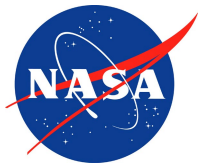


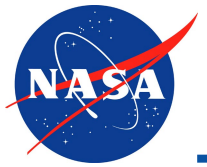
Clustering Days with Similar Airport Weather Conditions

Shon Grabbe and Banavar Sridhar
NASA Ames Research Center

Avijit Mukherjee
University of California - Santa Cruz

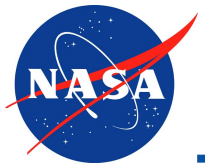
Motivation

- Develop automation to support post-operations analyses and day-of-operations planning
- Limited what-if capabilities for developing and planning traffic flow management initiatives
- Human operators often rely on past experience and intuition when developing initiatives



Research Achievements

- Airport-level clustering identified hours with similar probabilities of Ground Delay Programs occurring
- Consistently low usage of Ground Delay Programs under good weather and nighttime operations
- Consistency of Ground Delay Program usage is airport dependent under bad weather operations



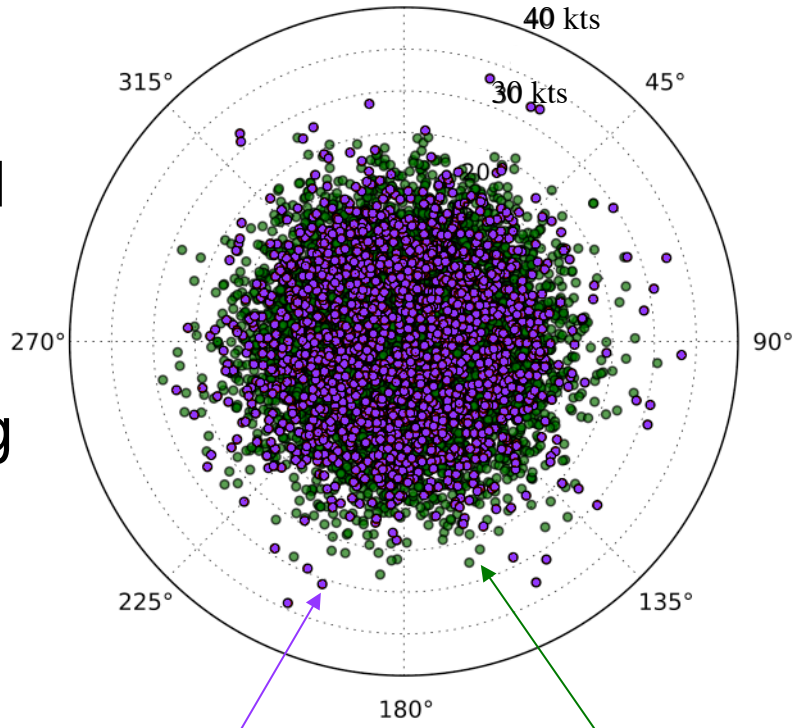
Outline

- Key Challenges
- Classification and Clustering Methodology
- Experimental Setup
- Results
- Potential Operational Use
- Conclusions

Key Challenges

- Noisy or incomplete data
- Data used for decision making may not be recorded (e.g., video camera feeds)
- Inconsistent decision making under similar weather and traffic conditions

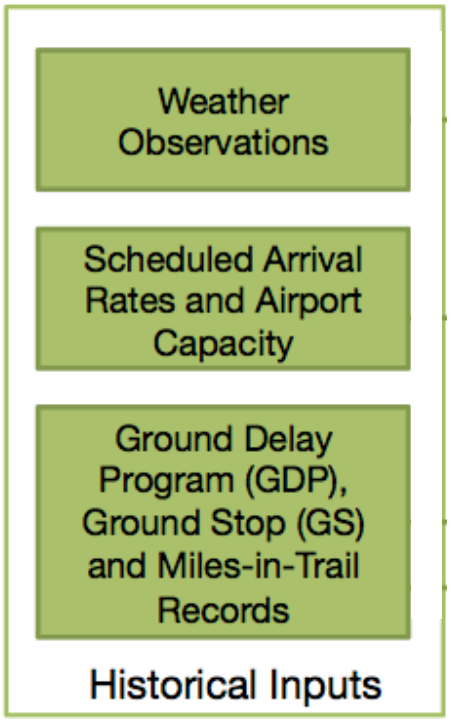
EWR Hourly Wind Magnitude and Direction for 2011



