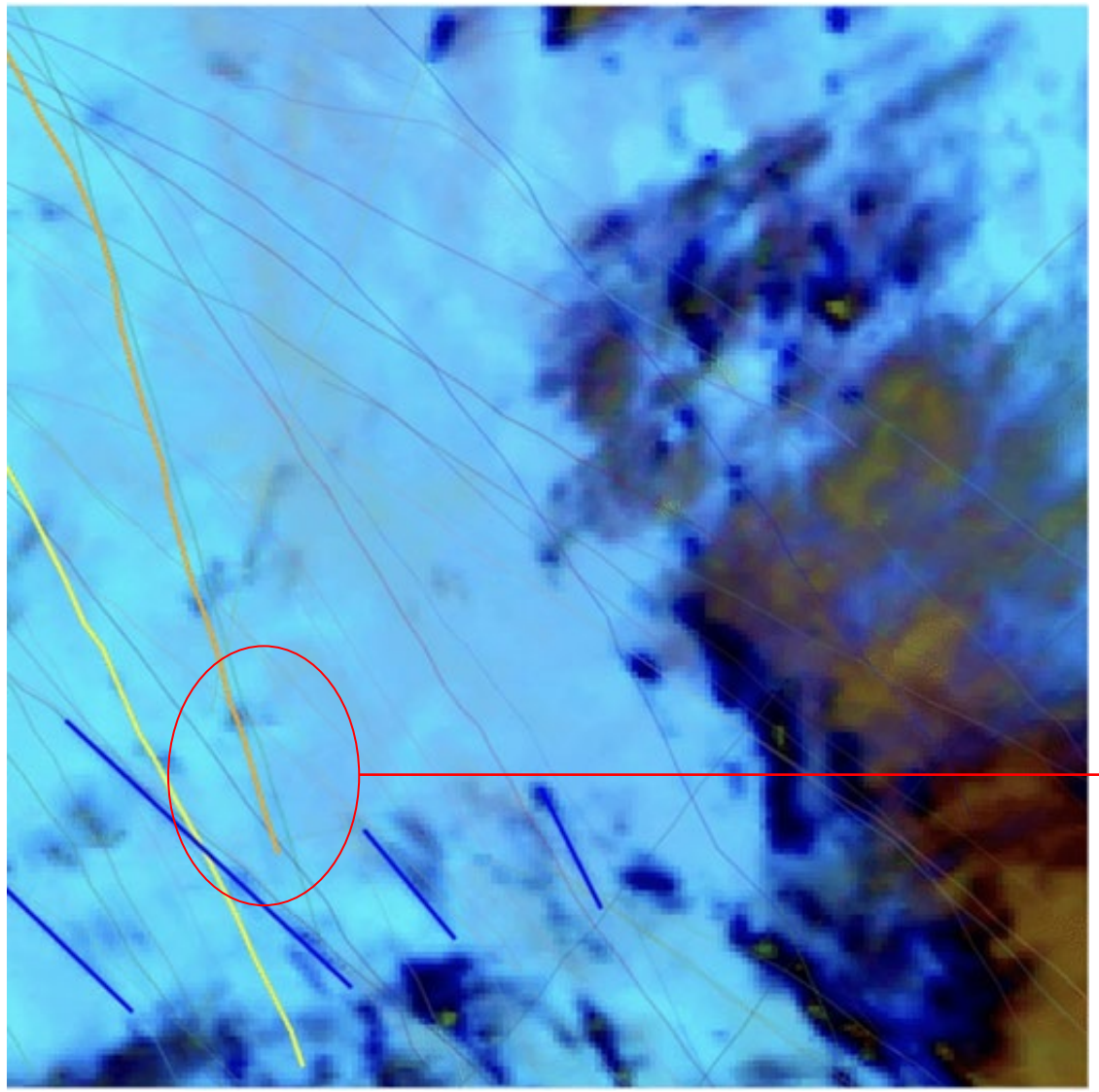
Dark - Other Advected Flights

Blue lines - Detected contrails

Positive example of contrail formation

Google Research

Forecasting - inputs for a learned model



Yellow - Flight path

Orange - Advected Flight Path

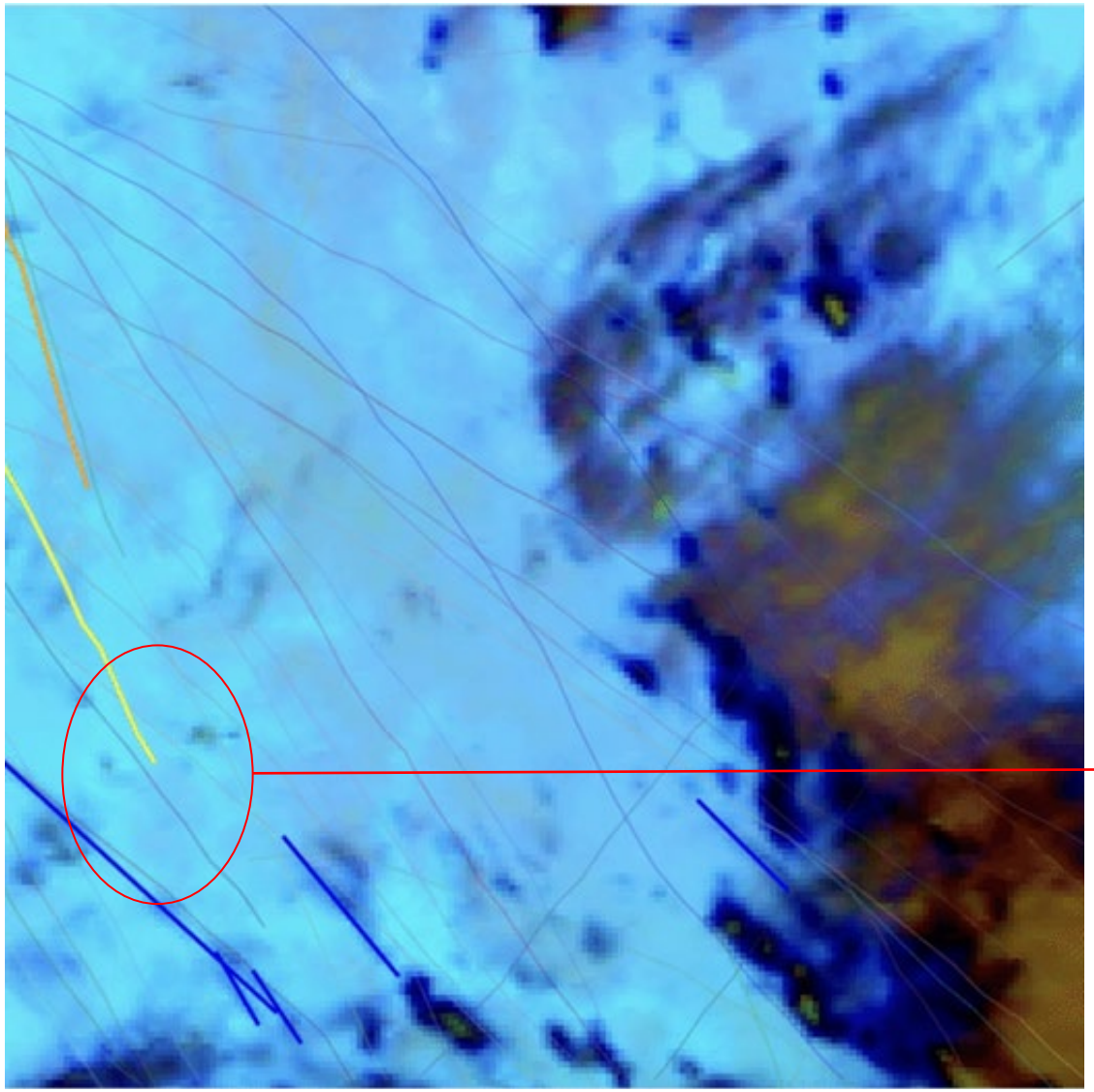
Dark - Other Advected Flights

Blue lines - Detected contrails

Sub-pixel contrail in ice-supersaturated humidity conditions

Google Research

Forecasting - inputs for a learned model



Yellow - Flight path

Orange - Advected Flight Path

Dark - Other Advected Flights

Blue lines - Detected contrails

Schmidt-Appleman criteria satisfied

Forecasting - inputs and architecture

Model inputs are primarily statistics of interpolated **numerical weather data** (typically ECMWF), found along the advection path of the flight segment:

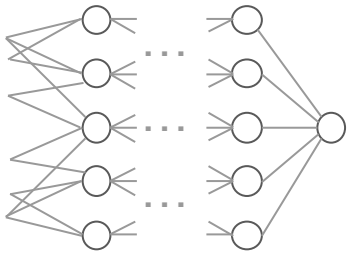
- relative humidity over ice
- wind vector components
- temperature
- specific humidity
- vorticity
- divergence
- cloud ice water content
- cloud snow water content
- fraction of cloud cover

X

- minimum
- 5th percentile
- 25th percentile
- mean
- 75th percentile
- max
- std-deviation

~100 features

4-layer neural net



binary label

and some other metadata:

- latitude
- longitude
- pressure-altitude of flight
- day of year
- local solar time

Forecasting - imbalanced data

Only ~3% of all flight segments over CONUS form a detected contrail.

