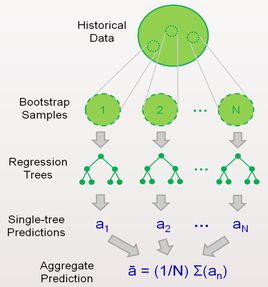
Creating Weather Impact Scores from Current Weather Forecast Products for a Probabilistic Airport Capacity Model

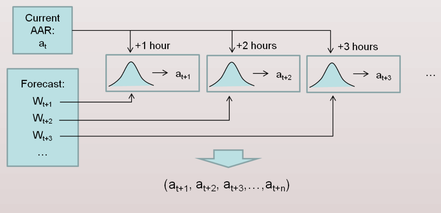
Jonathan Cunningham, Chris Provan, and Lara Cook*
Mosaic ATM, Leesburg, Virginia

Background

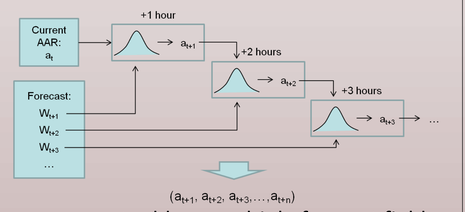
- The Weather Translation Model for GDP Planning (WTMG) is a two-part, self-training statistical model that focuses on predicting the airport arrival rate (AAR) in the presence of weather for the purpose of planning Ground Delay Programs (GDP).
  - Prediction Model: Trained with historical weather forecasts and observed AARs and using Matlab's *TreeBagger* class , a bootstrapped class regression tree methodology is used to create deterministic AAR predictions.
  - Sampling Model: Builds an empirical error distribution around each deterministic AAR prediction and creates a set of capacity scenarios from the current time period to ten hours into the future.



- WTMG runs in two modes:
  - Static Mode – generates each future hour's probabilistic AAR prediction based on the forecast information available at the time of the prediction.



  - Dynamic Mode – generates each future hour's probabilistic AAR prediction based on the previous hour's sample AAR.



- The regression tree methodology, however, cannot address multiple forecast fields whose impacts on arrival operations are dependent on one another. A set of three numeric impact scores were developed to replace the hourly forecast fields of the LAMP and the TAF.

Ceiling/Visibility Impact Score

- TAF forecasts are translated into LAMP's ceiling and visibility threshold categories.
- For each lead time from forecast issue time, scores are computed by running a linear regression on the binary variables representing each ceiling and visibility category as predictors of AAR reduction.
- The highest ceiling and visibility categories are fixed at zero.

Wind Impact Score

- Using runway heading information at each airport, individual tailwind and crosswind scores are computed for each runway based on the tailwind and crosswind components of the forecast wind speed and direction.
- The scores are based on respective lower and upper impact bounds of 5 and 10 knots for tailwind and 15 and 25 knots for crosswind. This scales the tailwind and crosswind scores between 0 and 1.
- Each runway is assigned an impact score computed as the maximum of the tailwind and crosswind scores for that runway.
- The overall wind impact score for the airport is computed as the mean of the runway scores for all runways at the airport multiplied by the mean observed AAR for the airport.

Weather Impact Score

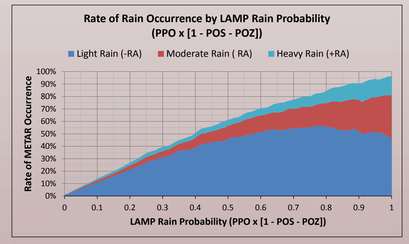
TAF

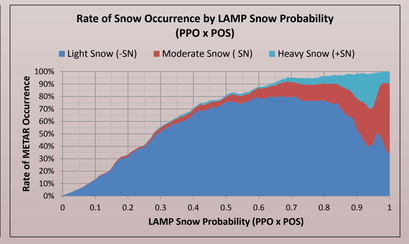
- Based on the significant weather factors and associated modifiers found TAF forecast.
- For each lead time from forecast issue time, the factors are converted to a impact score by running a linear regression on the binary indicators showing the presence of each weather factor in the forecast as a predictor of AAR.