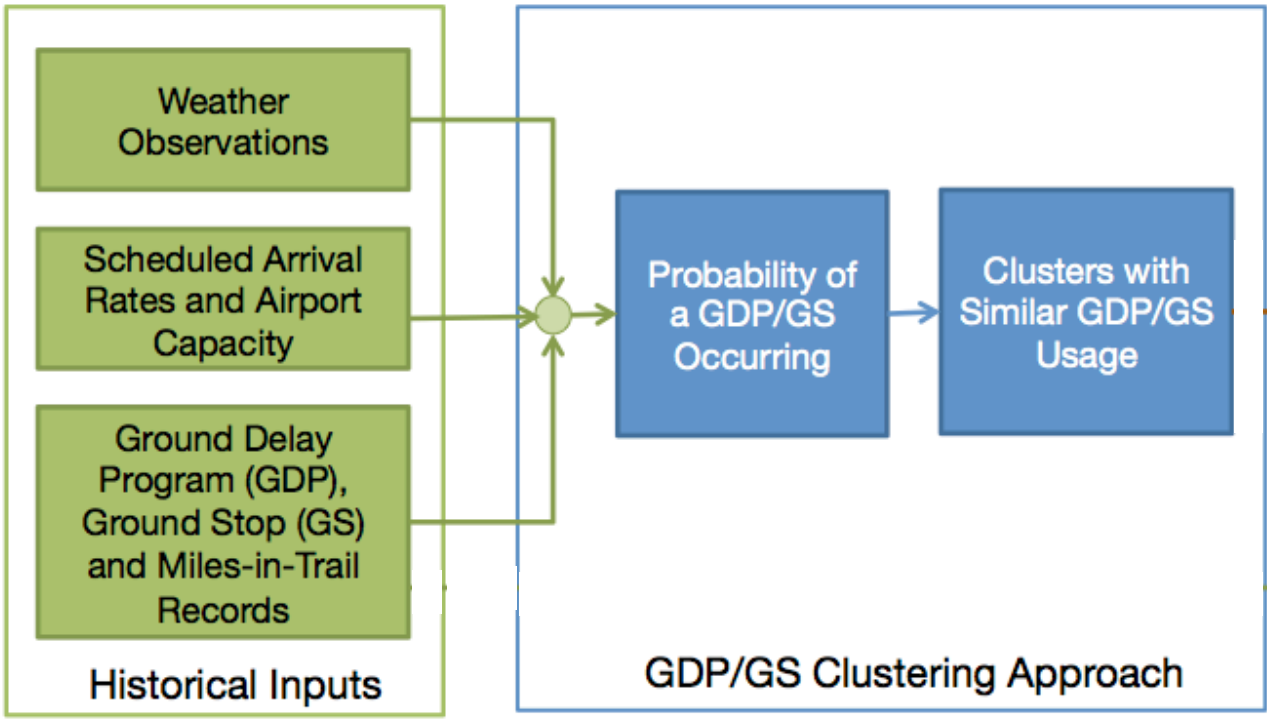
Ground Delay Program

No Ground Delay Program

Classification and Clustering Methodology

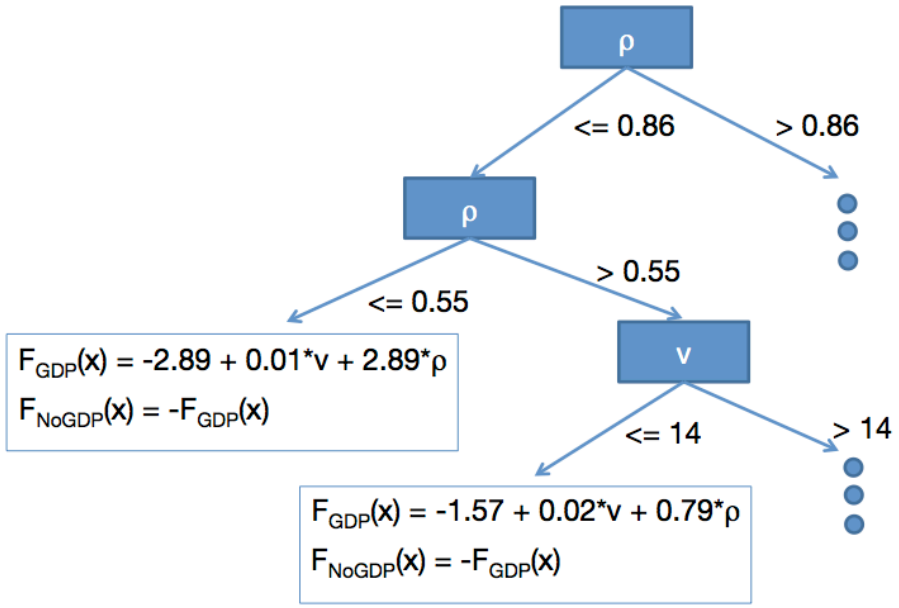


Classification and Clustering Methodology



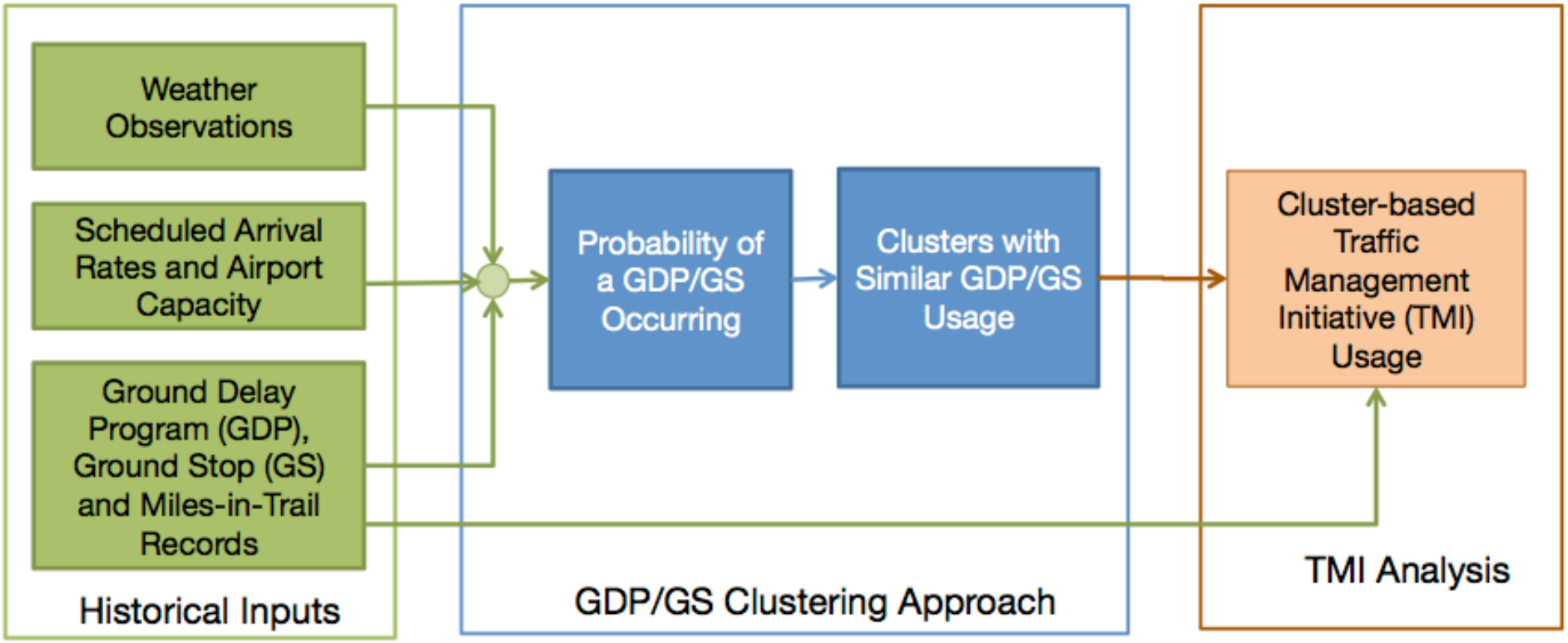
Probability of a Ground Delay Program

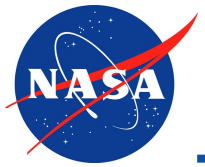
- Logistic Model Tree used to calculate the probability of a Ground Delay Program occurring
- Supervised learning technique
  - Tree induction models the complex structures in the data
  - Leaves of the tree contain simple logistic regression models



