Predicting Airport Runway Configuration

A Discrete-Choice Modeling Approach

Jacob Avery and Hamsa Balakrishnan
Department of Aeronautics and Astronautics
Massachusetts Institute of Technology
Cambridge, MA, USA

Abstract—**The runway configuration is a key driver of airport capacity at any time. Several factors, such as weather conditions (wind and visibility), traffic demand, air traffic controller workload, and the coordination of flows with neighboring airports influence the selection of runway configuration.**

This paper identifies a discrete-choice model of the configuration selection process from empirical data. The model reflects the importance of various factors in terms of a utility function. Given the weather, traffic demand and the current runway configuration, the model provides a probabilistic forecast of the runway configuration at the next 15-minute interval. This prediction is then extended to obtain the 3-hour probabilistic forecast of runway configuration. The proposed approach is illustrated using case studies based on data from LaGuardia (LGA) and San Francisco (SFO) airports, first by assuming perfect knowledge of weather and demand 3-hours in advance, and then using the Terminal Aerodrome Forecasts (TAFs). The results show that given the actual traffic demand and weather conditions 3 hours in advance, the model predicts the correct runway configuration at LGA with an accuracy of 82%, and at SFO with an accuracy of 85%. Given the forecast weather and scheduled demand, the accuracy of correct prediction of the runway configuration 3 hours in advance is 80% for LGA and 82% for SFO.

Keywords- runway configuration; air traffic control decision-making; discrete-choice models; data-driven modeling

I. INTRODUCTION

Airport congestion leads to significant flight delays at the busiest airports around the world. Such congestion occurs when the demand for aircraft operations exceeds the available airport capacity. Airport expansion projects to increase capacity tend to be expensive and take several years to complete; by contrast, the better utilization of existing airport capacity is a less expensive approach to mitigating congestion. The key driver of airport capacity at a given time is the active runway configuration [1], which is chosen by air traffic control personnel, taking into consideration many different factors such as wind speed, wind direction, meteorological conditions, arrival demand, departure demand, noise mitigation, and coordination with surrounding airports.

Airport capacity predictions are also needed for air traffic flow management [2, 3], airport surface operations scheduling [4], and system-wide simulations [5]. Since the capacity of an airport depends on the runway configuration being used, the forecast of the runway configuration is a key step toward predicting the capacity of an airport.

This paper develops a data-driven model of the runway configuration selection process using a discrete-choice modeling framework. The approach infers the air traffic controllers' utility functions that would best explain (that is, maximize the likelihood of) the observed decisions. The resultant model yields a probabilistic prediction of the runway configuration at any time, given a forecast of the influencing factors.

A. Related work

There are two main classes of models that have been developed for runway configuration selection: prescriptive models and descriptive models. The former aim to recommend an optimal runway configuration, subject to operational constraints. An early example of such a model is the Enhanced Preferential Runway Advisory System (ENPRAS) that was developed for Boston Logan International Airport [6]. Motivated by aircraft noise considerations, runway allocation systems were designed for Sydney and Brisbane airports [7]. More recently, several authors have considered the problem of optimally scheduling runway configurations, taking into account different models of weather forecasts and the loss of capacity during configuration switches [8, 9, 10, 11, 12].

Descriptive models analyze historical data in order to predict the runway configuration selected by the decision-makers. These models have received less attention, although data mining approaches have been used to forecast airport arrival rates, especially during Ground Delay Programs [13, 14]. A 24-hour forecast of runway configuration was developed for Amsterdam Schiphol airport, using a probabilistic weather forecast [15]. A logistic regression-based approach was used to develop a descriptive model of runway configuration selection at LaGuardia (LGA) and John F. Kennedy (JFK) airports, although this was not a predictive model [16]. Discrete-choice models of the runway configuration selection process have also been studied, and applied to LGA and Newark (EWR) airports [17, 18].

This paper extends the discrete-choice modeling approach [18], and applies it to LaGuardia (LGA) and San Francisco (SFO) airports. A key novelty in this paper is that the constraints pertaining to maximum tailwinds and crosswinds allowable for the use of a runway are learned from real data. The utility functions capture the importance of wind speed and direction, traffic demand, noise abatement procedures, the coordination of flows with neighboring airports, as well as

"inertia" (or the resistance to configuration changes). While the influence of inertia may be less important at long fore-cast horizons (when the key factors are likely to be wind conditions, visibility and demand), at shorter time-horizons such as up to 3 hours ahead, the resistance to configuration changes play a more important role. Without accounting for the inertia factor, tools that suggest runway configuration choices have been found to recommend significantly more frequent changes than was seen in actual operations [12]. The discrete-choice modeling framework helps accommodate the effect of inertia, in addition to the other influencing factors. This paper illustrates the proposed approach using case studies of LGA and SFO, first assuming a knowledge of the actual weather conditions and traffic demand 3 hours ahead, and then using the most recent Terminal Aerodrome Forecast (TAF) available 3 hours in advance. The results show that given the scheduled traffic demand, runway configuration being used and TAFs available 3 hours in advance, the configuration used during any 15-minute time-period can be predicted with an accuracy of 79-80% at both LGA and SFO.

B. Notation

Runway configurations are typically designated in the form 'A1, A2 | D1, D2' where A1 and A2 are the arrival runways, and D1 and D2 are the departure runways. The numbers for each active runway are reported based on their bearing from magnetic north (in degrees) divided by 10. Pairs of parallel runways are differentiated by 'R' and 'L'. The layouts of SFO and LGA are illustrated in Figs. 1 and 2, respectively.
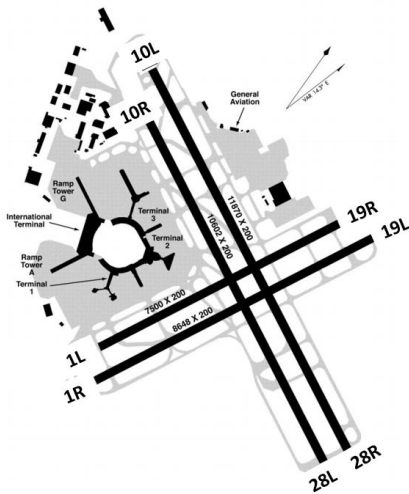

Figure 1: Layout of SFO airport.

