

Validation of Runway Capacity Models

Amy Kim and Mark Hansen

Department of Civil and Environmental Engineering
University of California, Berkeley
Berkeley, USA
amy_kim@berkeley.edu
mhansen@ce.berkeley.edu

Abstract— There are many runway capacity estimation models currently available today, and developers usually claim that their models have been validated. However, information about the validation process is often limited, and different models are validated at different levels of complexity. As a result, this paper proposes two validation methodologies that can be used to test model predictions against reality. We demonstrate the methods on two models--the Airfield Capacity Model (ACM) and Runway Simulator (rS)--and two airports—SFO and LAX. The results indicate that both models tend to over-predict capacities under good visibility conditions, and predict wider ranges of capacities than are seen empirically. Overall, capacity estimates from rS are typically more accurate than those from ACM.

Keywords--runway capacity; empirical estimation; capacity models; ACM; rS; validation; censored regression.

I. INTRODUCTION

There are many runway capacity estimation models commercially available and in use today. These models span a wide range of types, scope, and capabilities. Model developers usually claim that their models have been validated, but there are several issues that arise with these validation claims. Firstly, information on the calibration and validation processes used is often vague or unclear. Secondly, model validations are performed at differing levels of complexity. Finally, validation exercises were often carried out by the developers themselves.

In this paper, we propose and demonstrate two possible validation methods that can be used to compare the estimates from a runway capacity model against empirical counts of the number of operations. These methods account for the fact that capacity may not be directly observable, since it represents an upper limit rather than the actual number of operations. They are demonstrated on two models-- the Airfield Capacity Model (ACM) and Runway Simulator (rS). This paper will provide a description of the models and the validation methodology, describe the data used, present validation results and suggested directions for future work.

II. BACKGROUND

There are many phenomenon that can affect the number of aircraft able to land and depart at an airport. Those that commonly have the greatest effects are below [1]:

- Weather (visibility, cloud ceiling, precipitation), and subsequent meteorological condition designation;

- Air traffic control separation requirements;
- State and performance of ATM system;
- Number of runways in use and their geometric layout;
- Aircraft fleet mix and performance;
- Runway occupancy times;
- Overall arrival/departure split;
- Mix and sequencing of arrivals and departures on runways, and
- Controller workload.

Depending on their purpose, capacity models attempt to account for sets of the above factors. All factors, excluding the last, are typically included in a model. Controller workload is a subjective measure and therefore more difficult to account for.

III. DESCRIPTION OF CAPACITY MODELS

A. Overview

The runway capacity models commercially available today span a wide range of scopes, capabilities, and complexities [2]. Models can be categorized in several ways; here we categorize them by three important aspects: calculation method, stochastic capability, and model scope. The first two are independent of one another; however, they serve to isolate key differences between models.

Runway capacity models calculate capacity analytically or through simulation. Analytical models are mathematical representations of operations, and can be implemented using a calculator or spreadsheet. They rely on a set of key capacity-affecting inputs and variables to quickly, simply, and efficiently estimate the average behavior of entities (in this case, aircraft). Simulation models attempt to characterize changing conditions over time. They can be further categorized as macroscopic, mesoscopic, and microscopic. Macroscopic models are like analytical models in that they rely on key variables to represent the average behavior of entities. However, they are updated with changing information in discrete time steps. In microscopic simulation, aircraft (for instance) are represented individually, and the model creates and records their interactions with one another and their environment. Microscopic models tend to be more

IV. DATA

comprehensive in accounting for capacity-affecting factors. Mesoscopic models combine elements of both macro and microscopic models. All the above models can be placed on a sliding scale of computational complexity, from the simple (analytical) to the highly complex (microscopic simulation).

Models can be deterministic, or stochastic to varying degrees. The degree of stochasticity depends on how many parameters (that do vary in actual operations) are treated as random variables in a given model.

Lastly, runway capacity models' scopes can range from being able to represent runway operations only to aircraft operations at gates, on aprons, taxiways, and in airspace. Very sophisticated models can incorporate numerous complex factors and operations that affect capacity, even beyond those listed in the previous section.

B. Airfield Capacity Model (ACM)

The ACM was initially developed by a consortium in the late 1970s and then modified by the FAA and MITRE CAASD, with the last modification made in 1981. It is an analytic model that calculates the hourly capacity of runway systems given continuous demand [3]. It asks the user for basic information about capacity-affecting characteristics, which it then converts to numerical inputs for its calculations. The ACM can estimate capacities for 15 simple runway configurations, from a single runway to 4 runways in varying configurations. The model's default assumption is that there is a 5% probability of violating separation standards, and this is used to determine runway aircraft spacing.

The ACM was validated in the early 1980s by the FAA; capacity estimates for certain runway configurations were deemed to be reasonably accurate. More information on the validation work is difficult to obtain. It is mainly used by the FAA and their consultants [4].

C. Runway Simulator (rS)

rS was developed by MITRE CAASD, and is an intermediate effort between a simple analytical model and a complex discrete event simulation model. rS simulates individual aircraft movements on runways and airspace in the immediate vicinity of the airport, under continuous demand. Like many simulation models, rS is based on "blocking" rules, meaning that it is built on a link-node system where each link can only hold a pre-specified maximum number of aircraft at any given time. rS is capable of estimating both capacity and delay (which requires input of a schedule). rS requires a basic set of operational inputs (not very different from ACM) although it does require more physical parameter inputs. Users can set up an analysis in rS relatively quickly in comparison to other more complex simulation models.

rS was validated by MITRE by comparing capacity results from rS to those of ACM for a number of simple scenarios [5]. As basic calculations are found to be correct, they are assumed to remain so for more complex scenarios. In addition, the animation can be viewed to insure that all ATC rules specified are followed correctly. The program is mainly used for in-house studies, although the Federal Aviation Administration (FAA) has begun using it as well.