$$\Pr(G = GDP \mid X = x) = \frac{\exp(F_{GDP}(x))}{\exp(F_{GDP}(x)) + \exp(F_{NoGDP}(x))}$$

Classification and Clustering Methodology





Outline

- Key Challenges
- Classification and Clustering Methodology
- Experimental Setup
- Results
- Potential Operational Use
- Conclusions

2012 Hourly Prediction Attributes

Airports:

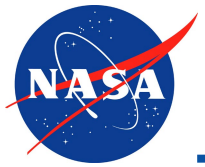
Newark Liberty International Airport (EWR) and Chicago O'Hare International Airport (ORD)

Weather Attributes (13 attributes):

- Wind direction, wind speed, probability of precipitation, ceiling height, visibility, probability of freezing precipitation, etc.
- Center Weather Impacted Traffic Index

Air Traffic Attributes:

Arrival demand and capacity

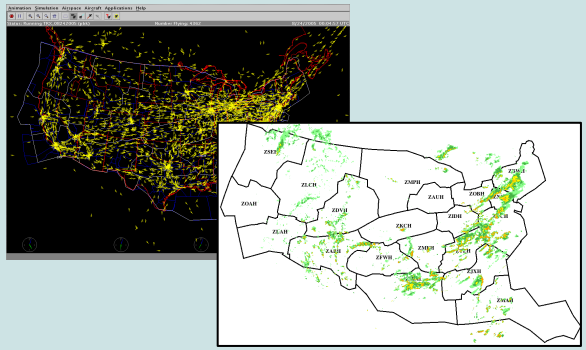


Outline

- Key Challenges
- Classification and Clustering Methodology
- Experimental Setup
- Results
  - Chicago O'Hare International Airport
  - Newark Liberty International Airport
- Potential Operational Use
- Conclusions

ORD: 5 Clusters

Historic Weather and Traffic



Morning and Nighttime Operations
58% of Observations

Fair Weather and High Daytime Operations
21% of Observations

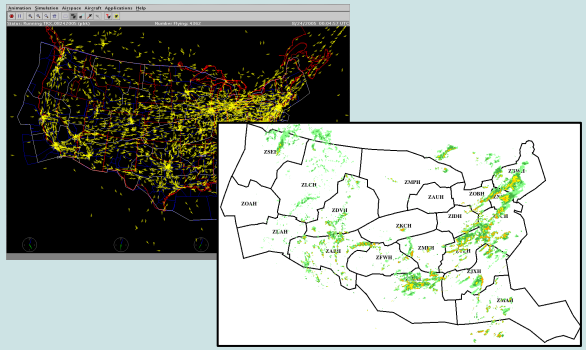
Fair Weather and Low Daytime Operations
11% of Observations