Theoretically, an airport with N runways has $O(6^N)$ possible configurations, since each runway can be used for arrivals, departures or both, and in either direction. However, only 5-10 configurations are typically used at an airport. In addition, due to the additional coordination required during switches, runway configurations only change 1-3 times per day on average.

Table I shows the frequencies with which the most commonly-used configurations at SFO and LGA were ob-
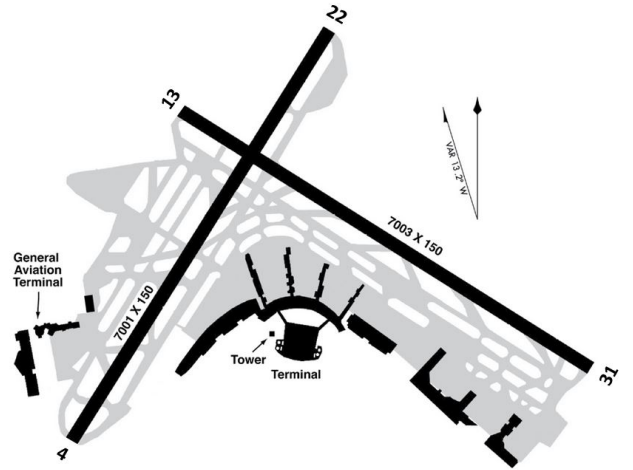

Figure 2: Layout of LGA Airport.

served in 2011. Hours between midnight and 6AM are not included.

Table I: Configurations observed at SFO and LGA in 2011.

Airport	Configuration	Frequency
SFO	28R,28L|1R,1L	18,952
	28R,28L|28R,28L	2,490
	28R/L|1R,1L	1,627
	19R,19L|10R,10L	752
LGA	31|4	6,772
	22|13	5,679
	22|31	4,488
	4|13	3,325
	31|31	1,483
	22,31|31	820
	4|4	813

II. METHODOLOGY

A. Discrete-choice modeling framework

Discrete-choice models are behavioral models that describe the choice selection of a decision maker, or the nominal decision selection among an exhaustive set of possible alternative options, called the choice set [19]. Each alternative in the choice set is assigned a utility function based on defining attributes that are related to the decision selection process. At any given time, the feasible alternative with the maximum utility is assumed to be selected by the decision maker.

The utility function is modeled as stochastic random variable, with an observed (deterministic) component, V, and a stochastic error component, ε. For the nth selection, given a set of feasible alternatives C_n, the utility of choice $c_i \in C_n$ is represented as

$$U_{n,i} = V_{n,i} + \varepsilon_{n,i}. \quad (1)$$

The decision maker selects the alternative with maximum utility, that is, $c_j \in C_n$ such that

$$j = \underset{i:c_i \in C_n}{\operatorname{argmax}}(U_{n,i}) \quad (2)$$

The observable component of the utility function is defined as a linear function of the observed vector of attributes, $\vec{X}_{n,i}$. The attributes include the different factors that can influence

the decision. They are weighted by the values in vector, $\vec{\beta}_{n,i}$, and include alternative specific constants, $\alpha_{n,i}$, as follows:

$$V_{n,i} = \alpha_{n,i} + [\vec{\beta}_{n,i} \cdot \vec{X}_{n,i}]. \tag{3}$$

The random error component of the utility function reflects all measurement errors, including unobserved attributes, variations between different decision-makers, proxy variable effects, and reporting errors. The error term is assumed to be distributed according to a Type I Extreme Value (or Gumbel) distribution with a location parameter of zero, that is:

$$f(x) = \mu e^{-\mu(x-\eta)} e^{-e^{\mu(x-\eta)}} \tag{4}$$

where μ is the scale parameter and η is the location parameter. The location parameter is set to zero when defining the discrete choice models. The Gumbel distribution is used to approximate a normal distribution due to its computational advantages. The Multinomial Logit (MNL) model assumes that the error components of each utility function are independent from one another, as shown in Fig. 3. Under the assumptions
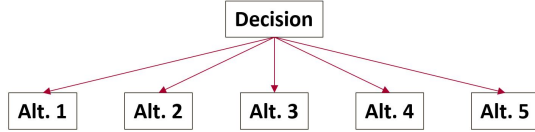

Figure 3: Schematic of a MNL model structure.

of the MNL model, the probability that choice i is chosen during the nth selection is given by

$$P_{n,i} = \frac{e^{V_{n,i}}}{\sum_{j:c_j \in C_n} [e^{V_{n,j}}]}. \tag{5}$$

The independence among the error terms of each utility function in the MNL model assumes that all correlation among alternatives has been captured by the attributes included in the utility function [19]. The Nested Logit (NL) model relaxes this assumption by grouping alternatives into subsets, or nests (denoted B_k), which have correlation between their error terms (Fig. 4).
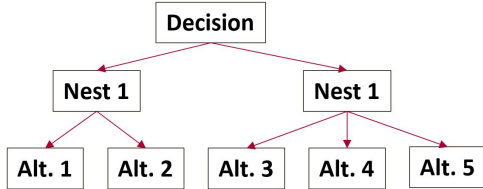

Figure 4: Example of a NL model structure.

The NL model splits the observable part of the utility function into a component that is common among the alternatives within a nest, and a component that varies between the different alternatives in a nest. The NL model can then be treated as nested MNL models using conditional probabilities. The probability that a specific alternative is chosen is given by the probability that its nest is chosen, multiplied by the

probability that the specific alternative is chosen from among the alternatives in that nest. In other words

$$P_n(c_i) = P_n(c_i|B_k)P_n(B_k), \tag{6}$$

$$\text{where} \quad P_n(B_k) = \frac{\exp(I_{n,k})}{\sum_{l=1}^{K}[\exp(I_{n,l})]} \tag{7}$$

$$P_n(c_i|B_k) = \frac{\exp(\mu_k V_{n,i})}{\sum_{j \in B_k}[\exp(\mu_k V_{n,j})]} \tag{8}$$

$$I_{n,k} = \frac{1}{\mu_k}\ln\left(\sum_{j \in B_k} \exp(\mu_k V_{n,i})\right). \tag{9}$$

Equation (7) has an additional term in the numerator called the inclusive value, that acts as a bridge between the lower level MNL models within each nest, and the upper level.