Data was obtained from the Aviation System Performance Metrics (ASPM) database, which is part of the FAA's Operations and Performance Data system. Data from the "Download/Airport" section of the ASPM database was used for this analysis. This data includes hourly and quarter-hourly arrival and departure counts, demands, various weather conditions, and visibility conditions (either visual (VFR) or instrument (IFR) flight rules). It also provides detailed information on individual flights. ASPM count data are based on individual aircraft landing and take-off times as supplied through Airline Service Quality Performance (ASQP) data or Enhanced Traffic Management System (ETMS) messages. The data is available for 77 major airports in the United States.

To understand our methodology and its results, it is necessary to understand the demand data in our data set. Conceptually it is the number of flights that "want" to perform an arrival or departure information within a particular time period. It is based on the updated flight plan just before a flight is due to take off at the origin airport. In most cases, a flight counts toward demand beginning in the time period it is planned to land or take-off, and continuing through the time period when it actually does so. The only exception is when it arrives or takes off in a time period earlier than planned, in which case it is counted toward the demand in this earlier time period. This procedure ensures that the count never exceeds the demand. When the count and demand are equal, no flights are forced to wait until the next time period to perform their desired operation, while when demand exceeds count, there is delay. A shortcoming of this method for determining demand is that demand is not updated based on delays that are incurred 1) between the time the flight plan is filed and the aircraft is taxiing for take-off (departure demand) or 2) en route to the destination airport (arrival demand)¹. The implication is that the airfield demand may in actuality be lower than the ASPM demand data reports. This can lead to incorrectly attributing a difference between count and demand to a capacity constraint. This is not taken into consideration in the ensuing analysis, but has been done so previously by Hansen [6].

Quarter-hour and individual flight data from 2006 was obtained for both SFO and LAX. However, the runway configurations identified in the LAX ASPM data were found to be incorrect, so it was replaced with runway configuration data from Performance Data Analysis and Reporting System (PDARS). PDARS is joint NASA-FAA effort developed by ATAC Corporation. The database is fed by radar track and flight plan information directly from Automatic Radar Terminal System (ARTS) computers at Terminal Radar Approach Control (TRACON) facilities, and from the host computers at Air Route Traffic Control Centers (ARTCCs), which provide precise state information for each aircraft every 2 seconds. As PDARS data was readily available for January through March 2005, ASPM data for the corresponding time period was used instead of 2006 data. Also, the meteorological

¹ On the opposite end, the demand does not include the effects of ground delay programs (GDP), the effects of air traffic management (ATM), plus other mechanisms that would cause a flight to deviate from its schedule.

condition for each quarter hour was determined by checking the weather data against known thresholds at each airport [7].

V. METHODOLOGY

A. Experimental Procedure

Several steps were taken to perform the validation exercise. The first step involves choosing the hours to be analyzed, by grouping the quarter-hourly ASPM data into hourly bundles starting on the hour. Complete candidate hours for analysis were identified by filtering the hours through several criteria:

- The predominant runway configuration was in use for the entire hour (28L,28R | 1L,1R at SFO, 24R,25L | 24L,25R at LAX);
- The weather designation was VMC or IMC for the entire hour, and
- The hour falls within the period of the day with the highest average demands (which, based on the data, was found to be between 9 am and 2 pm at both airports).

The second step involves randomly drawing 50 hours (approximately half VMC and half IMC) from each filtered set. After filtering, the number of IMC hours at both SFO and LAX were low enough such that only 20 IMC hours were available for analysis. However, about 30 VMC hours were available.

The next step is to obtain capacity estimates from both ACM and rS for each of the 50 hours. Each hour can be distinguished from one another by meteorological condition, fleet mix, and arrival/departure split (in %), while runway configurations are held fixed at each airport's predominant configuration. As the purpose of this work was to assess model performance using minimal to no calibration, no additional edits were made after all inputs were complete.

The result is data set containing predicted capacities, observed counts, and other relevant variables for each of 50 hours at LAX, and likewise for SFO. These data serve as the basis for our two validation methods, which we now discuss.

B. Comparison of Predicted Capacities with Demand-unconstrained Counts

The first validation method is based on a simple comparison of the realized counts and the capacities predicted by the models. Recognizing that counts may reflect demand rather than capacity constraints, we selected those observations in which the demand exceeded the capacity. Since we chose hours during busy periods of the day, this turned out to be the majority of our observations. In addition to plotting demand-unconstrained counts against capacity, we calculated the Theil [8] inequality coefficient and its components for each model. Given a predicted and realized value for observation i , P_i and A_i , the coefficient is calculated as:

$$U = \frac{\sum_i (P_i - A_i)^2}{\sum_i A_i^2} \quad (1)$$

The inequality coefficient may be decomposed into three parts: bias or error in central tendency, U_m ; unequal variation, U_s ; and incomplete covariation, U_c . These components, normalized so that they sum to 1, are given by:

$$U_m = \frac{(\bar{P} - \bar{A})^2}{\frac{1}{n} \sum (P_i - A_i)^2} \quad (2)$$

$$U_s = \frac{(s_P - s_A)^2}{\frac{1}{n} \sum (P_i - A_i)^2} \quad (3)$$

$$U_c = \frac{2(1 - r)s_P s_A}{\frac{1}{n} \sum (P_i - A_i)^2} \quad (4)$$

In these expressions s_P and s_A are the standard deviations of the predicted and actual values, r is the correlation coefficient between P and A , and n is the sample size.

