A Recursive Multi-step Machine Learning Approach for Airport Configuration Prediction

Shaymaa Khater* , Juan Rebollo*, William J. Coupe†

* Mosaic ATM, †NASA Ames Research Center

Context

Airport configuration selection is a complex decision making process that involves operational and human factors

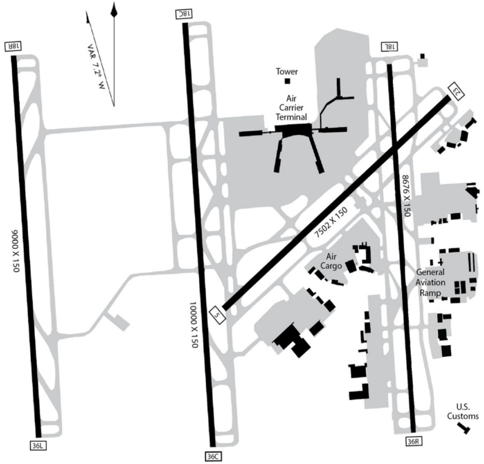
Objective: predict sequence of airport configurations at a given airport over time, out to six hours from current time

Two major challenges:

- There are many valid airport configurations, requiring novel encoding techniques.
- Relationship between configurations in subsequent time periods (i.e., autocorrelation)

Example airport configurations

Airport	Configuration
CLT	D_18C_18L_A_18C_18L_18R
	D_36C_36R_A_36C_36L_36R
	D_18L_A_18L
	D_36C_A_36C_36L
JFK	D_22R_A_22L_22R
	D_13R_A_13L
	D_31L_A_31L_31R

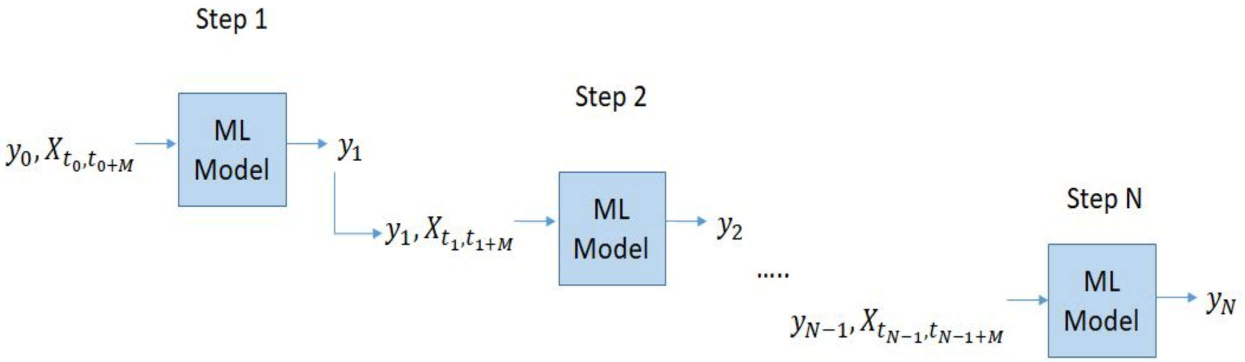


Source: https://www.faa.gov/airports/planning_capacity/profiles/media/CLT-Airport-Capacity-Profile-2015.pdf

Approach

Recursive multi step forecasting approach

- Prediction from the prior step is used as input to generate prediction for the following time step.
- Having the prior configuration values as input guarantees the stability of the predicted configuration



Recursive Multi-step Forecasting.
M: is the number of steps in the model look ahead
N: is the number of steps in the prediction look ahead,
Y: the target variable
$X_{ta,tb}$: is the feature vector with feature data from time step t_a to time step t_b

ML Model: Random Forest classifier model.

Data & Features

- Weather Forecast data
  - Obtained from Localized Aviation MOS Program (LAMP).
  - Products in the LAMP system are updated hourly and are valid over a 25-hour period.
- Airport configuration
  - Obtained from Data-link Automatic Terminal Information Service (D-ATIS) messages
  - Defines our target variable and the current configuration feature

Feature	Example Value
Cloud ceiling forecast	8
Lightning probability forecast	L
Precipitation in forecast	TRUE
Temperature forecast	70
Visibility forecast	7
Wind direction forecast	12
Wind speed forecast	20
Wind gust forecast	4
Current airport configuration	D_17R_18L_A_17C_18R