B. Maximum-likelihood estimation of model parameters

Maximum-likelihood estimates of the linear weighting parameters, alternative specific constants, and scale parameters are estimated from the training data. The maximum-likelihood function is defined as the joint probability that the vector of sample data will occur, given a vector of parameters $\vec{\theta} = <\alpha, \vec{\beta}, \mu>$ as follows,

$$\mathscr{L}(\vec{\theta}) = P(\vec{X}; \vec{\theta}). \tag{10}$$

The estimated parameters are those that maximize the likelihood of the observations:

$$(\widehat{\alpha}, \widehat{\vec{\beta}}, \widehat{\mu}) = \arg\max_{\alpha, \vec{\beta}, \mu}(\mathscr{L}(\alpha, \vec{\beta}, \mu)). \tag{11}$$

The resulting nonlinear optimization problem is solved computationally using an open-source software package called BIOGEME [20].

C. Statistical tests

The discrete-choice models for SFO and LGA were identified iteratively, and variables were added or removed based on their statistical significance. Different network structures were also tested and evaluated for statistical significance. The statistical significance of different attributes with respect to the training data were determined using the t-test. The significance of their effects on the overall model were tested using a likelihood-ratio test. Likelihood-ratio tests were also used to evaluate the effect of adding nests to NL models [19].

III. APPLICATION TO RUNWAY CONFIGURATION SELECTION AT LGA AND SFO

A. Training data

The training and test datasets were obtained from the Aviation System Performance Metrics (ASPM) database [21]. The data included the active runway configuration, the arrival and departure demand, ceiling and visibility conditions, and wind speed and direction, for each 15-minute interval. The training datasets used to determine the parameters of the LGA and SFO models were taken from year 2011.

The ASPM dataset gives both the wind direction, θ and wind speed, v, for each 15-minute interval. Fig. 5 illustrates

that the headwind and crosswind components are given by $v\cos(\varphi - \theta)$ and $v\sin(\varphi - \theta)$, respectively, where φ denotes the orientation of the runway. Tailwinds occur when the headwind function takes a negative value.
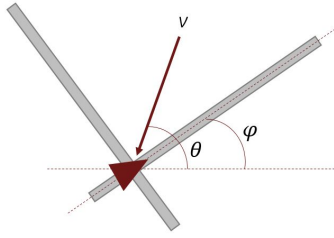

Figure 5: Determination of the headwind and crosswind components.

B. Attributes of the utility functions

The utility function is assumed to be a linear function of the observed vector of attributes, or factors that can influence the decision. For the runway configuration selection problem, the following attributes were considered.

- **Inertia:** The inertia variable reflects the preference of air traffic controllers to stay in the same configuration, since configuration changes require increased coordination among the different stakeholders, and reduce airport throughput [11]. The inertia variable is expected to have a positive impact on the utility function of the incumbent configuration.

- **Wind speed and direction:** Wind speed and direction are key factors that influence runway configuration. High tailwind or crosswind speeds are not favorable for operations, and can render some runways unusable. While prior work had used FAA-guidelines to determine the maximum allowable tailwind and crosswind components [18], this paper learns these threshold values from actual data (ASPM, 2011).

Figs. 6 and 7 show the identified ranges of feasible wind speed and direction for each runway at SFO and LGA, respectively. The tailwind and crosswind limits were taken as their 95th percentile threshold values in order to account for any reporting errors. While there is no limit set on headwind speeds for feasibility, the figures use a maximum headwind speed of 40 knots for illustrative purposes. The dashed lines show the 75th percentile of the head/tailwinds and crosswinds.

Even with the maximum tailwind and crosswind limitations, several runway configurations may be feasible at any time. Fig. 8 shows the actual observations of wind speed and direction at SFO, for each 15-minute interval in 2011. We see that the significant majority of points correspond to conditions in which all the runways are feasible (although some may be more favorable than others). Fig. 9 shows the same data for LGA in 2011.

Headwinds are expected to add a positive contribution to the runway configuration utility functions, while tailwinds are expected to add a negative contribution.
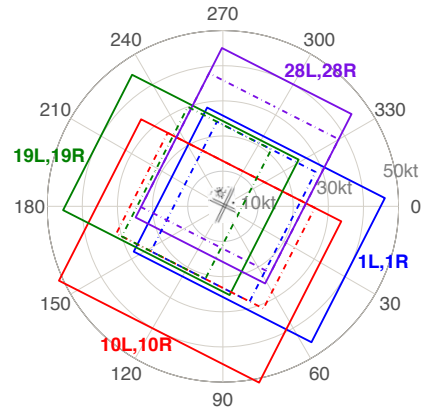

Figure 6: Wind speed and direction thresholds for different runway configurations at SFO, as estimated from 2011 data. The solid lines correspond to the 95th percentile, while the dashed lines correspond to the 75th percentile.
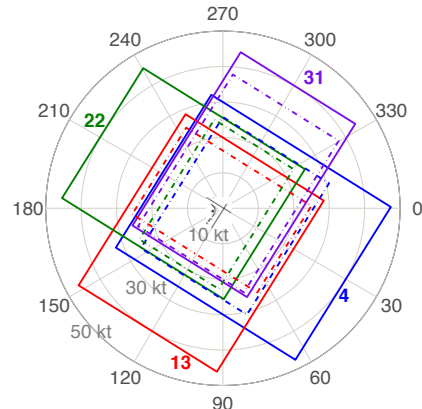

Figure 7: Wind speed and direction thresholds for different runway configurations at LGA, as estimated from 2011 data. The solid lines correspond to the 95th percentile, while the dashed lines correspond to the 75th percentile.