Reduced Ceilings and Elevated Probability of Precipitation
7% of Observations

Bad Weather
3% of Observations

ORD: GDP Probabilities

Historic Weather and Traffic



Morning and Nighttime Operations
<10% Probability of a GDP

Fair Weather and High Daytime Operations
<10% Probability of a GDP

Fair Weather and Low Daytime Operations
<10% Probability of a GDP

Reduced Ceilings and Elevated Probability of Precipitation
<20% Probability of a GDP

Bad Weather
11-100% Probability of a GDP

ORD: Take-away

- For good/fair weather and low arrival demand, the probability of a GDP occurring is low

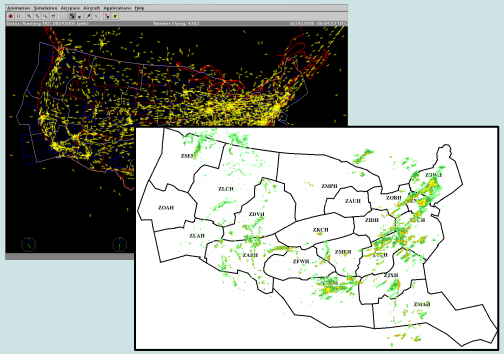
Fair Weather and High Daytime Operations
<10% Probability of a GDP

- For bad weather and daytime operations, GDP usage was very inconsistent

Bad Weather
11-100% Probability of a GDP

EWR: 13 Clusters

Historic Weather and Traffic



Nighttime Operations
26% of Observations

(2) Morning Operations
29% of Observations

(7) Winter Weather and/or Reduced Ceilings
39% of Observations

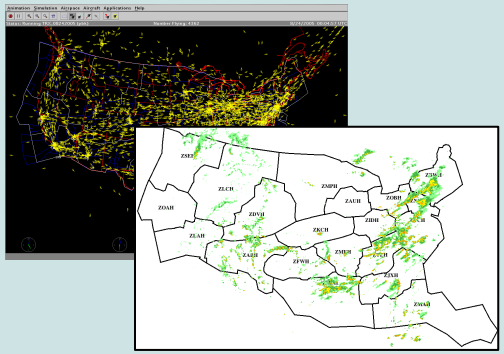
Elevated En Route Weather and Precipitation
2% of Observations

Heavy Arrival Demand with Elevated Precipitation
2% of Observations

Bad Weather
2% of Observations

EWR Clusters: GDP Probabilities

Historic Weather and Traffic



Nighttime Operations
< 10% Probability of a GDP

(2) Morning Operations
< 20% Probability of a GDP

(7) Winter Weather and/or Reduced Ceilings
21-90% Probability of a GDP

Elevated En Route Weather and Precipitation
51-70% Probability of a GDP

Heavy Arrival Demand with Elevated Precipitation
61-80% Probability of a GDP

Bad Weather
>91% Probability of a GDP

EWR: Take-away

- For good weather and low arrival demand, the probability of a GDP occurring is low

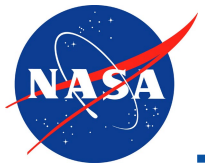
(2) Morning Operations
< 20% Probability of a GDP

- For winter weather and low ceiling operations, the use of GDPs is highly variable

(7) Elevated Freezing Precipitation and Reduced Ceilings
20-90% Probability of a GDP