Data & Features

- Future arrival/departure count
  - Use estimated time of arrivals (ETAs) from different systems and defined a set of rules.
  - The obtained physics-based ETAs are used to calculate future arrival counts input to our model
- Time of the day

Feature	Example Value
Predicted future arrival counts	10
Predicted future departure counts	15
Time of the day	2021-01-10 10:00:00

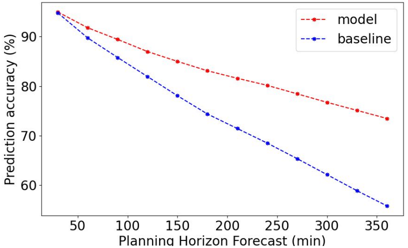
Each of these features provided at multiple lookahead times to generate prediction

Experimental Design

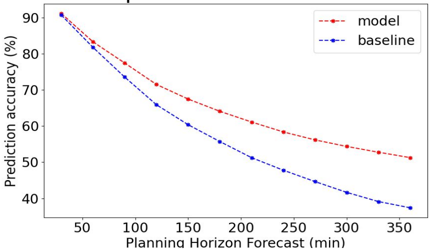
- Training/Testing data
  - Dates range: August- Dec 2019
  - 80% of the data for training and 20% for testing.
  - Data was split on a weekly basis to avoid correlation and data leakage between training and testing
- For airports:
  - DFW
  - DAL
  - CLT
  - JFK
  - EWR
  - LGA
- XGBoost Classifier algorithm used for training models