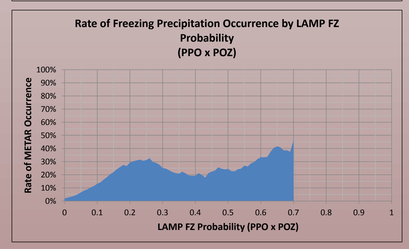
TAF	EWR (1 hr)	EWR (6 hr)	ORD (1 hr)	ORD (6 hr)
Light (-)	-1.9	-1.6	-5.8	-6.2
Heavy (+)	3.4	2.0	17.8	n/s
RA	4.3	4.0	21.5	21.5
DZ	5.8	5.2	24.7	26.9
TS	n/s	n/s	4.6	n/s
FZ	7.5	13.7	12.6	4.5
SN	9.8	9.6	27.6	28.1
PL	15.1	13.9	21.5	34.7
GS	n/s	n/s	33.8	n/s

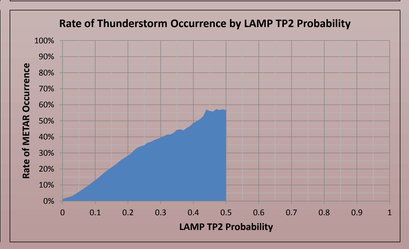
LAMP

- Using the raw and conditional probabilities, absolute probabilities are computed for rain, snow, and freezing precipitation.
- Although LAMP does not include severity indicators, there is correlation between high probability forecast and increased severity in the observed weather.







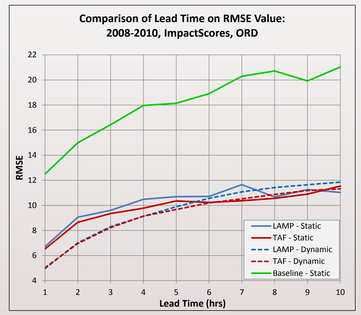


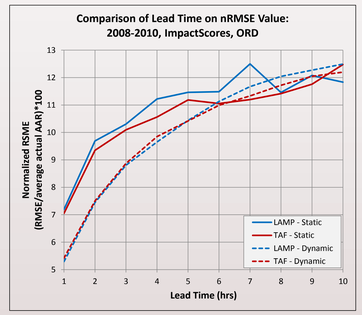
- To convert to an impact score, a linear regression is run on the absolute probabilities of rain, snow, freezing precipitation, and thunderstorms which are used as indicators for the reduction of AAR.

LAMP	EWR (1 hr)	EWR (6 hr)	ORD (1 hr)	ORD (6 hr)
Rain	0.044	0.056	0.27	0.37
Snow	0.13	0.18	0.39	0.51
Freezing Precipitation	0.14	0.20	0.37	0.69
Thunderstorm	0.0059	0.15	0.40	0.58

Results

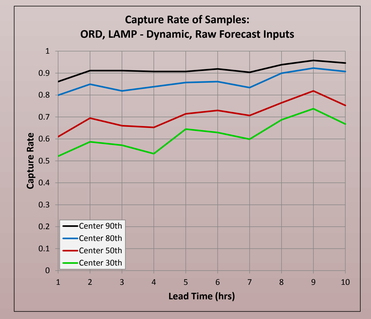
- Four different versions of the WTMG model were built and tested based on the two different modes and forecast products: LAMP – Static, LAMP – Dynamic, TAF – Static, and TAF – Dynamic.
- The primary metric used to evaluate the prediction model within WTMG is the root mean squared error (RMSE) between the predicted AAR and the actual AAR. The normalized RMSE (nRMSE) represents the RMSE as a percentage of the average AAR at the airport.

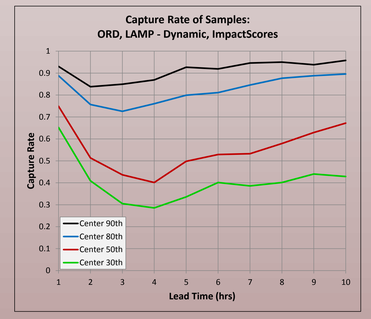




"Baseline – Static" represents baseline models that assume persistence of AAR through time.

- To evaluate the uncertainty generated by the sampling model, a capture rate was used. This method finds the frequency that the actual AAR was captured in the central x^{th} percentile of the sample AARs at each lead time.





For a perfect error distribution, the capture rate must match the percentile; e.g., the 50th percentile capture rate would be 0.5.

Conclusions

- Calculated weather impact scores provided a satisfactory translation of raw forecast fields for the purposes of WTMG.
- No significant difference was observed between the results of the deterministic prediction models that used raw forecast inputs and the prediction models that used impact score inputs.
- However, for the sampling models, there was a noted improvement in the impact score input models when compared to raw forecast input models.
- While additional work is needed to further enhance the uncertainty models, WTMG performed well and showed significant improvements over baseline weather-naïve models at all of the airports studied.

*Corresponding Author: Jonathan Cunningham, Mosaic ATM, Leesburg, VA *email:* jcunningham@mosaicatm.com