However, significant head winds may have an adverse effect on operations by decreasing spacing between pairs of aircraft on final approach, a phenomenon known as *compression* [22]. In order to account for this possibility, headwinds speeds in the top quartile (that is, above the 75th percentile) were treated as "high" headwinds. Compression was found not to have a statistically significant impact at SFO. However, as will be seen in Section III-C,
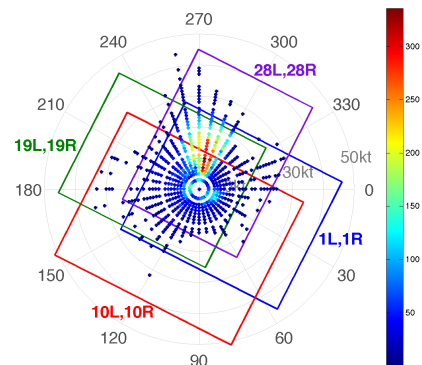

Figure 8: Wind speed and direction observed at SFO in 2011. The solid lines correspond to the maximum thresholds. The colors indicate the frequency of occurrence of different points.
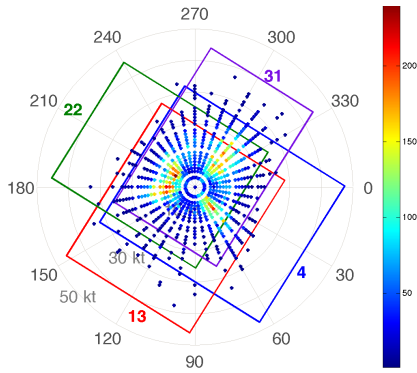
Figure 9: Wind speed and direction observed at LGA in 2011. The solid lines correspond to the maximum thresholds. The colors indicate the frequency of occurrence of different points.

- the weighting factor of the headwind attribute of the utility functions does decrease for headwind speeds in the top quartile at LGA.
- **Demand:** Airport arrival and departure demand play a significant role when picking a runway configuration. Specifically, in high demand situations, high capacity configurations are preferred. High capacity configurations usually include an extra arrival or departure runway.
- **Ceiling and visibility:** Meteorological conditions, as represented by the visibility and cloud ceiling, are an important consideration in the selection of runway configuration. Visual Meteorological Conditions (VMC) refer to times when the visibility is sufficient for pilots to maintain visual separation from the ground and other aircraft, while Instrument Meteorological Conditions (IMC) refer to times when pilots are required to primarily use their flight instruments. IMC is defined by a visibility less than 3 miles and a ceiling less than 1,000 ft [23].

   The LGA model incorporates variables for each utility function corresponding to visual and instrument conditions. In this manner, these variables will capture any preferences for one configuration over another under VMC or IMC.

   The optimal capacity configuration at SFO is 28R,28L|1R,1L, which involves simultaneous arrivals on closely-spaced parallel runways. Since the centerlines of 28R and 28L are only 750 ft apart, simultaneous landings are not possible in IMC [23]. One would therefore expect that the 28R/L|1R,1L configuration (which involves using only one of the two runways for arrivals) would be favored under IMC. In short, the VMC/IMC attribute is expected to have a very strong influence on the utility function of different runway configurations at SFO.

Table II: OCCURRENCES OF 28R,28L|1R,1L AND 28R/L|1R,1L AT SFO IN 2011.

Configuration	VMC	IMC
28R,28L|1R,1L	19,832	4,161
28R/L|1R,1L	3,466	1,844

Table II shows the relative use of each of these configurations under VMC and IMC. As shown, the single arrival runway configuration is used a greater fraction of the time

in IMC than in VMC. However, runway configuration 28R,28L|1R,1L is still used the majority of the time during IMC. During IMC periods, simultaneous (side-by) landings are not possible, and the airport operates almost as it would in a single arrival runway configuration. However, there may be a small increase in the airport capacity when 28R,28L|1R,1L is used under IMC. In order to evaluate these effects, new variables that combine the effects of visibility and demand were used in the SFO model. Four categorical variables were defined for periods of:

1) IMC + low demand,
2) IMC + high demand,
3) VMC + low demand,
4) VMC + high demand,

where a low (high) demand period was defined by an arrival demand less than (greater than) 5 flights per 15-minute interval.

- **Noise abatement procedures:** Noise abatement procedures are used at most major airports to reduce the impacts of noise on neighboring communities, especially during early morning and nighttime periods. At SFO, configurations that arrive and depart over the water are preferred over configurations that arrive and depart over the suburbs. At LGA, configurations with flight paths over the city and away from suburban areas are preferred during the nighttime hours. Variables were included in both the SFO and LGA models to account for these effects.
- **Coordination with neighboring airports:** In multi-airport terminal-areas, flows into and out of the different airports need to be coordinated. Therefore, in the LGA model, variables were added to test whether runway configuration decisions at LGA were affected by the runway configurations at JFK, at any given time.
- **Switch proximity:** Despite the resistance to configuration changes, changes in the attributes described above can necessitate a configuration switch. Certain configuration switches require more coordination relative to others. For instance, the addition of an arrival runway may be easier to implement than a change in the direction of operations. To account for these effects, variables were added that weight each utility function differently depending on the runway configuration in the previous time-step. These variables are fundamentally the same as the inertia variable, but operate on the utility functions of different alternatives instead of the utility function of the same runway configuration.

C. Estimated discrete-choice models

1) SFO model: The training data set used to estimate the parameters for the SFO model consisted of 24,820 decision selection periods. Runway configurations that were utilized more than 1

The estimated weighting parameters for the utility function (β values from (3)) are reported in Table III. The table
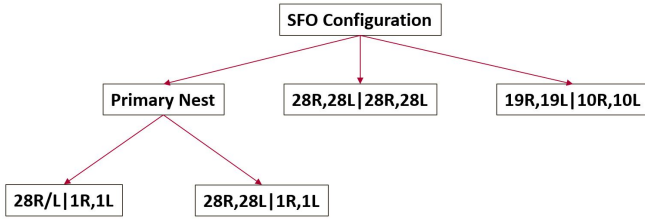
Figure 10: SFO model specification.

Tree: SFO Configuration → Primary Nest, 28R,28L|28R,28L, 19R,19L|10R,10L. Primary Nest → 28R/L|1R,1L, 28R,28L|1R,1L.

