

This work was written as part of one of the author's official duties as an Employee of the United States Government and is therefore a work of the United States Government. In accordance with 17 U.S.C. 105, no copyright protection is available for such works under U.S. Law.

Public Domain Mark 1.0

<https://creativecommons.org/publicdomain/mark/1.0/>

Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

1
2
3 **Doing More with Less: An Assessment of Capacity Utilization**

4 **Using Stochastic Frontier and Spectral Analysis Models in the Case of**

5 **Atlanta Hartsfield/Jackson International Airport**
6
7
8

9 **ABSTRACT:** This study proposes a methodology to measure and compare the
10 operational outcomes of airport capacity before and after the implementation of
11 improvement programs at discrete periods at Atlanta Hartsfield/Jackson International
12 Airport (ATL). In this analysis, the percent of airport capacity utilized is a function of
13 operations, total demand, taxi-out times, and ceiling/visibility conditions. The outputs of a
14 stochastic frontier model indicated that the mean technical efficiency slightly declined to
15 0.98 in the 2016 sample, down from 0.99 in both the 2012 and 2015 samples. However,
16 the decline in technical efficiency can be misleading. Moreover, the outcomes of spectral
17 analysis indicated that ATL was able to reduce the amplitude and number of peaks per
18 cycle in the 2016 compared with the other samples, which resulted in improved capacity
19 utilized and on-time performance in a context of increasing operations.

20 **Keywords:** Stochastic frontier analysis, spectral analysis, capacity utilization
21

INTRODUCTION

Capacity utilization represents an important topic for air traffic regulators, airport operators, and aviation practitioners in the U.S. because there have been few new runways put in service at large hub airports over the last fifteen years. In this article, capacity utilization is defined as the ratio of total operations (hourly arrivals plus departures) to declared airport capacity (hourly airport arrival plus departure rates). The theory of airport capacity and queues predicts that airport congestion and resulting delays are likely to rise as airport capacity becomes fully utilized (Young and Wells¹; Horonjief and McKelvey²; Daniel³; Kleinrock⁴). As more aircraft arrive and depart, especially at peak times, congestion at the airport's gate area and taxiways is likely to create take-off queues.

Over the last ten years, the U.S. Federal Aviation Administration (FAA) has been implementing NextGen programs to ensure the transition of the National Airspace System (NAS) from a radar-based to a satellite-based navigation system through improved surveillance, data and capacity management, and communication capabilities. NextGen programs include complex technical capabilities designed to modernize the National Airspace System (NAS) and alleviate capacity constraints, both en route and at congested airports. They are also designed to reduce conflicting traffic in and out of large metropolitan areas called 'metroplexes.' Metroplexes refer to metropolitan areas where the proximity of large hubs and general aviation airports require deconfliction of approach and departure procedures. This is all the more important at times of capacity constraints resulting from high demand, increased volume, and low ceiling and visibility, among others.

This study proposes a different methodology to measure operational program outcomes, which is of great interest to transportation practitioners, regulators, as well as performance and budget analysts. It fills a gap in program evaluation methodologies by comparing the results from a stochastic frontier analysis (SFA) with those of a spectral analysis (SA). Other methods to measure capital program outcomes may include

- Capital investment metrics (i.e., internal rate of return, payback periods, net present value)
- Earned Value Management (i.e., planned value, earned value, and actual cost)
- The balanced score cards to monitor performance in four key areas including financial, customer and stakeholders, internal process, and organizational capability, and
- Key performance indicators based on critical success factors that an organization strives to achieve.

In our approach, the SFA model estimates whether total operations, total demand, taxi-out times, and visual approach conditions can explain variations in the percentage of airport capacity utilized as a measure of ‘technical efficiency’ in SFA parlance. On the other hand, the spectral model focuses on the study of capacity utilized in the frequency rather the time domain, which will be explained later in this article.

SFA has often been used to compare the productivity of airports and to benchmark them. SFA research has focused mainly on three major areas: (1) airport performance (Barros and Dieke⁵; Oum et al.⁶; Diana⁷); (2) benchmarking (Yoshida⁸; Diana⁹; Martin and Roman¹⁰; Barros and Dieke¹¹; Morrison¹²; Diana¹³; Adler et al.¹⁴); and (3) airport

67 competition (Gillen and Lall¹⁵; Martin et al.¹⁶; Scotti et al.¹⁷). Welch and Ahmed¹⁸ used
68 spectral analysis to evaluate airport performance. They suggested that spectral analysis
69 could be used “to reveal relationships between delay and throughput that might help
70 airlines improve scheduling accuracy.” However, spectral analysis has not been used to
71 examine changes in program performance as in the present study.

72 This study examines whether airport capacity utilized at ATL had changed at three
73 discrete periods (March to May 2012, 2015, and 2016) during the implementation of
74 specific NextGen capabilities. A decrease in ‘technical efficiency’ in the SFA context does
75 not necessarily mean that an airport is not performing at efficient capacity. A one-hundred
76 percent capacity utilization is not desirable and would most likely result in congestion and
77 delays, which would severally constrain airport operations. Being on the ‘efficiency
78 frontier’ in case of capacity utilization does not necessarily imply that an airport is efficient
79 in maximizing the use of its scarce resources (see Lin et al.¹⁹). Some slack in capacity
80 utilization may be necessary for an airport to recover from unexpected events such
81 periods of peak demand or poor weather conditions.

82 Furthermore, the methodology described in this study can help determine whether
83 variations and amplitude in traffic flows have changed as a result of program
84 implementation. This is best accomplished in the frequency rather than the time domain
85 because the percent of capacity utilized becomes normalized and is easier to compare
86 across relevant time periods.

87 Spectral analysis is the process of breaking down a signal into its components at
88 various frequencies. Usually, spectral density estimation is useful to detect any
89 periodicities in a data distribution and peaks that correspond to identified periodicities. In

the present case, spectral analysis can supplement to SFA by measuring capacity utilized in the frequency domain and validate any change measured in the time domain.

MODEL VARIABLES

Variables and Data Sources

All the variables originated from the Aviation System Performance Metrics (ASPM) data warehouse²⁰. The data were collected during the local core hours of 07:00 to 21:59 for all weekdays at three periods of NextGen program implementation.

To identify the relevant model variables, several SFA models featuring different combinations of variables were evaluated. The variables in the model with the lowest Akaike Information Criterion were selected²¹. The Akaike Information Criterion is defined as $AIC = 2k - 2\ln(\hat{L})$ with k representing the number of estimated parameters and (\hat{L}) being the maximum value of the likelihood function. All the selected variables were significant at a 95 percent level.