This has implications…

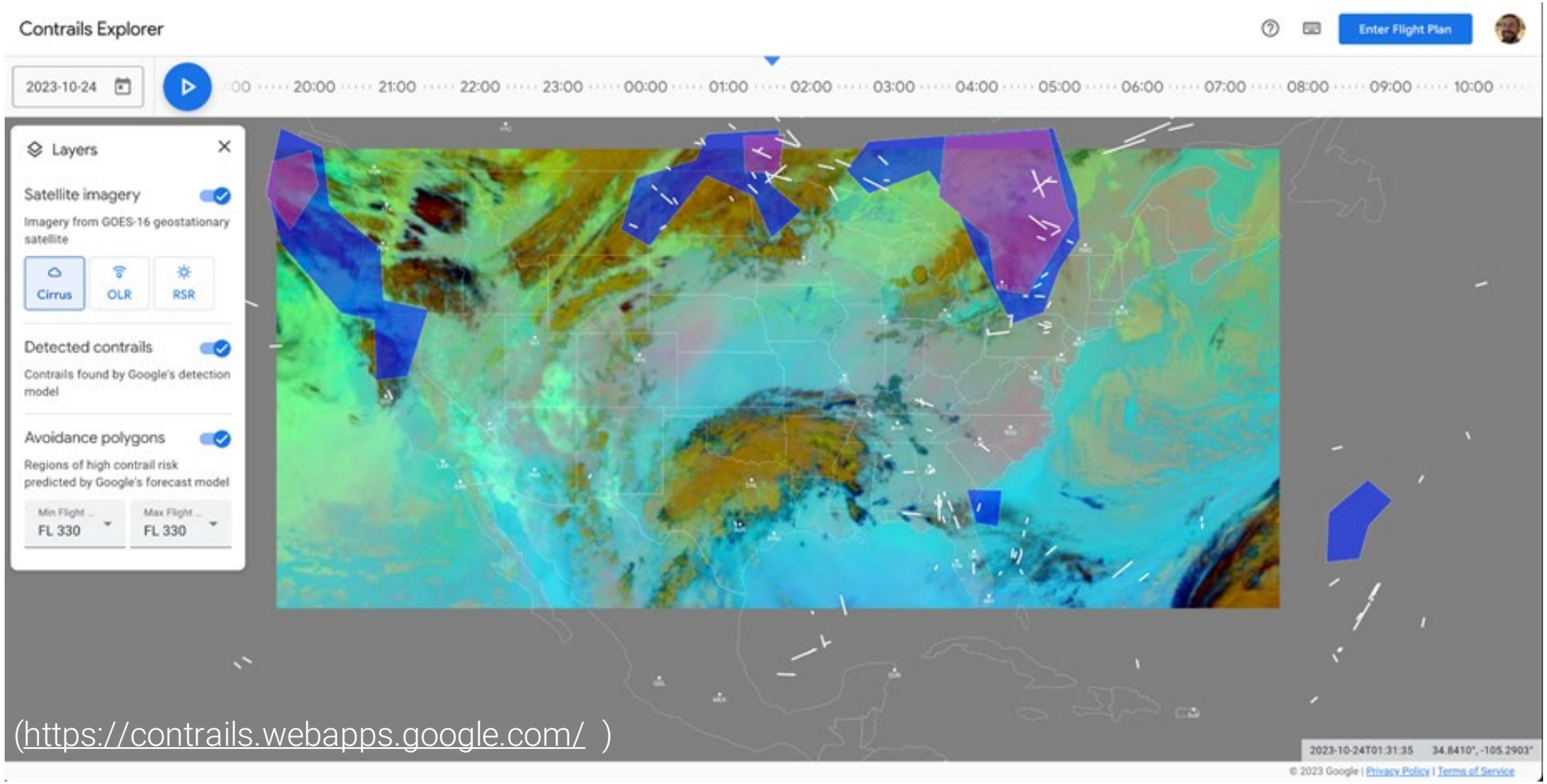
…for training the model:

- We primarily mitigate this by upsampling positives, until the training set is 50% positives.

…for evaluating the model:

- Precision/Recall metrics will appear higher when the positivity rate is higher, even for a model with the same accuracy.

Forecasting - inference



Also available in latitude x longitude x altitude NetCDF format on Google Cloud Storage

Google Research

Thank you, any final questions?

g.co/research/contrails
for further information and contact info