Airspace Technology Demonstration 2 (ATD-2)

Accuracy Comparison of Various Landing Time Prediction Sources

May 22, 2019
Quantify accuracy of different sources of landing (on) time predictions as actual arrival event approaches

• Want to make design decisions for fuser mediation rules informed by data about actual accuracy of various potential prediction sources

• Accuracy defined as difference between actual and prediction, particularly interested in how this evolves as actual event approaches

• Other sources could easily be included in this framework, e.g., operator-generated predictions
• Two weeks of arrivals to Charlotte from March 2019
• Comparing:
  – TBFM ETA
  – TBFM STA
  – TBFM STA (only when frozen)
  – TBFM STA (before frozen)
  – TFM ETA
• Measure error as actual landing time – prediction
• Sample every minute for every flight, then average in a variety of ways
How many flights have predictions?

• Consistent with expectations – TFM available well ahead of time, then TBFM ETA, then STAs begin appearing and are frozen
Mean is a potentially troublesome measure because positive and negative errors may cancel each other out.

Strange that accuracy gets worse as landing approaches!
What are the MAD errors?

- Plot shows mean of absolute difference (error), weighting positive and negative errors equally, but showing best measure of “average” error.

Recall that there are few flights with this prediction.

About 75 minutes out, TBFM ETA beats TFM ETA, with similar number having this prediction.
Are predictions worse before takeoff?

- Plot shows MAD leading up to takeoff time for same set of flights. Only TFM ETA widely available, but TBFM ETA becomes available 60-80 minutes pre-departure.
- Conclusion: predictions somewhat poor pre-departure.
Plot shows MAD leading up to landing time, but only includes flights that have already departed.

Post-departure, predictions are much better than pre.

Seems clear that TBFM provides best estimates at most lookaheads.

Are predictions better after takeoff?

Similar number of flights available at each epoch for these four predictions.

Note vertical scale only goes to 10 minutes.
• Prediction accuracy generally improves as landing time approaches, as expected
• Not all errors converge to zero
• Demonstration of feasibility of comparing landing time prediction accuracy of various data sources
  – *This work could be replicated with a big pile of data captured directly from SWIM feeds*
  – *But, this work is significantly simpler when using data that has passed through the fuser / ATD-2 system*