The SFA model includes the following variables:

- The ***Percentage of Total Available Capacity Utilized*** as the dependent variable. It is the ratio of the total number of observed hourly arrivals and departures (operations) to the airport's stated arrival plus departure capacity (hourly called rates). The available airport capacity is the sum of airport arrival rates (AAR) and airport departure rates (ADR), which reflects the selection of specific configurations as well as expected or actual weather conditions (i.e., ceiling, visibility, wind speed and angle, among others). NextGen programs and capabilities are designed to maximize available capacity—especially at times when volume, demand, and

weather conditions are constrained—through reduced separation, timed departures, precision-based navigation, among others. The percentage of airport capacity utilized represents an important metric because it underlines how well an airport anticipates traffic flows and demand based on published schedules, demand, runway configurations, as well as weather conditions at the airport and en route.

- **Total Operations** consist of the number of arrivals and departures observed during an hour.
- **Total Demand** includes arrival and departure demand. The former refers to the number of aircraft that have left the gates at the origin airports but have not yet landed at ATL (from gate-out to wheels-on). The latter consists of the number of planes that left the gates at ATL but have not yet taken off to their destination (from gate-out to wheels-off).
- **Taxi-Out Time** is the number of minutes it takes for an aircraft to move from gate-out to wheels-off. Taxi-out time is an indicator of surface congestion: taxi-out time is likely to increase as more aircraft are waiting in line for departure, which has a direct impact on capacity utilization. Taxi-in time was not significant at a 95 percent level in the three periods and, therefore, was not selected as part of the SFA model.
- **Visual Approach Conditions** refer to periods when ceiling is higher than 3,600 feet and visibility higher than 7 nautical miles at ATL. The variable is coded as ‘1’ for visual approach conditions or VAC and ‘0’ for instrument approach conditions or IAC in the data samples.

The inefficiency frontier model discussed in the next section includes the following additional variables:

- **On-Time Gate Arrivals.** It is the percentage of aircraft that arrive at the gate no later than the time filed in the last flight plan before takeoff. On-time performance can also be compared with airlines' published schedules in the Computer Reservation System (CRS). However, studies have shown that schedules may be padded to account for events such as headwind and surface congestion, for instance (Wu²²).
- **On-Time Gate Departures.** It is the percentage of aircraft that depart from the gate no later than the time filed in the last plan before takeoff.

The samples covered the months of March to May: the incidence of thunderstorms is lower during these months and they are not in the peak holiday or summer traffic seasons. Poor weather disruptions and peak traffic may affect the 'noise' component or v of the error terms, which is explained in the 'Model Assumption' section. The selection of the three years (2012, 2015, and 2016) provides three windows of observations during the implementation of key NextGen capabilities outlined in Table 1, mainly Automated Terminal Proximity Alert, wake vortex recategorization, and adjacent center metering.

155

Table 1. Timeline of NextGen Capabilities Deployed at ATL

Year	Capabilities
2006	Airport Surface Detection Equipment — Model X (ASDE-X) Area Navigation (RNAV) Global Positioning System (GPS) Approaches
2009	Area Navigation (RNAV) Standard Instrument Departures (SIDs) Required Navigation Performance (RNP) Authorization Required (AR) Approaches Adapted for Adjacent Center Metering (ACM)
2010	Area Navigation (RNAV) Standard Terminal Arrival Routes (STARs) Area Navigation (RNAV) Standard Terminal Arrival Routes (STARs) External Surface Data Release
2011	Expanded Low-Visibility Operations Using Lower Runway Visual Range (RVR) Minima Area Navigation (RNAV) Standard Terminal Arrival Routes (STARs) Optimized Profile Descents (OPDs) Equivalent Lateral Spacing Operations (ELSO)
2012	Automated Terminal Proximity Alert (ATPA) Phase 1 Expanded Low-Visibility Operations Using Lower RVR Minima
2013	Area Navigation (RNAV) Standard Terminal Arrival Routes (STARs) Optimized Profile Descents (OPDs) Expanded Low-Visibility Operations Using Lower RVR Minima Deployment of Time Based Flow Management (TBFM)
2014	Wake Re-Categorization Phase 1 — Aircraft Re-Categorization Situational Awareness and Alerting of Ground Vehicles Area Navigation (RNAV) Global Positioning System (GPS) Approaches
2015	Addition of Adjacent Center Metering (ACM) from Houston Air Route Traffic Control Center (ZHU) Qualifies for Independent Runway Standards in Order 7110.65
2016	Departure Clearance Tower Service Initial Operating Capability Area Navigation (RNAV) Standard Instrument Departures (SIDs)

Note: This timeline reflects programmatic milestones and excludes capabilities implemented across the National Airspace System. Information as of September 29, 2016.

156

157 Source: Federal Aviation Administration

158 Comparison of Variables among the Three Samples

159 Table 2 compares the values of the selected variables in the SFA model. The
160 takeoff weight determines the following categories:

- 161 ● 300,000 lbs. for heavy equipment (i.e., Boeing 777 or Airbus A 330)
- 162 ● 109,000 lbs. for Boeing 757-200

- From 41,000 to less than 300,000 lbs. for large aircraft (i.e., Boeing 737 or Airbus A 320)
- Below 41,000 lbs. for all other aircraft.

The counts of aircraft in the traffic mix originated from Traffic Flow Management System (TFMS) and represent the total number of arrivals at ATL identified by an AZ message in TFMS flight records during the sampled periods. Variations in the heavy and Boeing 757 aircraft types reflect Delta Air Lines' fleet restructuring after its merger with Northwest Airlines in 2012. However, this study does not focus on any specific carrier operations since the implementation of NextGen capabilities involves all carriers in collaboration with air traffic control.

When comparing user classes, Table 2 shows that the largest decline over the three sampled periods occurred in the air taxi category (mainly commuter flights), whereas the number of air carriers was the highest in the 2016 sample during the peak hours of 07:00 to 21:59 (local time). Scheduled air carriers are governed by 14 CFR Part 121 of the Federal Aviation Regulations, while air taxi and air charter operations are governed by 14 CFR Part 135.