Experimental Results -2019

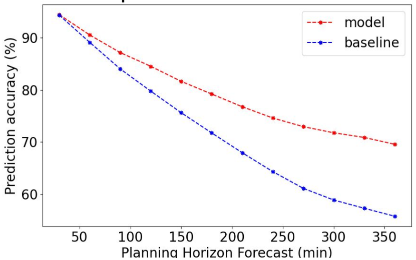
CLT airport



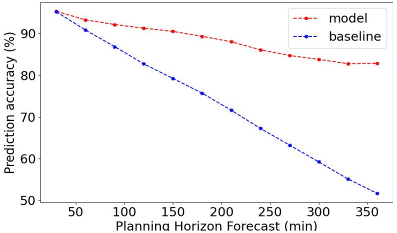
JFK airport



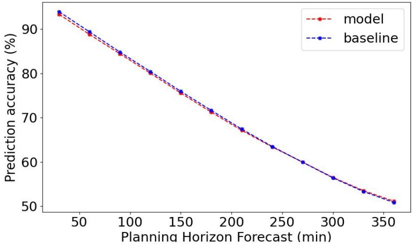
EWR airport



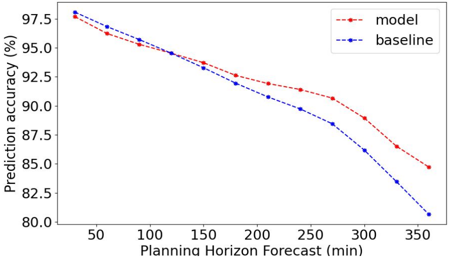
DFW airport



DAL airport



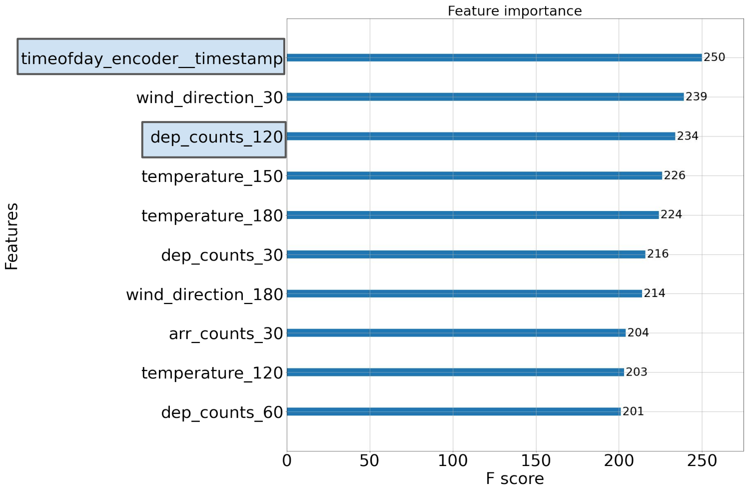
LGA airport



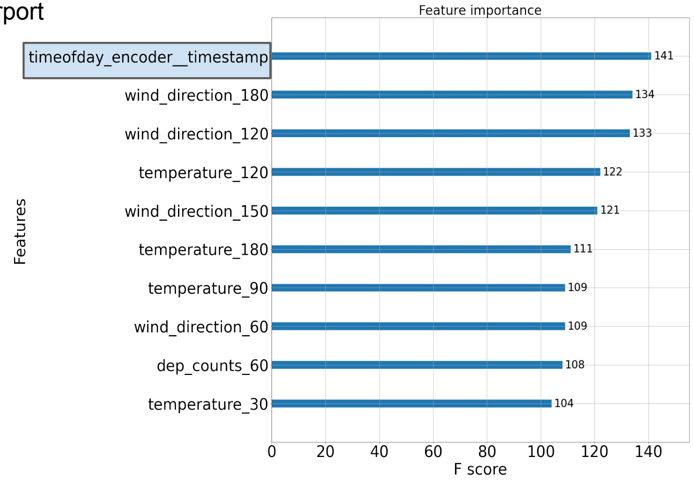
- Naive baseline assumes no changes in the configuration.

Experimental Results -Features Importance

DFW airport



DAL airport



- Traffic (i.e, arrival and departure counts), the time of the day and wind are the most important features deriving the prediction.
- Poor performance at DAL is due to the lack of predictive power of the traffic features.

Conclusion

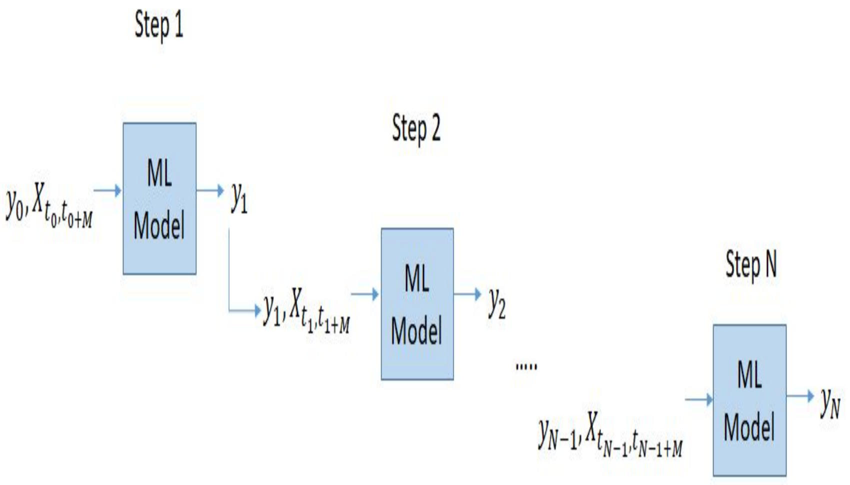
- The ML model performs similarly to the baseline for short horizon of time, but outperform it for longer horizon.
- The approach guarantees the stability of the predicted configuration by taking into account the predicted configuration at the previous time step.
- Traffic counts, time of the day, wind speed and directions features are the most important features affecting the airport configuration changes.
- The proposed model can be integrated in real time system and deployed for many airports.

Thank you

Backup - Approach

- Define three parameters
  - Step size : 30 mins
  - Overall lookahead time: 6 hrs
  - Prediction model's lookahead time: 3 hrs

For airport configuration, the first step value is the current configuration



Recursive Multi-step Forecasting.
M: is the number of steps in the model look ahead
N: is the number of steps in the prediction look ahead,
Y: the target variable
$X_{ta,tb}$: is the feature vector with feature data from time step t_a to time step t_b

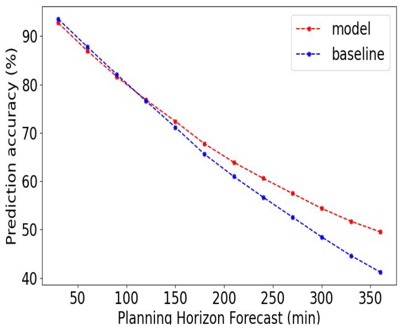
ML Model: Random Forest classifier model.

Experimental Results -2020

- In DFW, there is large gap between the importance of the time of the day feature and the other features.

- As a result of low traffic levels , traffic features don’t have an important role in model’s prediction accuracy.

DFW airport



DAL airport