C. Censored Regression Model

We also used censored regression to evaluate the two models. A censored regression model is equivalent to an ordinary least squares (OLS) regression model in that it relates a dependent random variable Y to a set of independent variables X_1, X_2, \dots, X_n [9]. However, in censored regression it is assumed that Y cannot be observed beyond some minimum or maximum threshold value (or both). For instance, if a value of Y is larger than the maximum threshold value Y_{max} , then only Y_{max} is observed. The true value of Y - the latent variable Y^* - cannot always be observed due to this censoring effect, although X_1, X_2, \dots, X_n are always observable. Tobit regression accounts for this by ensuring that the parameters of the regression model estimate the effects of X_1, X_2, \dots, X_n on the latent variable Y^* and not on the censored (observed) variable Y . In this analysis the dependent observed variable Y is the airport's arrival or departure capacity. We are limited in our ability to measure capacity in that aircraft counts from our data cannot exceed the demand in any given time period t , despite the fact that capacity may actually be greater than the demand in that time period. Nor can throughput or capacity be less than zero. As a result, the observed throughput in t is censored from above, and there are two situations that can arise [6].

$$C_o(t) = \begin{cases} C_o^*(t), & \text{if } 0 < C_o^*(t) < C_o^U(t) \\ C_o^U(t), & \text{if } C_o^*(t) \geq C_o^U(t) \end{cases} \quad (5)$$

Where

$C_o(t)$ is the "observed" capacity for operation type o (arrivals or departures) in time interval t ,

$C_o^*(t)$ is the true (or latent) capacity for operation type o in time t , and

$C_o^U(t)$ is the upper bound of observable capacity (i.e. demand) for operation type o in time t .

In the first scenario of (5) counts are less than the demand ($C_o^U(t)$); in this case capacity can be equated to the count. The second scenario is the upper censor where counts equal

demand, and therefore capacity is measured to be this demand (although it could in reality be higher, therefore the censoring effect).

The basic model specification is introduced here.

$$C_o(t) = \beta_0 + \beta_1 * Mod_{x,o}(t) + \varepsilon \quad (6)$$

$$Q_o(t) = \min [D_o(t), C_o(t)] \quad (7)$$

Where

β_0 , β_1 , and σ_o are estimated parameters,

$Mod_{x,o}(t)$ is the capacity estimate from model x (ACM or rS), for operation o in t ,

$Q_o(t)$ is the throughput for operation o in t ,

$D_o(t)$ is the demand for operation o in t , and

ε is the iid error term, distributed Normal with mean 0 and variance σ_o^2 .

The model parameters are estimated from the data using maximum likelihood estimation (MLE).

If a given model yielded perfect capacity predictions, we would expect $\beta_o \rightarrow 0$, $\beta_1 \rightarrow 1$, and $\sigma_o^2 \rightarrow 0$. Thus the coefficients yielded by estimating these regressions provide a basis for scoring the validity of the models. Before moving on, it should be clarified that the subjects of this discussion are capacity estimates from three models: ACM and rS, as well as the empirical capacity regression model. As seen, the ACM and rS model estimates are obtained and then used as explanatory variables in the regression model. To avoid confusion, the capacity regression model will be referred to as the empirical or

regression model, while the ACM and rS models (if not referred to by name individually) will be called the test models.

D. Results

1) Predicted Capacities versus Unconstrained Counts:

Figs. 1 and 2 compare predicted capacities and realized counts for each model and each airport. The rS model yields better agreement with observed counts in the case of SFO (Fig. 1), while neither model does very well for LAX (Fig. 2).

Table I presents the Theil analysis results. For SFO, the rS arrival and departure inequality coefficients are much less than those of ACM, confirming its better predictive capability. The primary sources of inequality are also different, with bias (U_m) the major one for the ACM model as opposed to incomplete covariation (U_c) for the rS model. In the case of LAX, the inequality coefficients for the two models are comparable, as are the inequality proportions. In general, unequal variation (U_s) is the smallest contributor to the inequality of the predicted and actual data sets. Aggregating across the two airports, the rS model emerges as the better predictor, primarily because it exhibits less bias.

Arrival and departure capacity predictive performance appears to be highly correlated. Inequality coefficients for arrivals and departures are generally of very similar magnitudes, as are the inequality proportions.

2) *Regression Model I:* The results of the basic model (Equations 6 & 7) for ACM and rS are reported in Tables II, III, and IV. Table II contains results for SFO, Table III for LAX, and Table IV for a model that contains data from both airports. Recall that the empirical model results are based on about 30 VMC and 20 IMC observations at each airport.