Table 2. Sample Variables and Variations

Variables	March to May			Variation	
	2012	2015	2016	2016-2015	2016-2012
Average Percent Capacity Utilized	71.44	63.14	63.16	0.02%	-8.28%
Operations	2,362	2,206	2,255	2.22%	-4.53%
Total Demand	2,527	2,396	2,357	-1.63%	-6.73%
Taxi-Out Time (min)	19.08	17.04	16.96	-0.08	-2.12
Percent VAC	82.61	70.58	84.28		
Traffic Mix					
Heavy	6,083	5,402	5,429	0.50%	-10.75%
Boeing 757	16,945	12,710	9,937	-21.82%	-41.36%
Large	57,363	70,017	75,368	7.64%	31.39%
Other	38,285	23,461	24,232	3.29%	-36.71%
User Class					
Air Carrier	111,285	108,355	112,032	3.39%	0.67%
Air Taxi	5,739	1,730	1,374	-20.58%	-76.06%
Freight	773	821	863	5.12%	11.64%
General Aviation	711	607	632	4.12%	-11.11%
Military	113	26	30	15.38%	-73.45%
Other	55	51	35	-31.37%	-36.36%
Runway Configurations					
26R, 27L, 28 26L, 27R	59%	51%	61%	10%	2%
8L, 9R, 10 8R, 9L	35%	47%	32%	-14%	-3%
26R, 27L, 28 26L, 27R, 28	2%	1%	4%	3%	2%
Other Configurations	4%	2%	3%	1%	-2%
Total	100%	100%	100%		

Note: 180 observed periods from 07:00 to 21:59 local time

Source: FAA

185

186

Source: FAA

187

188

189

190

191

192

The main differences in the utilization of runway configurations happened in 2015 when more operations had a west-east flow pattern. Changes in the traffic mix (types of aircraft) may help explain changes in airport throughputs and differences among the model outputs, especially after the implementation of wake vortex re-categorization in June 2014. The separation between aircraft depends on aircraft classification based on takeoff weight, which affects runway occupancy and capacity utilization. The compression

of inter-departure and arrival times at peak times and more homogenous aircraft types (especially for the large aircraft category) may explain changes in capacity utilization as well as on-time performance improvement. The percentage of on-time gate arrivals and departures was included in the inefficiency frontier model as a way of assessing the potential impact of delays compared with flight plans.

MODEL ASSUMPTIONS

Stochastic Frontier Analysis Model

The SFA model makes it possible to measure the impact of selected operational variables on the dependent variables as well as to identify the error components that may explain technical inefficiency.

The production function can have two specific functional forms (see Battese and Broca²³). The first assumes that

$$f(x) = f(x; \beta) \quad (1)$$

with unknown parameters β . The second is a Cobb-Douglas function such that

$$y = \beta_0 + x_1^{\beta_1} + x_2^{\beta_2} + \dots + x_n^{\beta_n} \quad (2)$$

with unknown parameters $\beta_0, \beta_1, \dots, \beta_n$. Maximum likelihood determines the value of the estimates based on observed data.

This analysis utilized the ‘frontier’ package in R (Coelli and Henningsen²⁴) as well as the QLIM procedure in SAS. In the more traditional OLS model, inefficiency represents any deviation from the mean assuming $v \sim N(0, \sigma^2)$. In the SFA model, deviations are the results of ‘noise’ (v) and inefficiency (u). Considering eq. (1), the SFA model becomes

214
$$f(x) = f(x; \beta) + v - u \text{ or } y = \exp(x \beta) + v - u \quad (3)$$

215 with u as the efficiency term such that $u \sim N + (0, \sigma_u^2)$ and v as the error term such that v
 216 $\sim N(0, \sigma_v^2)$. The combined error term is $\varepsilon = v - u$. While v has a normal distribution, u has
 217 a half-normal distribution in the present model. The parameter v includes factors that
 218 airports can control such as number of departures and arrivals, the choice of runway
 219 configurations, and the airport departure and arrival rates. For a clear exposition of the
 220 density function of the error term v and the inefficiency term u , refer to Battese and Coelli²⁵
 221 and Bogetoft and Otto²⁶.

222 The terms u and v are independent. When $u = 0$, there is no inefficiency, which is
 223 not the case if $u > 0$. According to Schmidt and Sickles²⁷, technical efficiency implies
 224 “failure to produce maximal output, given the set of inputs used.” According to Battese²⁸,
 225 “technical efficiency of an individual firm is defined in terms of the ratio of the observed
 226 output to the corresponding frontier output, conditional on the levels of inputs used by that
 227 firm.” For instance, if total available capacity (arrival plus departure rates) is 100 during a
 228 given hour and the total number of arrivals and departures is 80, then the percentage of
 229 total airport capacity utilized is 80/100, that is, 80 percent.

230 In this study, the SFA model is defined as follows:

231
$$\ln(y) = f(\ln(x); \beta) + v - u \quad (4)$$

232 where $\ln(\text{Capacity Utilized}) =$

233
$$\alpha + \beta_1 \ln(\text{Operations}) + \beta_2 \ln(\text{Total Demand}) + \beta_3 \ln(\text{Taxi-Out Time}) + \beta_4 \text{VAC} + \varepsilon \quad (5)$$

234 with α as the intercept, β_n as the variable estimates, and $\varepsilon = v - u$.

235 Based on Battese and Coelli²⁹, the inefficiency frontier is defined as

$$U_{it} = z_{it} \delta W_{it} \quad (6)$$

where W_{it} is a random variable for the i th unit at time t , $z_{it} \delta$ is the mean of the normal distribution truncated at zero, and $W_{it} \geq -z_{it} \delta$. As a result,

$$TE_{it} = \exp(-U_{it}) = \exp(-z_{it} \delta - W_{it}) \quad (7)$$

In this study, the inefficiency frontier model (referred to as the ‘efficiency effects frontier’ in the ‘frontier’ package for R) is

$$U_{it} = z_{it} \alpha + z_{it} \delta_1 \text{On-Time Arrivals}_{it} + z_{it} \delta_2 \text{On-Time Departures}_{it} \quad (8)$$

with α as the intercept.

Spectral Analysis

The purpose of spectral analysis is to determine whether a decrease in technical efficiency in the time domain translates into changes in capacity utilization in the frequency domain. Using Fast Fourier Transform, the hourly distribution of total capacity utilized can be decomposed into a spectrum of frequencies over a specific range or window. Capacity utilized represents a ‘signal’ including noise (inefficiencies) whose frequency content can be compared among samples. While the SFA model evaluates changes in the efficiency frontier through a series of metrics detailed in the next section, spectral analysis makes it possible to transform a time series into several sine and cosine waves of different frequencies. According to Engle³⁰, the spectrum is “a plot of the squared amplitude of each component against the frequency of that component [...] The spectrum is a decomposition of the variance into the components contributed by each frequency.”