Fig. 11, the chosen model had a nested logit structure with a single nest containing all alternatives using runway 22 for arrivals. The scale parameter for the nest was $\mu_{ARR22} = 1.1$. The estimated values of the of the weights on different attributes (i.e., $\vec{\beta}$), their standard errors and t-statistics are presented in Table IV.
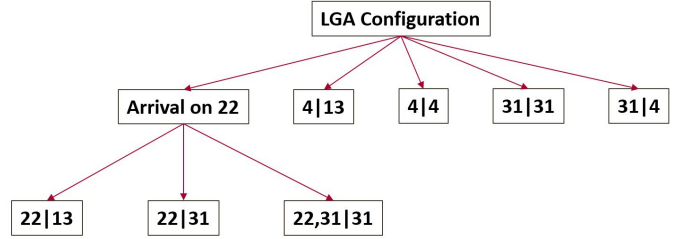
Figure 11: LGA model specification.

Tree: LGA Configuration → Arrival on 22, 4|13, 4|4, 31|31, 31|4. Arrival on 22 → 22|13, 22|31, 22,31|31.

includes the estimated weight of each attribute in the utility function, the standard error of the estimate, and its t-statistic. Parameters that did not have a statistically significant t-statistic (absolute value less than 1.96) were removed from the model, except in cases where removing these variables could bias the predictions.

Table III: ESTIMATED UTILITY FUNCTION WEIGHTS FOR SFO.

Parameters	Value	Std. error	t-statistic
Inertia parameters			
28R,28L|1R,1L	4.48	0.139	31.87
19R,19L|10R,10L	3.20	0.376	8.64
28R,28L|28R,28L	4.35	0.209	20.82
28R/L|1R,1L	4.48	0.139	31.87
Wind parameters			
Headwind on arrival Runway	0.0415	0.0131	3.2
Tailwind on arrival Runway	-0.0415	0.0131	-3.2
Headwind on departure Runway	0.0608	0.0076	8.14
Tailwind on departure Runway	-0.0608	0.0076	-8.14
Noise parameters			
Depart Runway 28 during evening	-0.356	0.176	-3.23
Arrive Runway 10 during evening	-0.356	0.176	-3.23
Switch proximity parameters			
28R,28L|1R,1L to 28R,28L|28R,28L	-0.775	0.649	-1.22
28R/L|1R,1L to 28R,28L|28R,28L	-0.775	0.649	-1.22
19R,19L|10R,10L to 28R,28L|1R,1L	-0.775	0.649	-1.22
19R,19L|10R,10L to 28R/L|1R,1L	-0.775	0.649	-1.22
19R,19L|10R,10L to 28R,28L|28R,28L	-0.946	0.601	-1.57
28R,28L|28R,28L to 28R,28L|1R,1L	1.33	0.251	4.97
28R,28L|28R,28L to 28R/L|1R,1L	1.33	0.251	4.97
Demand/visibility parameters			
VMC + high demand; 28R/L|1R,1L	-1.66	0.39	-3.44
IMC + low demand; 28R/L|1R,1L	0.327	0.381	0.93

The estimated values of $\vec{\beta}$ shown in Table III indicate that the inertia parameters are the most important factor in the SFO runway configuration selection process. Wind also has a heavy influence on the canidate utility functions, espscially during unfavorable weather conditions. Low pressure systems typically have high circulating winds that disrupt the typical high wind from the San Bruno Gap (which corresponds to headwinds on runway 28). When this occurs, air traffic controllers are highly inclined to change the configuration [24]. The noise parameter favors configurations in which arrivals and departures fly over the Bay (instead of over the suburbs) during the evening and morning hours. As expected, the single arrival runway configuration, 28R/L|1R,1L, gets a positive contribution to its utility function during periods of low demand and IMC, and a negative contribution to its utility function during periods of high demand and VMC.

2) LGA model: The training data set used to estimate the parameters for the LGA model consisted of 26,203 decision selection periods. Again, only runway configurations that were utilized more than 1% during the year were considered, resulting in 7 different runway configuration alternatives. Shown in

Table IV: Estimated utility function weights for LGA.

Parameters	Value	Std. error	t-statistic
Inertia parameters			
Config. 22|13	4.58	0.187	24.5
Config. 22|31	7.41	0.36	20.57
Config. 22,31|31	7.41	0.36	20.57
Config. 31|31	4.91	0.401	12.24
Config. 31|4	3.16	0.25	12.6
Config. 4|13	3.99	0.196	20.34
Config. 4|4	5.44	0.416	13.1
Wind parameters			
High headwind on arrival runway	0.0952	0.0161	5.89
Normal headwind on arrival runway	0.123	0.0197	6.26
Tailwind on arrival runway	-0.0946	0.0199	-4.74
Tailwind on departure runway	-0.211	0.0173	-12.2
Tailwind on extra arrival runway	-0.348	0.07	-4.97
Demand parameters			
Arrival demand; 31|31	-0.101	0.0312	-3.24
Arrival demand; 4|4	-0.0807	0.0327	-2.47
VMC/IMC parameters			
VMC on 31|31	2.09	0.402	5.19
VMC on 31|4	1.36	0.231	5.9
Switch proximity parameters			
31|4 to 31|31	-1.4	0.463	-3.03
4|13 to 31|31	-2.52	0.714	-3.53
4|4 to 31|31	-1.32	0.747	-1.77
22|13 to 31|31	-1.99	0.577	-3.45
4|13 to 31|4	-2.19	0.368	-5.94
4|4 to 31|4	-1.05	0.515	-2.04
22|13 to 31|4	-2.14	0.355	-6.04
4|13 to 4|4	-1.6	0.443	-3.61
22|13 to 4|4	-1.92	0.532	-3.6
31|31 to 22|13	-1.05	0.573	-1.84
Inter-airport coordination parameters			
JFK arr. on 13; LGA arr. on 22/ dep. on 4	0.85	0.308	2.76
JFK arr. on 13; LGA arr. on 31/ dep. on 13	1.27	0.464	2.75
JFK dep. on 13; LGA dep. on 31	-1.99	0.224	-8.88
JFK arr. on 13; LGA arr. on 4	-0.448	0.172	-2.6
JFK arr. on 13; LGA dep. on 31	-1.61	0.222	-7.26
JFK arr. on 13; LGA arr. on 31/dep. on 13	0.796	0.25	3.19
JFK dep. on 4; LGA dep. on 31	-2.5	0.341	-7.34
JFK arr. on 4; LGA arr. on 22/ dep. on 4	-0.737	0.293	-2.51
JFK dep. on 13; LGA arr. on 4	-1.15	0.312	-3.68