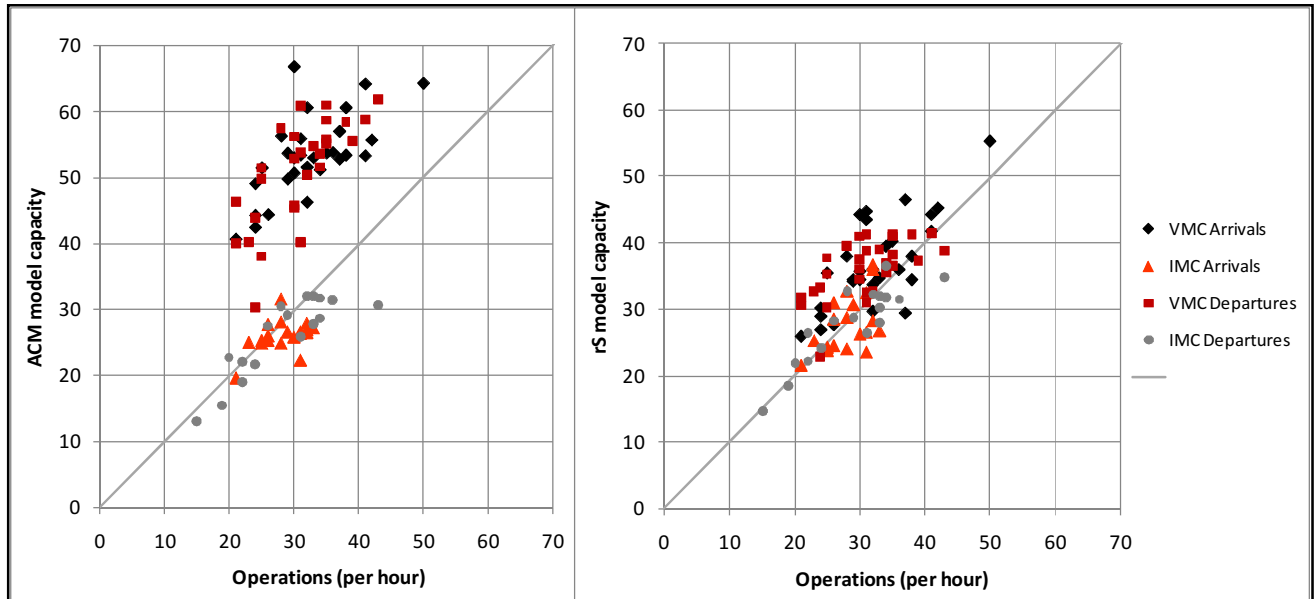


Figure 1. Model Capacity Estimates versus Unconstrained Counts, SFO

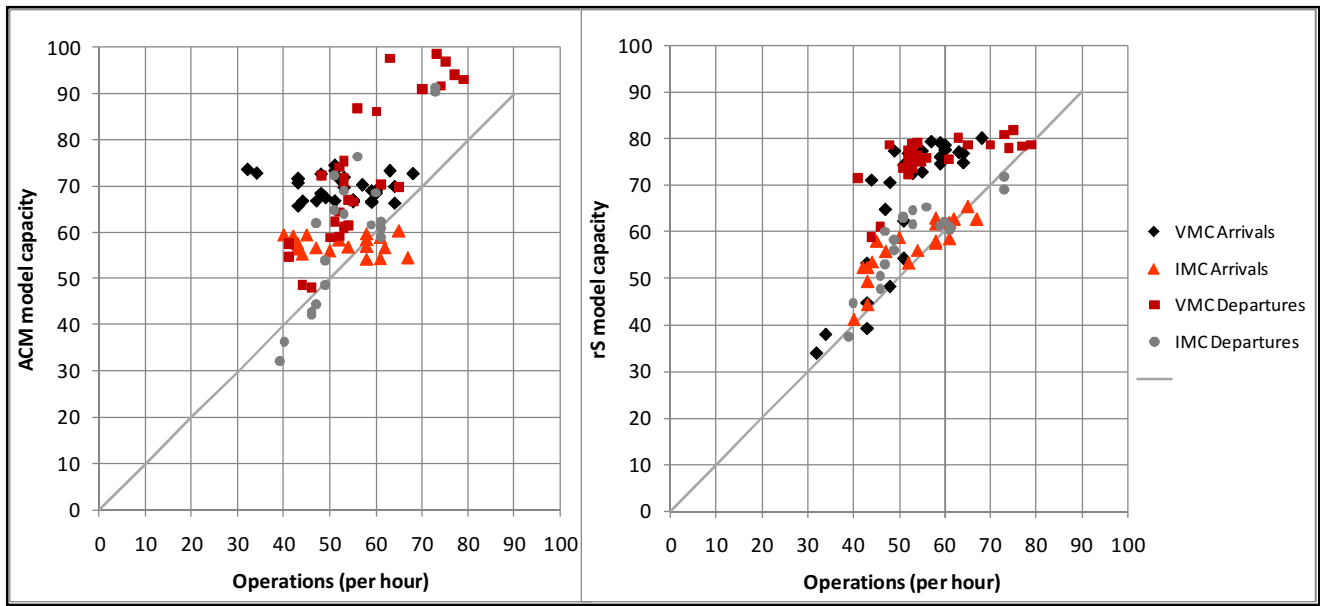


Figure 2. Model Capacity Estimates versus Unconstrained Counts, LAX

TABLE I. PREDICTION-REALIZATION ANALYSIS

ACM		RMS error		Inequality Coefficient (U)		Inequality Proportions					
						Arr			Dep		
<i>Air</i>	<i>MC</i>	<i>Arr</i>	<i>Dep</i>	<i>Arr</i>	<i>Dep</i>	U_m	U_s	U_c	U_m	U_s	U_c
<i>SFO</i>	<i>VMC</i>	21.3	20.6	0.646	0.657	0.941	0.000	0.059	0.924	0.013	0.063
	<i>IMC</i>	3.9	4.3	0.138	0.145	0.360	0.061	0.579	0.323	0.088	0.589
<i>SFO Total</i>		16.8	16.4	0.536	0.536	0.478	0.255	0.268	0.474	0.218	0.309
<i>LAX</i>	<i>VMC</i>	19.9	17.8	0.375	0.303	0.763	0.092	0.145	0.799	0.061	0.140
	<i>IMC</i>	10.4	11.1	0.197	0.201	0.244	0.407	0.348	0.279	0.404	0.318
<i>LAX Total</i>		16.8	15.4	0.317	0.269	0.555	0.017	0.428	0.586	0.190	0.224
<i>Total</i>		16.8	15.9	0.389	0.350	0.514	0.028	0.458	0.524	0.092	0.384
rS		RMS error		Inequality Coefficient (U)		Inequality Proportions					
						Arr			Dep		
<i>Air</i>	<i>MC</i>	<i>Arr</i>	<i>Dep</i>	<i>Arr</i>	<i>Dep</i>	U_m	U_s	U_c	U_m	U_s	U_c
<i>SFO</i>	<i>VMC</i>	6.6	6.9	0.200	0.220	0.461	0.003	0.536	0.578	0.038	0.385
	<i>IMC</i>	3.8	3.4	0.133	0.115	0.006	0.059	0.935	0.006	0.059	0.935
<i>SFO Total</i>		5.6	5.8	0.180	0.190	0.211	0.077	0.712	0.261	0.000	0.739
<i>LAX</i>	<i>VMC</i>	16.8	19.6	0.318	0.333	0.744	0.105	0.151	0.813	0.073	0.114
	<i>IMC</i>	6.0	6.4	0.113	0.116	0.373	0.145	0.482	0.403	0.018	0.578
<i>LAX Total</i>		13.6	15.6	0.256	0.272	0.561	0.096	0.343	0.601	0.002	0.397
<i>Total</i>		10.3	11.6	0.238	0.255	0.371	0.224	0.405	0.403	0.128	0.468

TABLE II. MODEL I RESULTS FOR SFO

ACM	Departure			Arrival		
	Estimate	Error	t-stat	Estimate	Error	t-stat
β_0	20.95	2.604	8.04	22.18	2.004	11.06
β_1	0.23	0.055	4.22	0.21	0.052	4.03
σ_0	1.71	0.100	17.12	1.59	0.105	15.08
rS	Departure			Arrival		
	Estimate	Error	t-stat	Estimate	Error	t-stat
β_0	7.79	2.579	3.02	12.41	3.023	4.10
β_1	0.69	0.080	8.61	0.56	0.093	5.98
σ_0	1.56	0.096	16.20	1.38	0.078	17.80

TABLE III. MODEL I RESULTS FOR LAX

ACM	Departure			Arrival		