This analysis utilized the SciPy and QuantEcon packages in the Python programming language (SciPy³¹; Sargent and Stachurski³²). Based on the finite Fourier transform, the series x_t can be expressed as the sum of sine and cosine waves such that

$$x_t = \frac{a_0}{2} + \sum_{k=1}^m [a_k \cos(w_k t) + b_k \sin(w_k t)] \quad (9)$$

with t representing time; x_t , the data; n , the number of observations; m , the number of frequencies in the Fourier decomposition; a_0 , the mean term; a_k , the cosine coefficient; b_k , the sine coefficients; and w_k , the Fourier frequencies. The amplitude periodogram J_k is defined as:

$$J_k = \frac{n}{2} (a_k^2 + b_k^2) \quad (10)$$

The frequency ranges from 0 to π and is expressed in radians. If d is the number of observations per unit of time, then the period of the cycle is computed as $P =$

$$\frac{2\pi}{(d * frequency)} \simeq 0.37. \text{ The frequency in cycles per 15-hour period equals } Frequency * \left(\frac{d=15}{2\pi}\right) \simeq 2.67.$$

Spectrum analysis can help determine whether the amplitude of capacity utilization may have changed when comparing the three samples. Multiple peaks during the observed cycle would support the hypothesis of more periods of constrained versus unconstrained capacity, while the spectral density would assess the amplitude of these periods. The assumption is that NextGen programs would reduce the amplitude and number of peaks by improving capacity utilization.

MODEL OUTPUTS

Stochastic Frontier Analysis Model Outputs and Interpretation

Table 3 provides information about the endogenous variable, the number of observations in the three samples and the optimization method to derive the estimates. Table 3 indicates a decrease in the average percentage of capacity utilized over the 2012, 2015, and 2016 samples, respectively 71.24, 63.41, and 63.16 percent (when transforming the log values). The reduction in the Akaike and Schwarz information criteria values shows an improvement in the relative quality of the 2016 sample model compared with the other two sample models.

Table 3. The Percentage of Total Capacity Utilized

Variable	Mean (March-May)		
	2012	2015	2016
Log(Capacity Utilized)	4.2661	4.1496	4.1456
Standard Error	0.2092	0.2310	0.2097
Number of Endogenous Variables	1	1	1
Number of Observations	1,380	1,380	1,380
Log Likelihood	2201	2030	808.93506
Maximum Absolute Gradient	0.2310	0.0748	0.0001
Number of Iterations	31	33	18
Optimization Method	Quasi-Newton	Quasi-Newton	Quasi-Newton
AIC	-4388	-4046	-1604
Schwarz Criterion	-4351	-4009	-1567
Sigma	0.0491	0.0556	0.1491
Lambda	0.0004	0.0004	1.7444

Table 4 provides the estimates and statistics for the SFA model. All the estimates were significant at a 95 percent level.

Table 4. SFA Model Estimates

Parameter	March-May 2012		March-May 2015		March-May 2016	
	Estimate	Approx Pr > t	Estimate	Approx Pr > t	Estimate	Approx Pr > t
Intercept	-0.9487	<.0001	-1.1294	<.0001	2.7201	<.0001
Log(Operations)	0.9196	<.0001	0.9559	<.0001	0.8030	<.0001
Log(Total Demand)	0.0814	<.0001	0.0409	<.0001	0.1364	<.0001
Log(Taxi-Out Time)	0.0791	<.0001	0.1474	<.0001	0.1943	<.0001
VMC=1	-0.0857	<.0001	-0.1353	<.0001	-0.1069	<.0001
σ_v	0.0491	<.0001	0.0556	<.0001	0.0741	<.0001
σ_u	0.00002	<.0001	0.00002	<.0001	0.12931	<.0001
$\sigma^2 = \sigma_u^2 / \sigma_v^2$	0.049		0.056		0.203	
Lambda (λ) = $\text{sqrt}(\sigma_u^2 / \sigma_v^2)$	0.0004		0.0004		1.744	
Gamma (γ) = $\sigma_u^2 / (\sigma_u^2 + \sigma_v^2)$	0.000		0.000		0.753	
Total Error Variance due to Inefficiency = $\lambda^2 / (\lambda^2 + 1)$	0.000		0.000		0.753	

If $u = 0$, there is no difference between the Ordinary Least Squares (OLS) model and SFA. The SFA model did not capture variation in inefficiency (σ_u) in 2012 and 2015. However, if $u > 0$, only the SFA model can indicate that the independent variables accounted for only 13.94 percent of the variation of inefficiency.

The estimates for the four independent variables can be interpreted as follows. If operations (arrivals and departures) increased by one percent, the percentage of total capacity utilized would increase by 0.92, 0.96, and 0.08 percent, respectively in the 2012, 2015, and 2016 samples, holding other variables constant. Although operations increased by 2.22 percent between the 2015 and 2016 periods, operations in the 2016 sample were still 4.53 percent lower than those in the 2012 period (Table 2). The changes from 2014 to 2016 may be attributed to the implementation of wake vortex recategorization and the addition of adjacent center metering from ZHU. The increased volume of operations had

a lesser impact on the elasticity of capacity utilized in the 2016 sample compared with the other two periods of study, holding other variables constant. This may be due to the combined impact of ATPA, wake vortex recategorization, and adjacent center metering, which may have created more slack in the system. Consequently, the percent capacity utilized is further away from the efficiency frontier, which measures full capacity utilization.

If total demand (arrival and departure demand) increased by one percent, we would expect the percentage of total capacity utilized to rise by 0.08, 0.04, and 0.14 percent, respectively, holding other variables constant. Despite an increase in operations between the 2015 and 2016 sample, total demand declined by 1.63 percent (Table 2). This suggests that the combined impact of implemented capability contributed to better allocation of aircraft to arrivals and departures between the hours of 07:00 to 21:59 in the 2016 sample compared with the 2015 sample. Wake vortex re-categorization may have improved the optimization of traffic mix at peak times.

Table 2 shows changes in the average minutes of taxi-out times, which is one of the key indicators of airport congestion. If taxi-out times increased by one percent, the percentage of total capacity utilized would increase respectively by 0.08, 0.15, and 0.19 percent respectively, holding other variables constant. As a result, utilization becomes more sensitive to aircraft separation.

If approach conditions were not limited by ceiling and visibility, then the percentage of capacity utilized would increase by 0.92, 0.87, and 0.89 percent respectively, holding other variables constant.

Table 4 contains some important information on u and v . In the 2016 sample, the variance for inefficiency ($\sigma_u^2 = 0.1293^2 \simeq 0.0167$) was about 204.15 percent higher than

the variance for the random errors ($\sigma_v^2 = 0.0741^2 \approx 0.0055$). The variance for inefficiency is computed as $\lambda^2 - 1 = 1.744^2 - 1 = 204.15$ percent. However, the variance for inefficiency was 0.99 percent lower in the 2016 sample than that of the random errors in the 2012 and 2015 periods.

