Predicting airport runway configurations for decision-support using supervised learning

Tejas G Puranik, Milad Memarzadeh and Krishna M Kalyanam
USRA, NASA Ames Research Center, tejas.puranik@nasa.gov
USRA, NASA Ames Research Center, milad.memarzadeh@nasa.gov
NASA Ames Research Center, krishna.m.kalyanam@nasa.gov

One of the most challenging tasks for air traffic controllers is runway configuration management (RCM). It deals with the optimal selection of runways to operate on (for arrivals and departures) based on traffic, surface wind speed, wind direction, other environmental variables, noise constraints, and several other airport-specific factors. It affects the efficiency of the National Airspace System (NAS) and both surface and airspace operations can benefit from better understanding future runway configurations.

A key tenet of the next generation air transportation system (NextGen) is deploying automation capabilities through decision-support tools (DST) to aid air traffic controllers. Among such automation capabilities is the application of data-driven machine learning approaches to complex problems in air traffic management (ATM). There are two main types of data-driven analyses around runway configuration management – **prescriptive** and **predictive** models. The main goal of **prescriptive** models is to recommend the *most optimal* runway configuration for a given situation based on prevailing conditions, intended system goals, and constraints. The techniques in this type of application would include optimization, reinforcement learning, etc. Air traffic controllers or other decision-makers are generally the intended stakeholders for such prescriptive models. **Predictive** models on the other hand, do not aim to improve existing decision-making frameworks but rather try to learn them and to better predict future configurations. These models try to capture existing heuristics typically through supervised machine learning models. Both air traffic management personnel and operators can benefit from predictive models through reduction of airport capacity uncertainty, better traffic flow management, and better utilization of resources and planning. Both prescriptive and predictive models are important and have received significant attention in literature, with more diverse applications seen in the prescriptive models.

In this paper, we present a comprehensive implementation of **predictive** models for runway configuration estimation from large volumes of historical data. Specifically, operational data from two full years (2018 and 2019) is collected, analyzed, and fused together to build the data product used in this work. The data set differs from prior work in the field in terms of its scope, resolution, and variety of factors collected and considered. Meteorological data is collected from two different sources – current weather conditions from METAR (Meteorological Terminal Aviation Routine Weather Report) and forecast weather conditions from Localized Aviation MOS Program (LAMP). Operational data from the Federal Aviation Administration (FAA) Aviation System Performance Metrics (ASPM) related to scheduled and actual number of arrivals and departures, average taxi times, etc. are collected. NASA's Sherlock Data Warehouse is used to identify critical information such as go-arounds, and other events that might impact RCM decision-making.

All data is collected and aggregated over 15-minute intervals throughout the two years. This provides a resolution like the timescales that might be necessary for runway configuration management decision-making. A variety of supervised learning algorithms are tested including Support Vector Machine, Random Forest, Gradient Boosting, etc. including tuning of the model hyperparameters. The modeling process is applied and presented on two representative U.S. airports – Charlotte Douglas International Airport (KCLT) and Denver International Airport (KDEN). The two airports present different levels of complexity in terms of the total number of configurations used and provide a balanced perspective on the generalizability of the developed approach to other airports in the NAS. Initial results are promising (F1 score of 0.91 at KCLT and 0.83 at KDEN) for data in the test set. The final paper will contain a comprehensive comparison between different models and model building strategies as well as further refined results. Most important predictors for each airport will be identified along with a discussion and recommendations on adapting the framework to other scenarios.