Measuring Performance at a Large Metropolitan Area: The Case of the DC (District of Columbia) Metroplex

by Tony Diana

Hierarchical linear models improve the measurement of performance when applied to a construct such as a metroplex. It compares the outcomes of a hierarchical linear model with those of a multiple regression model to evaluate whether meteorological conditions at individual airports and overall would explain variations in block delays. The study used the cases of the three largest airports in the DC Metroplex and concluded airborne delays had a significant random effect on block delays in spite of meteorological conditions at each airport. It pointed out that surface operations efficiency played a significant role in explaining variations in block delays.

INTRODUCTION

The National Airspace System (NAS) consists of a network of airports that serve the needs of a variety of users (i.e. air carrier, air taxi, general aviation, military, and freight operators). These airports are sometimes clustered within large metropolitan areas, also called "metroplexes," where a mix of users and aircraft sizes raises some challenges for air traffic control (ATC) as they share the same airspace.

A metroplex can be defined as a metropolitan area where access among larger hub and smaller general aviation airports in close proximity may be affected by interdependent and sometimes conflicting arrival and departure routes. The appendix provides an illustration of the arrival and departure streams at the DC Metroplex. The metroplex concept holds a central place in the Next Generation Air Transportation System, or "NextGen," that aims to transform the legacy radar-based air traffic control into the future satellite-based air-traffic-managed system. In 2009, the RTCA (Radio Technical Commission for Aeronautics)[1] Task Force 5 identified 21 areas "to optimize area navigation (RNAV) and required navigation procedures (RNP) operations, and institute tiger teams that focus on quality at each location as well as integrate procedure design to de-conflict airports and expand use of terminal separation rules."[2] Such is the case in this study of the District of Columbia Metroplex, or DC Metroplex, that includes (1) large hub airports such as Baltimore/Washington International Thurgood Marshall (BWI), Washington Reagan National (DCA), and Washington Dulles International (IAD) airports; (2) secondary airports such as Richmond International (RIC); and (3) general aviation airports such as Frederick Municipal (FDK), Martin State (MTN), and Manassas Regional (HEF) airports, among others.

The Federal Aviation Administration (FAA) initiated the airspace redesign project of the DC Metroplex in September 2010. After proceeding through the phases of study, design, environment, and safety management system, the implementation of airspace redesign started in July 2013. In the meantime, new arrival and departure procedures were implemented to support area navigation (RNAV) and optimized profile descent (OPD), as well as to minimize conflicting flight paths. Optimized profile descents allow aircraft to stay longer in level flight and to descend progressively from the top of descent to the runway threshold. OPDs serve two major purposes: (1) Reduce stepwise descent that increases fuel consumption and (2) minimize the need for frequent read-back communication between pilots and air traffic controllers.

The present analysis focuses on the cases of the three largest airports in the DC Metroplex (BWI, DCA, and IAD) at three different time periods: before the implementation of NextGen capabilities and new flight procedures (summer 2007) and afterward (summer 2012 and 2013). The summer months of June to August were selected because they are characterized by peak traffic and extreme weather such as thunderstorms. This study assumes that variations in block delays are likely to be impacted by en-route miles flown, aircraft speed, the counts of NAS-related delays, taxi-in and -out delays, as well as the percentage of operations in instrument meteorological conditions. The study also hypothesize that meteorological conditions at each airport and overall for the metroplex are likely to affect block delays when considering the random effects of airborne delays. This study attempts to answer the following research questions:

- How much of the variation in block delays (the difference between actual and filed block times) can be attributed to meteorological conditions during the hours of 07:00 to 21:59 for each sampled time period?
- Is the influence of any independent variables on block delays more likely to vary under specific meteorological conditions (instrument vs. visual) overall and at each individual airport for each sampled summer?
- Is there any significant change in block delays at the Metroplex and individual airport levels as new procedures and NextGen capabilities have been implemented?
- Is there any difference between the outcomes of a hierarchical linear model based on a mixed model and those of a multiple regression model?