The percentage of the total error variance due to inefficiency is computed as follows:

$$\% TEV = \frac{\lambda^2}{\lambda^2 + 1} (11)$$

In the 2012 and 2015 samples, total error variance could not explain inefficiency since gamma (γ) was zero. However, in the 2016 sample, the percent of the total error variance due to inefficiency was 75 percent, while the remaining 25 percent could be attributed to random variation. The proportion of the total variance due to inefficiency is $\sigma^2 = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2}$.

If $\gamma = 0$ (as it is the case for the 2012 and 2015 samples), v representing statistical noise is insignificant and there is no difference between OLS and SFA model estimates. On the other hand, if γ close to 1, then technical inefficiency explains most of the deviations from the production frontier (see Aigner et al.³³; Meeusen and van den Broeck³⁴ for further explanations of the production frontier).

Table 5 indicates that the marginal effect of on-time performance is not significant at a 95 percent level in the three samples. Table 5 adds some factors (z) that may affect the efficiency frontier (the percentage of flights arriving and departing on time when compared with the last plan filed before takeoff). The coefficients of the additional explanatory factors can be interpreted as the marginal effects on the relative change of

capacity utilized. At a 95 percent level, the percentage of on-time arrivals and departures did not explain variations in U_{it} in the three samples.

Table 5. Efficiency Effects Frontier Outputs

Variables	March-May 2012		March-May 2015		March-May 2016	
	Estimate	Pr(> z)	Estimate	Pr(> z)	Estimate	Pr(> z)
(Intercept)	-0.771	0.000	-1.314	0.185	2.421	0.000
log(ops)	0.800	0.000	0.918	0.222	0.102	0.000
log(tot_dem)	0.176	0.000	0.064	0.931	0.137	0.000
log(txout_tm)	0.036	0.010	0.206	0.825	0.231	0.000
Z_(Intercept)	-0.180	0.939	-0.056	0.768	-1048.100	0.208
Z_otm_arr	-0.012	0.163	0.002	0.490	13.491	0.206
Z_otm_dep	0.007	0.766	-0.001	0.825	-18.862	0.207
sigmaSq	0.003	0.000	0.007	0.043	172.200	0.208
gamma	0.027	0.000	0.000	1.000	1.000	0.000

Not significant at 95% level

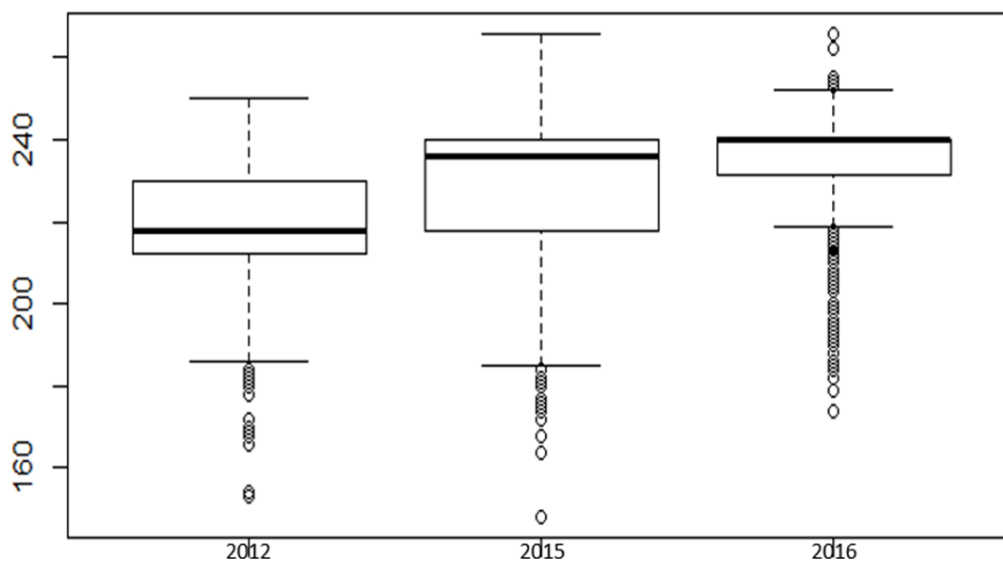
This analysis illustrates that ‘inefficiency’ parameters have to be interpreted with caution in the case of airport capacity utilization. Several conditions may justify the decline in the efficiency coefficients.

NextGen capabilities such as wake vortex re-categorization, Automated Terminal Proximity Alert (ATPA), and Equivalent Lateral Spacing Operations (ELSO) enabled ATL to reduce the percentage of total capacity utilized by compressing arrival and departure streams due to reduced separations and more efficient departure routes. This, in term, may explain why the airport did not reach full capacity utilization for extended periods and why total demand declined even though operations increased.

Second, capabilities such as wake vortex re-categorization allow controllers to increase peak arrival and departure throughputs. Figure 1 shows that the median available capacity (bold line within the boxplot) increased over the three observed

samples. It was 218, 236, and 240, respectively (source: ASPM). The interquartile range (difference between the 75th and 25th percentiles) reached a peak in the 2015 sample at 118, up from 97 in the 2012 sample. However, it went down to 92 in the 2016 sample. While the interquartile range declined in the 2016 sample, the boxplot also shows there were more outliers up to the 25th percentile compared with for the other two samples. The skewness coefficient was -0.94, -1.03, and -1.60, respectively, which means a longer tail to the left in the 2016 sample. The coefficient of kurtosis was respectively 1.96, 0.36, and 3.00, which implies an increased peakedness in the distribution of available total capacity when examined in the time domain.

Figure 1. ATL: Total Capacity (ADR plus AAR), March to May 2012, 2015, 2016



In the next section, we will assess the variation of total capacity utilized in the frequency domain to identify any change in patterns and amplitude.

Spectral Analysis Outputs and Interpretation

Table 6 provides two statistics to test for white noise as a measure of goodness-of-fit. The SPECTRA procedure in SAS provides Fisher's Kappa (κ) and Bartlett's

Kolmogorov-Smirnov (K-S) statistics. Readers interested in further details about these tests are referred to Bartlett³⁵; Durbin and Brown³⁶, and Fuller³⁷.