The estimated values of the weighting parameters once again indicate the importance of the inertia variables in the runway configuration selection process at LGA. The results show that the headwind parameters are statistically significant for the primary arrival runway, but not for the primary departure runway or the extra arrival runway. This finding suggests that the alignment of the primary arrival runway is more important than the departure or extra arrival runways. Furthermore, the influence of tailwinds was found to be statis-

tically significant for all runway configurations. Compression is found to have a small, but statistically significant effect on the runway configuration choice. The weighting parameter corresponding to high headwinds (in the top quartile) is found to be smaller than the one for normal headwinds, indicating that high headwinds (that can cause compression) are less preferable.

IV. PREDICTION OF RUNWAY CONFIGURATION

A. 3-hour forecast, assuming actual weather and demand

The estimated utility functions and the discrete-choice model can be used in (6)-(9) to determine the probability of choosing a runway configuration alternative during each 15-minute selection period, given the values of the attributes for that time-period. The attributes considered include the runway configuration in the previous time-period, the wind and weather conditions in the time-period being considered, as well as the demand in that time-period. In order to predict a runway configuration 3-hours in the future, one needs to determine the probability of choosing a particular configuration, considering not just the demand and weather in that time-period, but also all possible evolutions of the runway configuration in the next three hours. The runway configuration alternative with the maximum probability of being chosen is selected as the predicted runway configuration.

Bayes' rule can be recursively applied for each 15-minute interval to determine the probabilities of selecting a runway configuration at a 3-hour time horizon. For example, considering the attributes of the current time at SFO, the runway configuration alternatives 28R,28L|1R,1L, 19R,19L|10R,10L, 28R,28L|28R,28L, and 28R/L|1R,1L may have selection probabilities of 5%, 75%, 5%, and 15% respectively, for the next 15-minute interval. The probabilities of being in each configuration 30-min (i.e., 2 time-periods) from now will have to be conditioned on the runway configuration in the next time-period, and so on. In this manner, the probabilities of being in 28R,28L|1R,1L, 19R,19L|10R,10L, 28R,28L|28R,28L, or 28R/L|1R,1L 3-hours from now may have changed to 7%, 50%, 3%, and 40% respectively. Runway configuration 19R,19L|10R,10L is then taken as the 3-hour prediction because it has the highest probability.

The accuracy of the predictions are first evaluated assuming a perfect knowledge of the wind, visibility, and demand variables for the subsequent 3-hours. Accuracy is defined as the percentage of correct predictions for that configuration. In other words, if a runway configuration is predicted correctly for every 15-minute interval in which it is observed, the accuracy for that configuration would be 100%. In the previous example, if the actual configuration used 3-hours from now were 19R,19L|10R,10L, it would correspond to a correct prediction. The accuracy for configuration 19R,19L|10R,10L is given by the number of times it was correctly predicted, divided by the number of times that the configuration was observed in the year.

The test data for this validation exercise was the ASPM database for the year 2012. We note that the test set from year 2012 is an independent dataset from the training set from year 2011. The inputs included the actual arrival and departure demand, ceiling and visibility conditions, and wind speed and direction, for each 15-minute interval. The output was a prediction of the runway configuration 3-hours in the future. The results of the predictions, both for the next 15 minutes and 3 hours in the future, are shown in Table V for SFO and in Table VI for LGA. The tables show that the prediction accuracy at the 15-minute horizon is nearly 98% at both airports, given the actual weather and demand conditions.

Table V: PREDICTION ACCURACY (USING ACTUAL WEATHER AND DEMAND) FOR SFO IN 2012, FOR 15-MINUTE AND 3-HOUR PREDICTION HORIZONS.

Configuration	Frequency	Prediction accuracy	
		15 min	3 hr
28R,28L|1R,1L	21,582	99.1%	93.6%
28R/L|1R,1L	5,060	93.4%	63.7%
28R,28L|28R,28L	2,450	95.9%	58.0%
19R,19L|10R,10L	1,133	99.4%	76.3%
Total	**25,175**	**97.9%**	**85.1%**

A major challenge at SFO is accurately predicting between the configuration 28R,28L|1R,1L with arrivals on the closely-spaced parallel runways, and the single arrival runway configuration, 28R/L|1R,1L. As mentioned before, simultaneous (side-by) landings are not possible under IMC, and the airport operates almost as it would in a single arrival runway configuration, even in 28R,28L|1R,1L. The reported configurations in the ASPM data set do not differentiate between simultaneous and staggered parallel approaches, even though the latter has a capacity that would be close to 28R/L|1R,1L. This fact, along with the other similarities between these two runway configuration alternatives, makes it difficult to predict either of these alternatives accurately without introducing a selection bias. Accurately predicting 28R,28L|28R,28L is also challenging due to the limitations from the ASPM dataset. This runway configuration is typically only used for long-haul departures over the Pacific Ocean and to Hawaii, and the aggregate flight counts in ASPM are not sufficient to account for this factor. Despite these challenges, the overall accuracy for the 3-hour prediction of runway configuration at SFO is 80%, assuming perfect knowledge of future wind and weather conditions, and traffic demand.

Table VI: PREDICTION ACCURACY (USING ACTUAL WEATHER AND DEMAND) FOR LGA IN 2012, FOR 15-MINUTE AND 3-HOUR PREDICTION HORIZONS.