	Estimate	Error	t-stat	Estimate	Error	t-stat
β_0	22.66	2.712	8.35	60.68	13.709	4.43
β_1	0.50	0.043	11.58	-0.12	0.212	-0.56*
σ_0	1.69	0.082	20.52	2.18	0.082	26.66
rS	Departure			Arrival		
	Estimate	Error	t-stat	Estimate	Error	t-stat
β_0	20.97	5.565	3.77	22.29	3.558	6.27
β_1	0.52	0.090	5.82	0.49	0.057	8.53
σ_0	2.12	0.085	24.84	1.83	0.091	20.10

* Results are not significant at the 95% confidence level.

TABLE IV. MODEL I RESULTS FOR BOTH SFO AND LAX

ACM	Departure			Arrival		
	Estimate	Error	t-stat	Estimate	Error	t-stat
β_0	8.59	1.933	4.44	11.73	2.247	5.22
β_1	0.64	0.031	20.48	0.57	0.046	12.30
σ_0	2.14	0.056	38.24	2.28	0.067	33.84
rS	Departure			Arrival		
	Estimate	Error	t-stat	Estimate	Error	t-stat
β_0	7.86	1.714	4.59	10.04	1.358	7.40
β_1	0.71	0.040	17.69	0.66	0.032	20.82
σ_0	1.93	0.070	27.53	1.72	0.074	23.29

From the results, it can be observed that all β_0 estimates are much greater than 0, and all β_1 estimates are smaller than 1. The magnitudes of the β_1 estimates indicate that the ranges of capacity estimates from ACM and rS are both wider than the corresponding ranges of actual capacities. The β_0 estimates are relatively large due to the same reason. The implication of these results is that for a set of conditions where the predicted capacity is low, the actual capacity is likely to exceed the prediction, while when the predicted capacity is high, it is likely to exceed the realized capacity. At SFO, the rS regressions' β_0 estimates are lower and β_1 estimates are higher

than those of the ACM regressions, suggesting that the rS capacity model performs better. The standard deviations in the ACM regression model are also higher than those of the rS regression model. At LAX it appears that the estimates from the ACM and rS regressions are comparable, except those of the ACM arrivals regression. The ACM arrival capacity predictions proved very insensitive to the different input conditions of the 50 hourly samples in IMC and VMC. As a result, the β_1 estimate is insignificant and β_0 simple reflects the average empirical capacity for the entire 50 hour sample.