To answer these questions, a two-level hierarchical linear model utilizing fixed and random effects was selected. A hierarchical linear model implies that meteorological conditions are nested within each sampled airport. Hierarchical linear or multilevel analysis[3] models have been widely used in spatial analysis, sociology, and psychology, but not extensively in aviation. This study serves several purposes: (1) it illustrates how hierarchical analysis provides some better insight into the factors that explain block delays; (2) it takes into account multiple levels of measurement that would otherwise be “hidden” in an ordinary least-squares model; and (3) it includes fixed-effects and random-effects variables.

Beaubien et al. (2001) provided a review of hierarchical linear modeling techniques applied to commercial aviation research with pilot, crew, and airlines as the three levels of analysis. Haines et al. (2002) determined that chronic exposure to aircraft noise was likely to be associated with poor school performance in reading and mathematics performance. Castelli et al. (2003) resorted to multilevel analysis to evaluate the patterns of variation of price elasticity of demand among the various routes of an airline and concluded that the airfare elasticity of passenger demand significantly varied among the different routes of the airline. Gudmunsson (2004) evaluated the factors associated with airline performance using a two-level bottom–up hierarchical approach. Miranda et al. (2011) used multilevel analysis to investigate the blood lead levels among children living in six North Carolina counties resulting from the exposure to aviation gasoline exhaust. Chung and Wong (2011) investigated the impact of China-Taiwan non-stop routes on cross-strait air travel city pairs. Fidell et al. (2011) utilized multilevel models to assess the impact of annoyance with aircraft noise exposure across communities. Rozenblat et al. (2013) made use of multilevel clustering methods to delineate the effects of geographical distance, hubs, network densities, and bridges on worldwide air passenger traffic.

The next section will deal with the methodology and assumptions underlying the analytical models before discussing the model outcomes and providing some final remarks and implications for future research.

METHODOLOGY

The Models' Variables and Data Processing

The sample includes 828 observations equally divided into three summers (92 days each) and three airports. All variables originate from the Aviation System Performance Metrics (ASPM) data warehouse. ASPM reports operational and delay metrics from a variety of sources: OPSNET (Operational Network), Traffic Flow Management System (TFMS), and the Bureau of Transportation Statistics (BTS). Within each day, the selected variables were measured from the hours of 07:00 to 21:59 when traffic is most active. Each hour was flagged for instrument versus visual meteorological conditions. Since weather is the major driver of delays, this study assumed that it was not possible to measure variations in block delays as explained by the model without isolating the impact of meteorological conditions as a whole and for each selected airport in the metroplex.

The number of variables was determined by comparing Akaike's Information Criterion (AIC), Akaike's Information Criterion corrected for finite sample sizes (AICC) and Schwartz's Bayesian Information Criterion (BIC) of the various models. The lower the number of estimated parameters, the lower the value of the AIC and BIC and, as a result, the better the fit of the model. Other considered variables were the number of operations, gate arrival and departure delays, and excess miles flown. The final models included the following variables:

- *Block delays* represent the dependent variable. They are computed as the difference between actual and block times filed in the flight plan, in minutes. The flight plan used to compute block delays is the last one before takeoff. Block time measures the duration from gate-out (origin) and gate-in (destination) times. Block delays include all the flights that arrive at BWI, DCA, and IAD during the hours of 07:00 to 21:59 (local time). The comparison of actual with flight-planned times removes from the analysis any biases due to airlines' schedule padding if actual block times had been compared with scheduled block times. However, it is important to point out that the number of scheduled operations for the combined airports declined from 214,306 in summer 2007 to 207,127 in summer 2012 and 204,894 in summer 2013 (sources: Innovata flight schedules).
- *En-route miles flown* represent the distance about 100 nautical miles from the origin airport to 100 nautical miles from the destination airport.
- *Speed* is the average number of nautical miles flown per hour for aircraft flying into BWI, DCA, and IAD.
- *The counts of National Airspace System or NAS-related delays* account for "the delays and cancellations attributable to the national aviation system that refer to a broad set of conditions, such as non-extreme weather conditions, airport operations, heavy traffic volume, and air traffic control."[4] These delays often impact changes in the flight plan.
- *Taxi-out delays* are the differences between actual and unimpeded gate-out to wheels-off times, in minutes. Unimpeded taxi-out times are based on taxi-out times reported by the major carriers to BTS, and they estimate the time it takes for an aircraft to move from the gate departure to takeoff when there is only one aircraft ahead in the takeoff queue. The gate-out, wheels-off, wheels-in, and gate-in messages used to compute the actual times of the flight phases are compiled by ARINC (Aeronautical Radio, Incorporated), a division of Rockwell Collins (http://www.airinc.com) and recorded in the ASPM data warehouse.
- *Taxi-in delays* are the differences between actual and unimpeded wheels-on to gate-in times, in minutes. The computation of taxi-in delays are based on the same principles as taxi-out delays.
- *Airborne delays* measure the difference between actual airborne times and the flight plan's estimated time en route, in minutes. Airborne delays are sometimes used by ATC to provide safe separation, regulate speed, merge traffic, and avoid potential conflicts among flight paths.

Airborne delays characterize the random effect variable in the model, which implies setting up a common correlation among all observations having the same level of airborne delays.

- *The percentage of operations in instrument meteorological conditions* is derived from the minimum ceiling and visibility in effect at each facility summarized in Table 1. If the percentage of operations in instrument meteorological conditions during a specific hour is greater than 10%, then the dummy variable "IMC" for that flight record is coded as "1" and "0" otherwise.

Table 1: Selected Airports' Minimum Ceiling and Visibility

Airpot	Ceiling (ft)	Visibility (nm)
BWI	2,500	5
DCA	3,000	4
IAD	3,000	7

Source: FAA, ASPM

The hierarchical linear model estimates were derived with the MIXED procedure, while the multiple regression models used the REG procedure, both programmed in SAS®. Meteorological condition and airport are the classification variables that represent the two levels in this analysis. The hierarchical linear model estimates were generated with maximum likelihood. Fifteen iterations were required for convergence.

This study utilizes a mixed model that includes fixed-effects parameters (known explanatory variables) and covariance parameters that are useful when data are grouped into clusters (i.e., individual airport and meteorological conditions at selected airports) in which data are likely to be correlated. The clustering (nesting) of meteorological conditions into the airport variable creates additional potential variability and correlation. Although data are assumed to have a normal distribution, the mixed model allows correlation and heterogeneous variances.

The Hierarchical Model Assumptions and Specifications

An ordinary least-squares (OLS) regression model does not provide any indication of how the selected factors account for variations in block delays when data are sliced at different levels (i.e. by meteorological condition and by meteorological condition at each sampled airport). A hierarchical linear model enables analysts to account for the variations of block delays and to understand the contribution of meteorological conditions at each sampled airport to explain the variation in block delays. Readers interested in hierarchical linear or multilevel analytical models are referred to Bryk and Raudenbush (2001) and Hox (2010) for a clear exposition.

Hox (2010: 4) summarized the purpose of multilevel analysis in these terms: "The goal of the analysis is to determine the direct effect of individual- and group-level explanatory variables, and to determine if the explanatory variables at the group level serve as moderators of individual-level relationships." Hierarchical linear models provide the following benefits:

- They improve the estimation of effects within individual airports.
- They allow analysts to test the hypotheses of cross-level effects and the partitioning of variance and variance components among the two levels.
- One of the key assumptions of OLS is independence of observations. However, nesting meteorological conditions into airports creates dependencies in the data and may generate inaccurate estimates in non-hierarchical linear models. The dependence of observations may entail biased parameter estimates and standard errors in OLS models.
- It is important for analysts to understand variations at different hierarchical levels. Hierarchical linear analysis takes into account the correlated nesting of data, whether block delays vary based on meteorological conditions at a specific airport during a specific time period.