Table 6. White Noise Tests

Test for White Noise for Variable Capacity Utilized			
	2012	2015	2016
M-1	689	689	689
Max(P(*))	58020.34	44365.79	19835.22
Sum(P(*))	198790.4	203869	141159.6
Kappa	201.0963	149.9396	96.81573
Kolmogorov-Smirnov Test			
	2012	2015	2016
Test Statistics	0.41218	0.301913	0.209687
Approximate P-Value	<.0001	<.0001	<.0001

If observations in a time series are not autocorrelated, then J_k will have the same expected value for all k and the largest J_k values will be different from the mean of the J_k . When Fisher's Kappa test statistics is greater than the 5 percent critical value, then the null hypothesis that the percentage of total capacity utilized is white noise can be rejected. With $M = 689$, the Kappa statistics are all greater than the critical value of 9.47 ($M = 700$ with $Pr > 0.05$). The formula to derive Kappa is

$$\kappa = M * \frac{Max(P(*))}{Sum(P(*))} \quad (12)$$

The Kolmogorov-Smirnov statistic compares the normalized cumulative periodogram with the cumulative distribution function of a uniform $U(0,1)$ random variable. At a 95 percent level, we reject the null hypothesis of white noise in the three

samples since $p < 0.0001$. Therefore, the distribution of capacity utilized did not follow a normal distribution. The formula for the Bartlett's K-S test is

$$F_j = \frac{\sum_{k=1}^j J_k}{\sum_{k=1}^m J_k} \quad j = 1, 2, \dots, m-1 \quad (13)$$

with $m = n / 2$ for even number of observations and $m = n - 1$ for odd number of observations.

The periodogram and spectral density graphs are two important tools in spectral analysis. The QuantEcon package in Python served to generate Figure 2 to 4. The spectral density function shows the strength of the variations in capacity utilized as a function of frequency. The spectral density function identifies at which frequency range variations are going to be stronger. The periodogram is used to describe and identify the dominant cycles in a time series. The amplitude determines the maximum height of capacity utilized. In Figure 2 to 4, the periodograms show a dominant spike at a low frequency (around 1.1). The period for 1.1 is $1/1.1 = 0.91$, which means it takes 0.91 time periods for a complete cycle or 13.65 minutes. The graph utilized an ARMA (autoregressive moving average) process such that

$$X_t = \phi X_{t-1} + \varepsilon_t - \theta \varepsilon_{t-2} \quad (14)$$

with $n = 128$ randomly sampled observations, $\phi = 0.5$, $\theta = (0, -0.8)$, and ε_t as a white noise with unit variance. Figure 2 to 4 compare the periodogram to the actual spectral density. Φ is the autocorrelation value for the autocorrelated variable (autoregressive component), and θ is the autocorrelation value for the white noise of the model (moving average component).

Figure 2. Periodogram and Spectral Density of Percentage of Total Capacity Utilized by Frequency (March-May 2012)

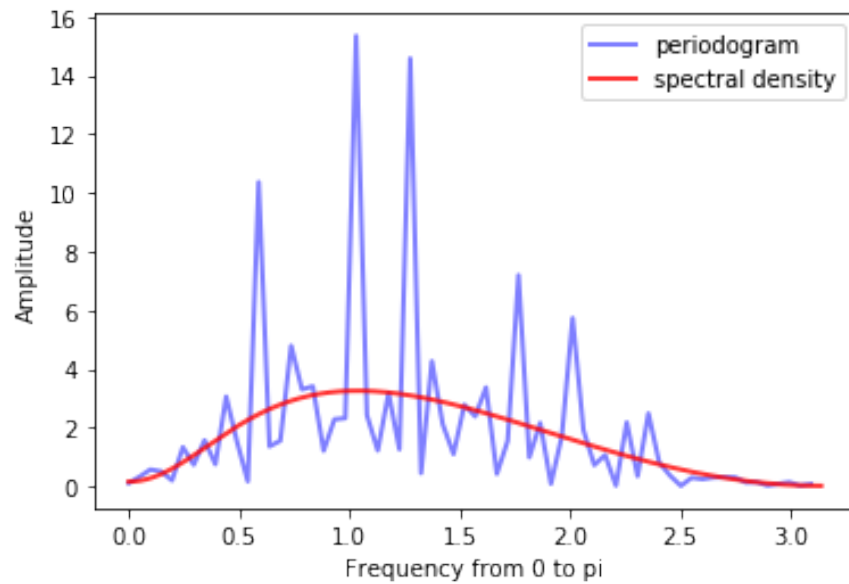


Figure 3. Periodogram and Spectral Density of Percentage of Total Capacity Utilized by Frequency (March-May 2015)

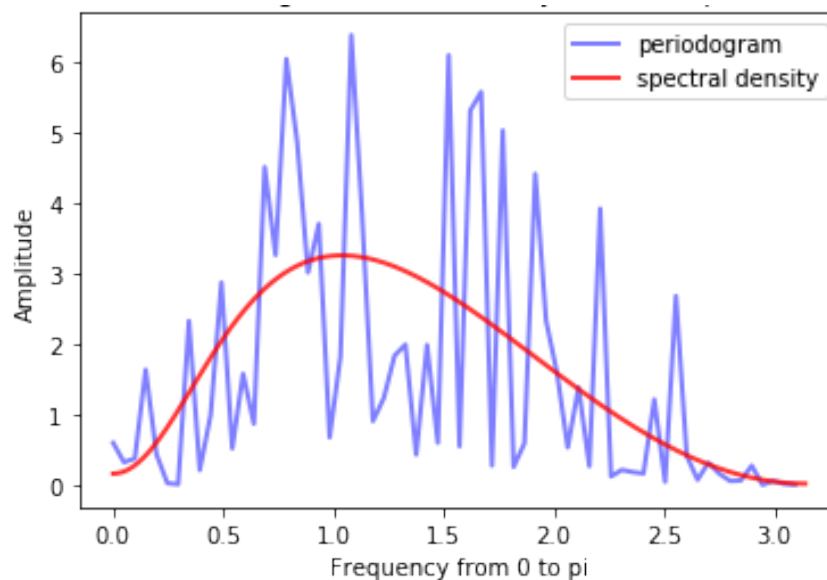
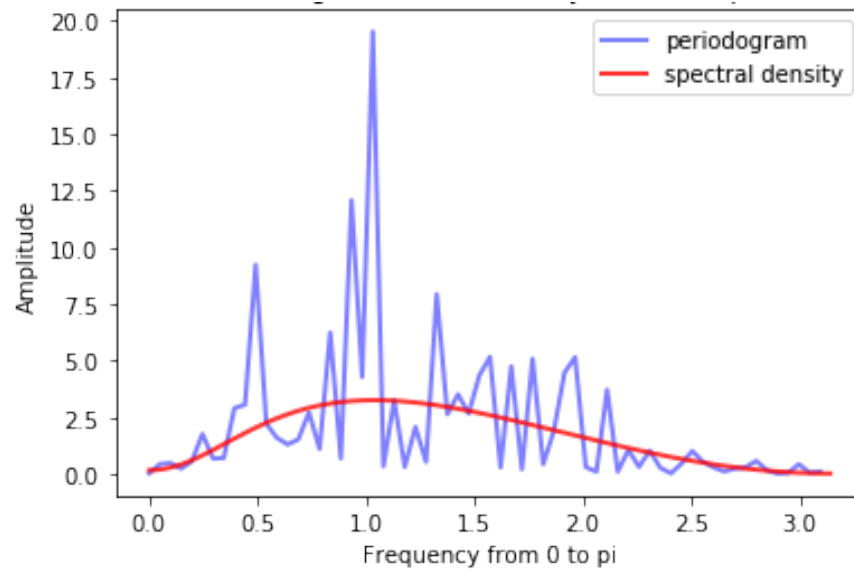