Similar patterns are observed in the results for the combined airport model (Table IV). The β_0 values tend to be smaller and the β_1 values larger than those of the other two models. Moreover, the rS and ACM coefficients are quite similar in the combined model. This implies that both models do fairly well in predicting the difference in capacity between SFO and LAX in their primary configurations. Where the rS prevails, at least for SFO, is in its ability to predict variations in capacity for a particular airport and configuration. .

3) *Regression Model II*: The basic empirical model was modified to include another parameter that distinguishes between VMC and IMC test model capacity estimates. This serves to isolate the effect of visibility condition has on the test models' predictive performance.

$$C_o(t) = \beta_0 + \beta_1 * Mod_{x,o}(t) + \beta_2 * I_o(VMC = 1) + \varepsilon \quad (8)$$

$$Q_o(t) = \min [D_o(t), C_o(t)] \quad (9)$$

Where

β_2 is an estimated parameter, and

$I_o(VMC=1)$ is an indicator variable set to 1 if operation type o in time t occurs under VMC conditions, and 0 if it occurs in IMC.

The results of this model are contained in Tables V, VI, and VII as well as Figs. 3 through 5. Each figure contains the regression model capacity predictions plotted against predictions obtains directly from the test models. Overall, the inclusion of the VMC indicator variable improves the estimates for β_0 and β_1 . This suggests that the test models are not particularly accurate in predicting differences between IMC and VMC capacities. β_2 is an estimate of the adjustment that must be made to VMC test model capacity estimates to reflect actual capacities. With its addition, β_0 decreases in most of the regressions, while β_1 increases towards 1. This implies that the test models can more accurately predict variations in capacity given a particular visibility condition, rather than the effect that visibility condition has on capacity.

The β_2 values in Tables V, VI and VII are each less than zero (except that of the ACM arrivals regression at LAX, which is subject to the same problems that were discussed in the previous section). This implies that both models overestimate the difference between VMC and IMC capacities. Figures 3-5 show that this is because they overestimate VMC results.

TABLE V. MODEL II RESULTS FOR SFO

ACM	Departure			Arrival		
	Estimate	Error	t-stat	Estimate	Error	t-stat
β_0	11.88	2.692	4.41	12.93	4.022	3.22
β_1	0.66	0.096	6.87	0.60	0.152	4.00
β_2	-14.21	2.778	-5.12	-12.26	3.780	-3.24
σ_0	1.45	0.092	15.72	1.44	0.105	13.77
rS	Departure			Arrival		
	Estimate	Error	t-stat	Estimate	Error	t-stat
β_0	1.35	3.255	0.42*	11.36	3.172	3.58
β_1	0.99	0.121	8.20	0.62	0.109	5.64
β_2	-6.20	1.695	1.69*	-1.51	1.319	-1.15*
σ_0	1.42	0.082	0.08*	1.37	0.081	16.85

* Results are not significant at the 95% confidence level.

TABLE VI. MODEL II RESULTS FOR LAX

ACM	Departure			Arrival		
	Estimate	Error	t-stat	Estimate	Error	t-stat
β_0	21.96	2.479	8.8	95.43	32.732	2.92
β_1	0.54	0.043	12.64	-0.74	0.572	-1.29*
β_2	-3.56	1.559	-2.28	8.87	7.176	1.24*
σ_0	1.63	0.087	18.86	2.16	0.080	27.07
rS	Departure			Arrival		
	Estimate	Error	t-stat	Estimate	Error	t-stat
β_0	-5.11	6.61	-0.77*	19.61	3.331	5.89
β_1	1.02	0.118	8.69	0.59	0.056	10.59
β_2	-13.69	2.634	-5.12	-6.69	1.747	-3.83
σ_0	1.97	0.099	19.78	1.69	0.077	22.05

* Results are not significant at the 95% confidence level.

TABLE VII. MODEL II RESULTS FOR BOTH SFO & LAX

ACM	Departure			Arrival		
	Estimate	Error	t-stat	Estimate	Error	t-stat
β_0	8.62	1.488	5.80	5.83	2.065	2.82
β_1	0.77	0.036	21.73	0.84	0.059	14.24
β_2	-11.89	1.580	-7.53	-14.69	2.064	-7.12
σ_0	1.88	0.061	30.81	2.04	0.075	27.31
rS	Departure			Arrival		
	Estimate	Error	t-stat	Estimate	Error	t-stat
β_0	9.41	1.695	5.55	11.37	1.320	8.61
β_1	0.75	0.037	20.12	0.70	0.031	22.88
β_2	-6.59	1.140	-5.78	-4.96	1.110	-4.47
σ_0	1.82	0.076	23.97	1.62	0.062	26.17

According to Table V ACM overestimates VMC capacities at SFO but underestimates lower IMC capacities. The rS model appears to do a better job of estimating capacities than ACM, although again it seems to slightly overestimate VMC capacities. The standard deviations for both the ACM and rS regressions are about the same.

At LAX it again appears that capacity estimates from rS more closely reflect empirical values than do the ACM results (Table VI). The complete failure of ACM to predict arrival capacities is apparent, while on the departure side it slightly overestimates the difference between VMC and IMC capacities and, as in the previous results, exaggerates capacity variability. rS greatly exaggerates the capacity difference between IMC and VMC at LAX, but aside from that does very well at predicting capacity variation, as implied by the β_1 coefficient matching the ideal value of 1.