Configuration	Frequency	Prediction accuracy	
		15 min	3 hr
22|13	8,220	98.3%	89.0%
31|4	6,454	98.0%	84.8%
4|13	4,851	98.0%	83.0%
22|31	2,938	97.3%	71.6%
31|31	2,136	96.5%	67.0%
22,31|31	1,838	96.8%	67.2%
4|4	795	96.5%	68.2%
Total	**27,232**	**97.8%**	**82.2%**

Even with these challenges, the accuracy of the predictive model for SFO is promising. The overall accuracy for predicting the active runway configuration on a three hour time horizon reached upwards of 80% assuming perfect knowledge of the future weather and demand. The overall 3-hour

prediction accuracy for LGA in 2012 was 82%, given the future wind, weather, and airport demand values. As with SFO, configurations that were seen more often than others exhibited a higher relative prediction accuracy. For comparison, prior research (with no look-ahead, and using logistic regression models) achieved a prediction accuracy of 75% at LGA and 63% at JFK, respectively [16].

B. 3-hour forecast using weather forecasts and scheduled demand

The discrete-choice models developed in this paper were combined with Terminal Aerodrome Forecast (TAF) data, in order to predict future runway configurations. This would correspond to a practical implementation, where the actual values of the attributes are not available, and instead weather forecasts and schedules have to be used. Table VII shows the 3-hour prediction performance for SFO in 2012 using TAF data (and scheduled demand), while Table VIII shows same results for LGA in 2012.

Table VII: PREDICTION ACCURACY AT THE 3 HR AND 6 HR FORECAST HORIZONS (USING TAF AND SCHEDULED DEMAND), FOR SFO IN 2012.

Configuration	Frequency	Prediction accuracy	
		3 hr	6 hr
28R,28L|1R,1L	22,173	94.3%	95.9%
28R/L|1R,1L	4,700	57.1%	39.5%
28R,28L|28R,28L	2,875	44.0%	8.5%
19R,19L|10R,10L	1120	30.3%	88.6%
Total	**30,868**	**81.8%**	**76.5%**

Table VIII: PREDICTION ACCURACY AT THE 3 HR AND 6 HR FORECAST HORIZONS (USING TAF AND SCHEDULED DEMAND), FOR LGA IN 2012.

Configuration	Frequency	Prediction accuracy	
		3 hr	6 hr
22|13	7,677	85.1%	78.1%
31|4	6,541	84.8%	81.2%
4|13	4,665	75.8%	58.6%
22|31	2,678	74.8%	57.6%
31|31	2,451	70.0%	39.2%
22,31|31	1,707	72.2%	52.2%
4|4	885	65.3%	38.9%
Total	**26,604**	**79.5%**	**66.8%**

The overall accuracy using TAF forecast data from 2012 was 80.4% for SFO, and 79% for LGA. It is promising that the accuracy of the model is not substantially degraded by using forecast data, which is inherently prone to error. As before, the accuracy of specific configurations increased as the frequency of their occurrence increased. These results are very promising, considering that logistic regression based models with no lookahead have been found to achieve a prediction accuracy of 75% at LGA [16].

C. Comparison with baseline heuristic

Since runway configuration changes only occur a handful of times in a day, a baseline heuristic would be to assume that the airport remains in the current configuration. A comparison of the discrete-choice model and such a constant baseline is shown in Figure 12, for increasing forecast horizons. We see that at short horizons, when the airport generally stays in the same configuration, the accuracy of the two models are comparative. However, as the forecast horizon increases, the

difference in the two models increases: At a 3-hour forecast horizon, there is a 5 percentage point improvement in the performance of the discrete-choice model, and at a 6-hour forecast horizon, the discrete-choice model outperforms the baseline by more than 10 percentage points.
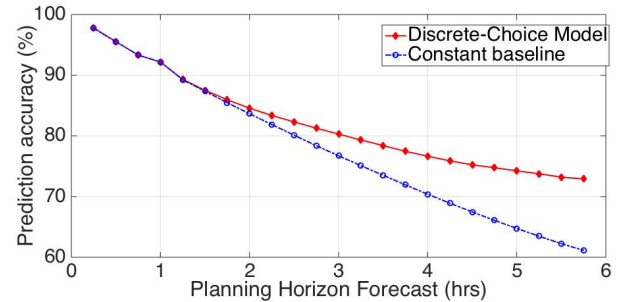

Figure 12: Prediction accuracy of discrete-choice model and a constant baseline model as functions of the forecast horizon, for LGA airport in 2012.

D. Limitations of approach and potential extensions

Discrete-choice models are inherently data-driven, and therefore only predict configurations that have been observed before. The models need to be re-estimated when there are major changes in the decision process, such as capacity enhancements or new procedures [25]. Similarly, configurations that are infrequently used are difficult to predict.

Also, discrete-choice models do not account for the variability among decision-makers, who may have varying levels of experience, diverse concerns, and different rationales for selecting a runway configuration. Only "nominal" behavior is captured by these models. The model assumes the presence of rational decision-makers who share a utility function that reflects nominal system behavior.

Bias can be another limitation when modeling the runway configuration selection process with a discrete choice model. Any errors or biases produced in the 15-minute selection probabilities are magnified when forecasting on a 3-hour time horizon. As a result, the inertia parameter frequently causes the utility function of the incumbent configuration to overpower the utility functions of the other alternatives. When the weather conditions do not heavily favor a configuration switch, these models will tend to predict that the configuration will remain the same. As a result, the prediction accuracy is reduced in time-periods close to a configuration switch.

Despite these limitations, the prediction performance of the proposed discrete-choice models suggests that they are a promising approach to predict runway configuration a few hours ahead of time. The inertia term could be further improved by limiting its effect as time progresses within a 3-hour prediction period. Biases within the estimated parameters could be reduced by estimating the utility parameters using a balanced data set. The effect of wind gusts, which are currently ignored, can also be easily included in the utility models.

V. CONCLUSIONS

This paper presented a framework to build a discrete-choice model of the configuration selection process using empirical

observations. The proposed model included a utility function that reflected the importance of various factors, including weather, wind speed and direction, arrival and departure demand, noise mitigation procedures, and coordination with neighboring airports. The resulting model considers the current runway configuration, weather and traffic demand, and provides a probabilistic forecast of the runway configuration at the next 15-minute interval. The predictive model was then extended to obtain the 3-hour probabilistic forecast of configuration.