Figure 4. Periodogram and Spectral Density of Percentage of Total Capacity Utilized by Frequency (March-May 2016)



In Figure 2 and 3, the periodograms show multiple spikes in the zero-to-pi frequency range compared with Figure 4. The spectral density function implies that the energy or strength of the variations in capacity utilized as a function of frequency was more evenly spread in Figure 2 and 4, whereas the variation in the spectral density decreased more exponentially in Figure 3. This may be due to shifts in arrival/departure patterns and more periods of approach in instrument conditions in the 2015 sample. In the 2016 sample, there was a clear peak in capacity utilization around the frequency of 1.1.

FINAL COMMENTS

This study proposed a methodology to evaluate operational program outcomes. First, it considered the stochastic frontier analysis model because it provides the possibility to compute an efficiency frontier. In theory, if an airport reaches its efficiency

frontier, then it is likely to be more technically efficient. However, this is not the case for airport capacity utilization. If an airport uses all its available capacity, then congestion and delays are likely to arise. The airport may not have enough slack to face sudden increases in demand or operations.

The technical efficiency of some capital programs is difficult to measure because it depends on specific conditions and periods when programs are best utilized and effective. A time-based analysis would not provide any insights on airport performance resulting from programs designed to address specific constraints. This paper argues that spectral analysis can supplement the SFA methodology because it normalizes the amplitude and variations of the target variable within frequency and cycles, which makes comparisons among periods much easier.

Although the mean efficiency coefficient of total available capacity utilized declined in the 2016 sample, spectral analysis shows that the number of peak periods in the frequency domain also declined from 2015 to 2016. NextGen capabilities such as wake vortex re-categorization may have contributed to a reduction in inter-departure and arrival times at peak periods and, thus, alleviated constraints on capacity at times of peak throughputs. Reductions in the amplitude of capacity utilization was also facilitated by a more homogenous traffic mix at departures. Finally, spectral analysis can help better understand the outcomes of the SFA model especially when the decline in the technical efficiency coefficient may not be intuitive.

REFERENCES

- [1] Young S, Wells A. Airport planning and management, 6th edition. New York: McGraw Hill; 2011.
- [2] Horonjeff R, McKelvey FX. Planning and design of airports, 5th edition. New York: McGraw-Hill; 2010.
- [3] Daniel JI. Congestion pricing and capacity of large hub airports: A bottleneck model with stochastic queues. *Econometrica: Journal of the Econometric Society*. 1995 Mar 1;327-70.
- [4] Kleinrock L. Queueing systems, volume 2: Computer applications. New York: Wiley; 1976 Apr.
- [5] Barros CP, Dieke PU. Measuring the economic efficiency of airports: A Simar–Wilson methodology analysis. *Transportation Research Part E: Logistics and Transportation Review*. 2008 Nov 1;44(6):1039-51.
- [6] Oum TH, Yan J, Yu C. Ownership forms matter for airport efficiency: A stochastic frontier investigation of worldwide airports. *Journal of urban economics*. 2008 Sep 1;64(2):422-35.
- [7] Diana T. Can we explain airport performance? A case study of selected New York airports using a stochastic frontier model. *Journal of Air Transport Management*. 2010 Nov 1;16(6):310-4.
- [8] Yoshida Y. Endogenous-weight TFP measurement: methodology and its application to Japanese-airport benchmarking. *Transportation Research Part E: Logistics and Transportation Review*. 2004 Mar 1;40(2):151-82.

- 486 [9] Diana T. Benchmarking airport efficiency: an application of data envelopment
487 analysis. *Air Traffic Control Quarterly*. 2006 Jul;14(3):183-202.
- 488 [10] Martín JC, Román C. A benchmarking analysis of Spanish commercial airports. A
489 comparison between SMOP and DEA ranking methods. *Networks and Spatial
490 Economics*. 2006 Jun 1;6(2):111-34.
- 491 [11] Barros CP, Dieke PU. Performance evaluation of Italian airports: A data
492 envelopment analysis. *Journal of Air Transport Management*. 2007 Jul
493 1;13(4):184-91.
- 494 [12] Morrison W. Understanding the complexities and challenges of airport performance
495 benchmarking. *Journal of Airport Management*. 2009 Jan 1;3(2):145-58.
- 496 [13] Diana T. Measuring and Benchmarking Airport Efficiency: An Application of Data
497 Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SF). *Airline
498 Industry: Strategies, Operations and Safety*. 2011:143-59.
- 499 [14] Adler N, Liebert V, Yazhemsky E. Benchmarking airports from a managerial
500 perspective. *Omega*. 2013 Apr 1;41(2):442-58.
- 501 [15] Gillen D, Lall A. Developing measures of airport productivity and performance: an
502 application of data envelopment analysis. *Transportation Research Part E:
503 Logistics and Transportation Review*. 1997 Dec 1;33(4):261-73.
- 504 [16] Martín JC, Román C, Voltes-Dorta A. A stochastic frontier analysis to estimate the
505 relative efficiency of Spanish airports. *Journal of Productivity Analysis*. 2009 Jun
506 1;31(3):163-76.

- [17] Scotti D, Malighetti P, Martini G, Volta N. The impact of airport competition on technical efficiency: A stochastic frontier analysis applied to Italian airport. Journal of Air Transport Management. 2012 Jul 1;22:9-15.
- [18] Welch J, Ahmed S. Spectral analysis of airport performance. In AIAA's 3rd Annual Aviation Technology, Integration, and Operations (ATIO) Forum(2003 Nov 17-19); 6715).
- [19] Lin ZF, Choo Y, Oum TH. Efficiency benchmarking of North American airports: Comparative results of productivity index, data envelopment analysis and stochastic frontier analysis. In Journal of the Transportation Research Forum 2013 Mar 1 (Vol. 52, No. 1).
- [20] U.S. Department of Transportation. [Internet]; 2020 [cited 2020 Dec 26]. Aviation System Performance Metrics. Available at: <https://aspm.faa.gov>.
- [21] Akaike H. A new look at the statistical model identification. IEEE transactions on automatic control. 1974 Dec;19(6):716-23
- [22] Wu, CL Airline operations and delay management: insights from airline economics, networks and strategic schedule planning. Farnham: Ashgate Publishing, Ltd; 2012.
- [23] Battese GE, Broca SS. Functional forms of stochastic frontier production functions and models for technical inefficiency effects: a comparative study for wheat farmers in Pakistan. Journal of productivity analysis. 1997 Nov 1;8(4):395-414.
- [24] Coelli T, Henningsen A. [Internet]; 2017 [cited 2020 Dec 26]. Frontier: Stochastic Frontier Analysis. R package version 1.1-2. Available from : <http://CRAN.R-Project.org/package=frontier>.