From the combined airports regression (Table VII), the performance of ACM and rS are more comparable. The rS model does a better job on the whole at predicting the difference in capacity between VMC and IMC. The models do equally well in predicting capacity variation from other sources, based on the β_1 results. There remains a tendency for the models to exaggerate capacity variation compared to what is actually observed. These results, like those in Table IV, are greatly influenced by the difference in capacity between SFO and LAX, and it is the ability of ACM to accurately predict that difference that makes it appear competitive with rS.

VI. CONCLUSIONS & FUTURE WORK

This paper has introduced two methodologies for validating capacity model results against empirical data. The validation results indicate that ACM and rS predict greater differences between average VMC and IMC capacity than the data indicates, as they typically appear to over-predict VMC capacities. The regression results also indicate that the models predict wider ranges of capacities than are seen empirically, as the β_1 coefficients are less than 1 in the majority of cases tested. Also, LAX arrivals capacities from ACM were insensitive to changing conditions, in comparison to the corresponding empirical capacities. Overall, it appears that rS estimates are typically better than those of ACM.

The work discussed in this paper can be continued and improve upon in several directions. It would be of interest to test other capacity models, particularly one of the more complex microscopic simulation models that are often used today. Capacities for additional runway configurations and other busy airports could be estimated and the regression model re-specified to include these. Also, arrival and departure capacity estimates could be tested together in one empirical model, to account for and assess their interaction effects. Finally, the capacity model estimation results could be compared against empirical capacity estimates based on other data sources such as PDARS.

A similar methodological framework is currently under development to assess empirically-derived en route capacity estimates against capacity predictions from existing controller workload models.