Case studies based on data from LaGuardia (LGA) and San Francisco (SFO) airports were used to evaluate the model. Given a perfect knowledge of weather and demand 3-hours in advance, the model was shown to predict the correct runway configuration at SFO with 81% accuracy, and LGA with 82% accuracy. For a practical implementation, weather forecasts in the form of TAF were used to predict the weather conditions 3-hours in advance. The results showed that even using weather forecasts and scheduled demand, the model can predict the correct runway configuration 3 hours in advance with an accuracy of 79% for LGA and 80% for SFO.

REFERENCES

[1] E. Gilbo, "Airport capacity: representation, estimation, optimization," *IEEE Transactions on Control System Technology*, vol. 1, no. 3, pp. 144–153, 1993.

[2] D. Bertsimas and S. Stock Patterson, "The air traffic flow management problem with enroute capacities," *Operations Research*, vol. 46, no. 3, pp. 406–422, May-June 1998.

[3] G. Lulli and A. Odoni, "The European Air Traffic Flow Management Problem," *Transportation Science*, vol. 41, no. 4, pp. 431–443, November 2007.

[4] Y. Jung, T. Hoang, J. Montoya, G. Gupta, W. Malik, L. Tobias, and H. Wang, "Performance evaluation of a surface traffic management tool for Dallas/Fort Worth International Airport," in *9th USA/Europe Air Traffic Management R&D Seminar*, 2011.

[5] J. Post, J. Gulding, K. Noonan, D. Murphy, J. Bonn, and M. Graham, "The modernized National Airspace System Performance Analysis Capability (NASPAC)," in *26th International Congress of the Aeronautical Sciences (ICAS)*, 2008.

[6] Flight Transportation Associates, "Enhanced Preferential Runway Advisory System (ENPRAS)." [Online]. Available: http://www.ftausa.com/enpras.htm

[7] D. G. Southgate and S. J. Sedgwick, "Time stamped aircraft noise prediction replacing the 'average day with the 'composite year," in *INTER-NOISE*, December 2006.

[8] L. Li and J.-P. Clarke, "A stochastic model of runway configuration planning," in *AIAA Guidance, Navigation and Control Conference*, August 2010.

[9] C. A. Provan and S. C. Atkins, "Optimization models for strategic runway configuration management under weather uncertainty," in *AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, 2010.

[10] C. Weld, M. Duarte, and R. Kincaid, "A runway configuration management model with marginally decreasing transition capacities," *Advances in Operations Research*, 2010.

[11] M. J. Frankovich, D. Bertsimas, and A. R. Odoni, "Optimal selection of airport runway configurations," *Operations Research*, vol. 59, no. 6, pp. 1407–1419, November-December 2011.

[12] R. Oseguera-Lohr, N. Phojanamongkolkij, G. Lohr, and J. W. Fenbert, "Benefits assessment for tactical runway configuration management tool," in *AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, 2013.

[13] P.-C. B. Liu, M. Hansen, and A. Mukherjee, "Scenario-based air traffic flow management: From theory to practice," *Transportation Research Part B*, vol. 42, pp. 685–702, 2008.

[14] G. Buxi and M. Hansen, "Generating probabilistic capacity profiles from weather forecast: A design-of-experiment approach," in *Ninth USA/Europe Air Traffic Management Research and Development Seminar (ATM2011)*, 2011.

[15] H. Hesselink and J. Nibourg, "Probabilistic 2-day forecast of runway use: Efficient and safe runway allocation based on weather forecast," in *Ninth USA/Europe Air Traffic Management Research and Development Seminar (ATM2011)*, 2011.

[16] S. Houston and D. Murphy, "Predicting runway configurations at airports," in *Transportation Research Board (TRB) Annual Meeting*, 2012, paper number 12-3682.

[17] V. Ramanujam and H. Balakrishnan, "Estimation of maximum-likelihood discrete-choice models of the runway configuration selection process," in *American Control Conference*, 2011.

[18] ——, "Data-driven modeling of the airport configuration selection process," *IEEE Transactions on Human-Machine Systems*, 2014, under review.

[19] M. Ben-Akiva and S. Lerman, *Discrete choice analysis: theory and application to travel demand*. MIT Press, 1985.

[20] M. Bierlaire, "BIOGEME: a free package for the estimation of discrete choice models," in *3rd Swiss Transport Research Conference*, 2003.

[21] Federal Aviation Administration, "Aviation System Performance Metrics (ASPM) database." [Online]. Available: aspm.faa.gov

[22] R. A. DeLaura, R. F. Ferris, F. M. Robasky, S. W. Troxel, and N. K. Underhill, "Initial assessment of wind forecasts for Airport Acceptance Rate (AAR) and Ground Delay Program (GDP) planning," Lincoln Laboratory, Tech. Rep. ATC-414, January 2014.

[23] Federal Aviation Administration, "Airport capacity benchmark report," 2004.

[24] The Boeing Company, "Airport noise and emissions regulations." [Online]. Available: www.boeing.com/commercial/noise/san_francisco.html

[25] Federal Aviation Administration, "FAA JO 7110.308 CHG 3: 1.5-Nautical Mile dependent approaches to parallel runways spaced less than 2,500 feet apart," 2004.

AUTHOR BIOGRAPHIES

Jacob Avery is an SM student in the Department of Aeronautics and Astronautics at MIT. He received a Bachelor of Science in Mechanical Engineering with highest honors at the University of Illinois at Urbana-Champaign. His research is in the design and analysis of algorithms to accurately model the runway configuration selection process. He is the recipient of a Graduate Research Award from the Transportation Research Board's Airport Cooperative Research Program (ACRP).

Hamsa Balakrishnan is an Associate Professor of Aeronautics and Astronautics at the Massachusetts Institute of Technology. Her research is in the design, analysis, and implementation of control and optimization algorithms for large-scale cyber-physical infrastructures, with an emphasis on air transportation systems. She received her PhD in Aeronautics and Astronautics from Stanford University. She is a recipient of the Donald P. Eckman Award from the American Automatic Control Council (2014), the Lawrence Sperry Award from the American Institute of Aeronautics and Astronautics (2012), the CNA Award for Operational Analysis (2012), the Best Paper Award at the FAA/Eurocontrol Air Traffic Management R&D Seminar (2011), and an NSF CAREER Award (2008).