- 530 [25] Battese GE, Coelli TJ. A model for technical inefficiency effects in a stochastic
531 frontier production function for panel data. *Empirical economics*. 1995 Jun
532 1;20(2):325-32.
- 533 [26] Bogetoft P, Otto L. *Benchmarking with DEA, SFA, and R*. Springer Science &
534 Business Media; 2010 Nov 19.
- 535 [27] Schmidt P, Sickles RC. Production frontiers and panel data. *Journal of Business &*
536 *Economic Statistics*. 1984 Oct 1;2(4):367-74.
- 537 [28] Battese GE. Frontier production functions and technical efficiency: a survey of
538 empirical applications in agricultural economics. *Agricultural economics*. 1992
539 Oct;7(3-4):185-208.
- 540 [29] Battese GE, Coelli TJ. *Op. cit.*
- 541 [30] Engle RF. Interpreting spectral analyses in terms of time-domain models. In *Annals*
542 *of Economic and Social Measurement*, Volume 5, number 1 1976 Jan 1 (pp. 89-
543 109). NBER.
- 544 [31] SciPy.org. [Internet]; 2017 [cited 2020 Dec 26]. Signal processing module, version
545 0.12.0. Available at:
546 [https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.periodogram.ht](https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.periodogram.html)
547 [ml](https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.periodogram.html)
- 548 [32] Sargent TJ, Stachurski J. [Internet]; 2014 [cited 2020 Dec 26]. Quantitative
549 Economics. Python package version 0.3.4. Available at:
550 <https://lectures.quantecon.org>.
- 551 [33] Aigner D, Lovell CK, Schmidt P. Formulation and estimation of stochastic frontier
552 production function models. *Journal of econometrics*. 1977 Jul 1;6(1):21-37.

[34] Meeusen W, van den Broeck J. Technical efficiency and dimension of the firm: Some results on the use of frontier production functions. Empirical economics. 1977 Jun 1;2(2):109-22.

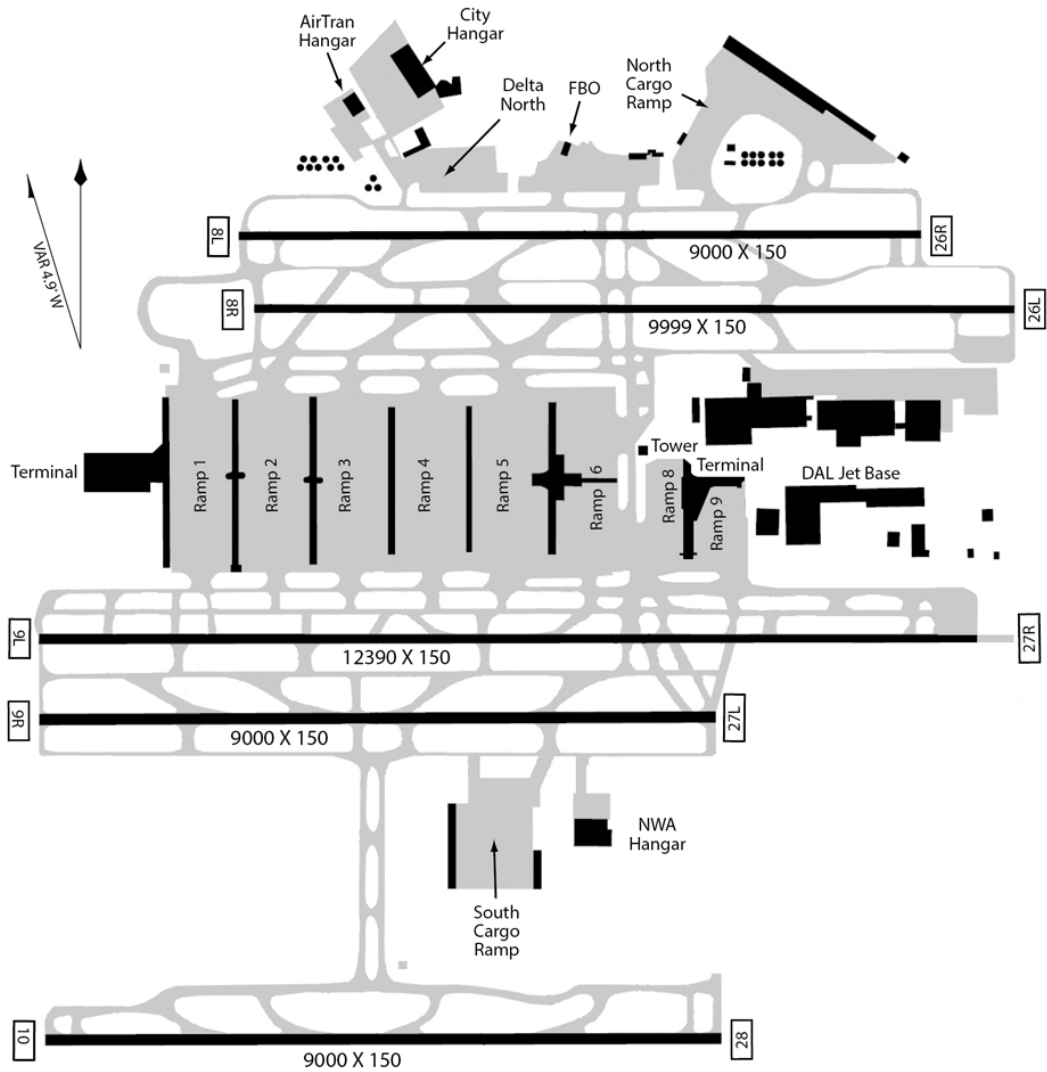
[35] Bartlett MS. An introduction to stochastic processes. Cambridge: Cambridge University Press; 1966.

[36] Durbin J, Brown R. Tests of serial independence based on the cumulated periodogram. Bulletin of the International Statistical Institute. 1967;42:1039-48.

[37] Fuller WA. Introduction to statistical time series. New York: John Wiley & Sons; 1976.

APPENDIX

ATL Runway Configuration



For illustration purposes only.

Source: Federal Aviation Administration