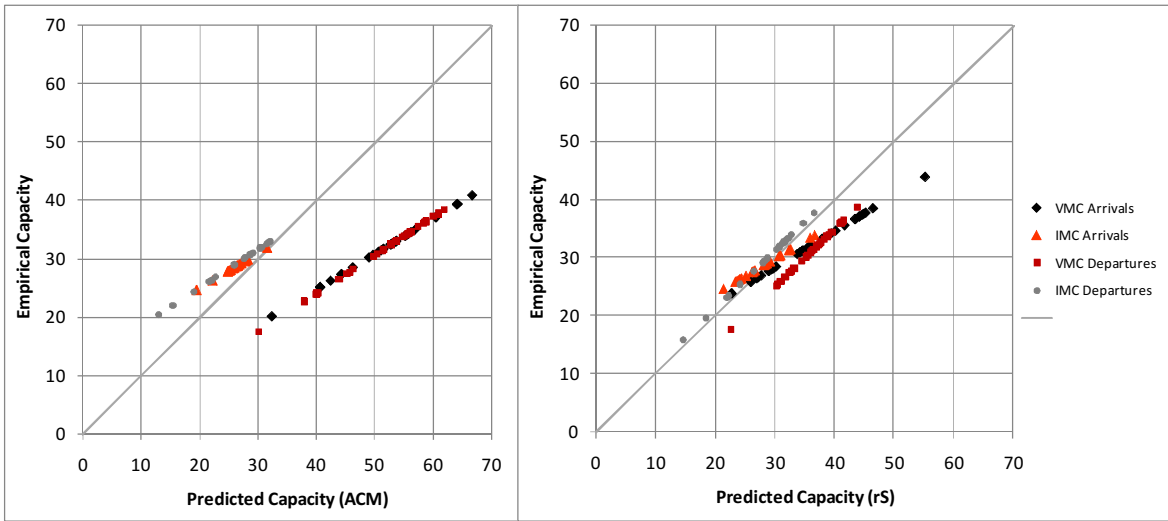


Figure 3. Model II Results for SFO

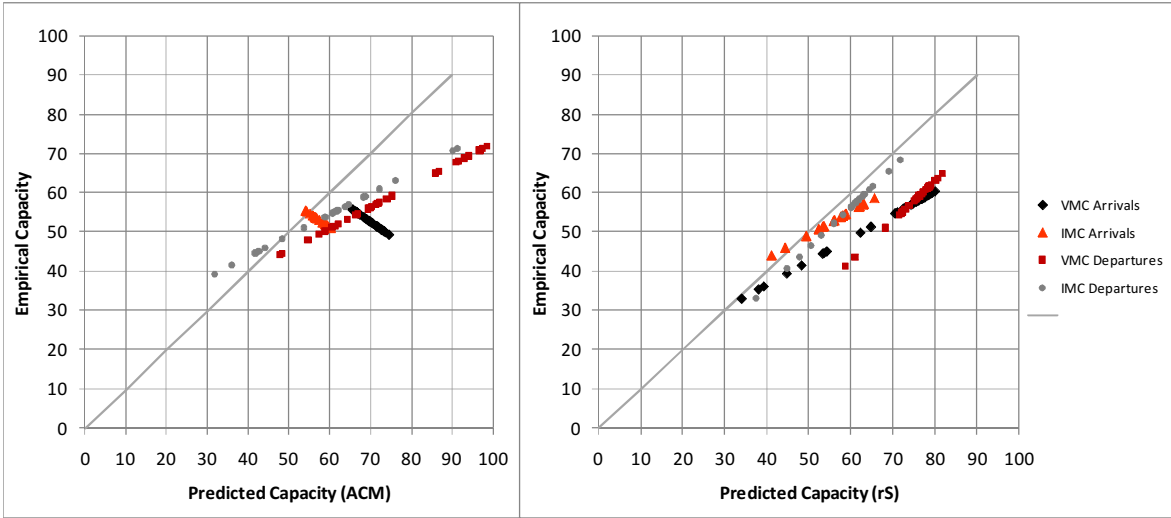


Figure 4. Model II Results for LAX

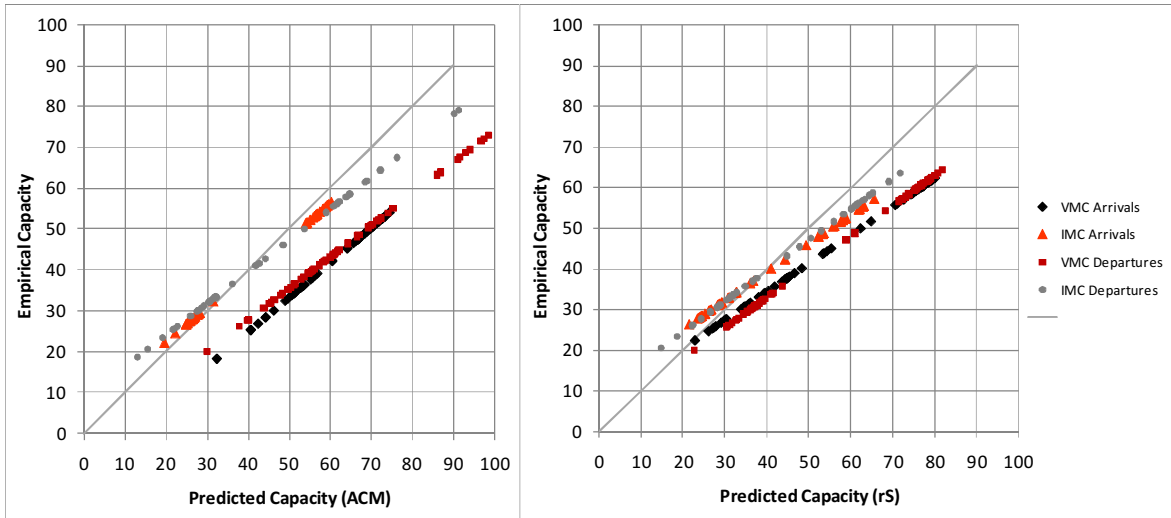


Figure 5. Model II Results for Both SFO & LAX

ACKNOWLEDGMENT

The authors would like to thank Joe Post at the FAA for his support of this research study.

REFERENCES

- [1] DeNeufville, Richard, and Amedeo Odoni. Airport Systems: Planning, Design, and Management. McGraw-Hill, 2003.
- [2] Lucic, Panta, Melissa Ohsfeldt, Mark Rodgers, and Alexander Klein. "Airport runway capacity model review." Research report by CSSI and Air Traffic Analysis, Inc. for FAA ATO-P Performance Analysis and Strategy, August 2007.
- [3] Odoni, Amedeo et al. "Existing and required modeling capabilities for evaluating ATM systems and concepts". Research report by the International Center for Air Transportation at MIT, for AATT Program (NASA), March 1997, <http://web.mit.edu/aeroastro/www/labs/AATT/index.html>.
- [4] Ball, Carl T. Model users' manual for airfield capacity and delay models. Washington, D.C.: Dept. of Transportation, Federal Aviation Administration, Systems Research and Development Service, 1976.
- [5] Barrer, John, and Peter Kuzminski. "Airport Capacity Analysis Through Simulation (ACATS): User's guide". Technical Report, Mitre Corporation, 2006.
- [6] Hansen, Mark. "Post-deployment analysis of capacity and delay impacts of an airport enhancement: case of a new runway at Detroit." Air Traffic Control Quarterly, Vol. 12, 2004, pp. 339-365.
- [7] Federal Aviation Administration, Airport Capacity Benchmark Report 2004, http://www.faa.gov/about/office_org/headquarters_offices/ato/publications/bench/2004download.htm.
- [8] Theil, Henri. Applied Econometric Forecasting. North-Holland Publishing Company, 1966.
- [9] Greene, William H. Econometric Analysis, 4th ed. New Jersey: Prentice Hall, 2000.

AUTHOR BIOGRAPHY

Amy Kim is currently a PhD student at the University of California at Berkeley. She received an undergraduate degree in Civil Engineering from the University of Waterloo in Canada in 2001, and an M.S. degree from UC Berkeley in 2002. Prior to returning to Berkeley in 2006, Amy worked as a traffic engineer and planner for two consulting companies – Dowling Associates and Delcan Corporation. Some of her research interests include air traffic operations, capacity analysis, and modeling.

Mark Hansen is a Professor of Civil and Environmental Engineering at the University of California, Berkeley. He graduated from Yale with a Bachelor's degree in Physics and Philosophy in 1980, and has a PhD in Engineering Science and a Masters in City and Regional Planning from UC Berkeley. Prior to graduate school, Dr. Hansen worked as a physicist at the Environmental Protection Agency. Since joining the Berkeley faculty in 1988, he has led transportation research projects in urban transportation planning, air transport systems modeling, air traffic flow management, aviation systems performance analysis, aviation safety, and air transport economics. He has taught graduate and undergraduate transportation courses in economics, systems analysis, planning, probability and statistics, and air transportation. Professor Hansen is the Berkeley co-director of the National Center of Excellence in Aviation Operations Research, a multi-university consortium sponsored by the Federal Aviation Administration. He is Chair of Transportation Research Board Committee AV-060, Airport and Airspace Capacity and Delay.