- For bad weather and daytime operations, a GDP is almost always used

Bad Weather
>91% Probability of a GDP

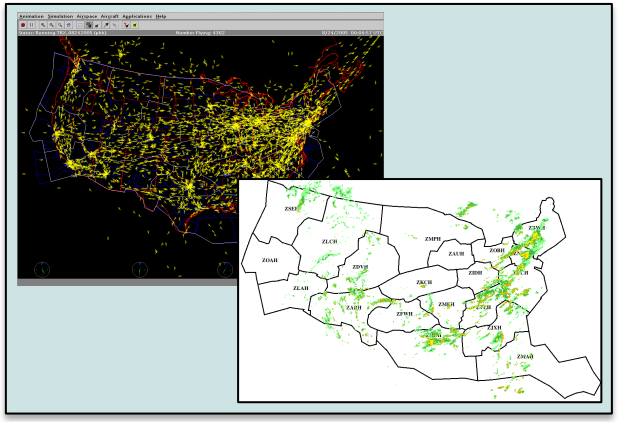


Outline

- Key Challenges
- Classification and Clustering Methodology
- Experimental Setup
- Results
- Potential Operational Use
- Conclusions

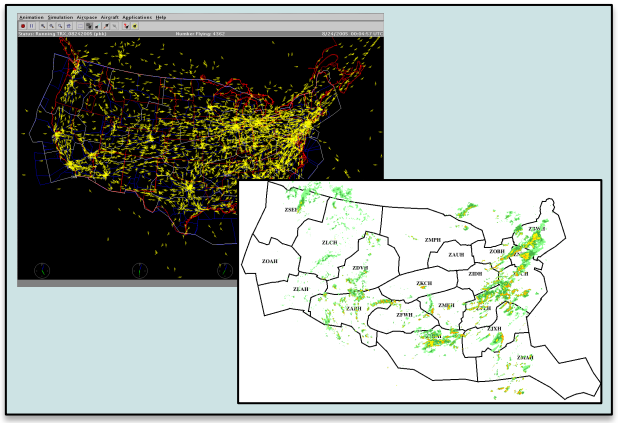
Potential Operational Use

Today's Weather and Traffic Observations and Forecasts



Potential Operational Use

Today's Weather and Traffic Observations and Forecasts



Clusters from Historic Weather and Traffic Observations and Forecasts

Morning and Nighttime Operations
58% of Observations

Fair Weather and High Daytime Operations
21% of Observations

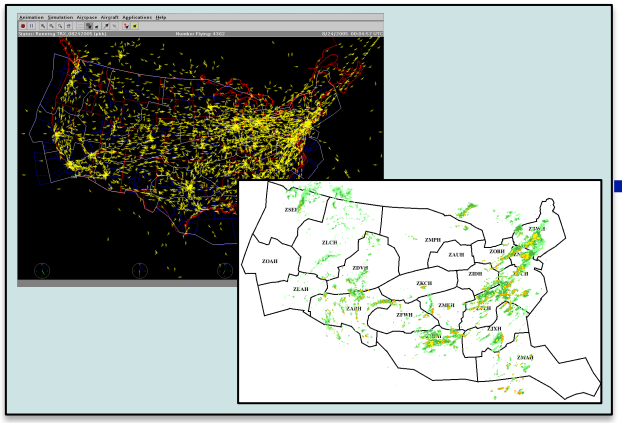
Fair Weather and Low Daytime Operations
11% of Observations

Reduced Ceilings and Elevated Probability of Precipitation
7% of Observations

Bad Weather
3% of Observations

Potential Operational Use

Today's Weather and Traffic Observations and Forecasts



Clusters from Historic Weather and Traffic Observations and Forecasts

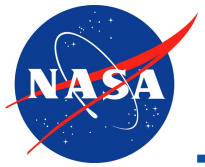
Morning and Nighttime Operations
58% of Observations

Fair Weather and High Daytime Operations
21% of Observations

Fair Weather and Low Daytime Operations
11% of Observations

Reduced Ceilings and Elevated Probability of Precipitation
7% of Observations

Bad Weather
3% of Observations



Conclusions

- Airport-level clustering identified hours with similar probabilities of Ground Delay Programs occurring
- Consistently low usage of Ground Delay Programs under good weather and nighttime operations
- Consistency of Ground Delay Program usage is airport dependent under bad weather operations