The hierarchical linear model that features a random intercept and slopes for each time period features two levels:

Level 1 for Each Selected Time Period:

(1) $Y_{ij} = \beta_{0j} + \beta_{1j} X_{1j} + \ldots + \beta_{ij} X_{ij} + \varepsilon_{ij}$

Where $\varepsilon_{ij} \sim N(0, \sigma^2)$, Y_{ij} represents block delays and X_{ij}, the factors in specific meteorological conditions nested at j airport that explain variations in block delays.

Level 2 for Each Selected Time Period:

(2) $\beta_{0j} = \Upsilon_{00} + U_{0j}$

(3) $\beta_{1j} = \Upsilon_{10} + U_{1j}$

Where $\begin{bmatrix} U_{0j} \\ U_{1J} \end{bmatrix} = U_j \sim N\left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_0^2 & \tau_{10} \\ \tau_{10} & \tau_1^2 \end{bmatrix}\right]$ independent over j and with ε_{ij}

τ_0^2 represents the variance of the level two residuals U_{0j} from predicting the level 1 intercept (β_{0j}). τ_1^2 is the variance of the level 2 residuals U_{1j} from predicting the level 1 slope (β_{1j}). τ_{10} is the covariance between U_{0j} and U_{1j}. The $cov(U_{0j}, U_{1j}) = cov(\beta_{0j}, \beta_{1j}) = \tau_{10}$.

Based on the equations (1) to (3), the hierarchical linear model can be expressed as:

(4) $Y_{ij} = \Upsilon_{00} + \Upsilon_{10} X_{ij} + U_{0j} + U_{1j} + \varepsilon_{ij}$

Where $\Upsilon_{00} + \Upsilon_{10}X_{ij}$ represents the fixed component and $U_{0j} + U_{1J} + \varepsilon_{ij}$, the random one. The variance of the random-effects parameters are known as variance components.

The specification of fixed and random components within the hierarchical linear model represents the major difference with the multiple regression models for each airport. The empirical model is equation (5):

(5) Block Delays$_{ij}$ = β_{0j} + β_{1j}Enroute Miles Flown$_{ij}$ + β_{2j}Speed$_{ij}$+ β_{3j} NAS Delays$_{ij}$ + β_{4j}Taxi-Out Delays$_{ij}$+ β_{5j}Taxi-In Delays$_{ij}$ + β_{6j}IMC$_{ij}$ + ε_{ij} for j airport.

FINDINGS AND INTERPRETATION

Goodness of Fit

The first step in the analysis consists in evaluating the goodness of fit of the hierarchical and multiple regression models.

In Table 2, the improvement in the -2 Log likelihood value over the iterations for each sample suggests there is "a significant improvement over the null model consisting of no random effects and a homogeneous residual error."[5]

Table 2: Fit Statistics for the Hierarchical Models

Fit Statistics			
Criteria	**Summer ’07**	**Summer ’12**	**Summer ’13**
-2 Log Likelihood	828.0	673.9	701.3
AIC (Smaller is Better)	850.0	693.9	721.3
AICC (Smaller is Better)	851.0	694.7	722.2
BIC (Smaller is Better)	835.6	680.8	708.3

Table 3: Fit Statistics for the Multiple Regression Models

Fit Statistics			
Criteria	**Summer ’07**	**Summer ’12**	**Summer ’13**
R^2	0.8660	0.7665	0.8408
Adjusted R^2	0.8630	0.7613	0.8373
F Value	289.63	147.16	236.83
Pr > F	<.0001	<.0001	<.0001

At a 95% confidence level, the hypothesis that any estimate in the three models equals zero can be rejected since $p < .0001$. The coefficients of determination indicate that the selected independent variables account for a large portion of the variation in block delays, albeit to a greater extent in summer 2007 and 2013 than in summer 2012. However, the multiple regression models do not specify (1) to what extent overall meteorological conditions and those at each selected airport may have accounted for any variation in block delays and (2) whether airborne delays linked to traffic management initiatives may have randomly affected block delays.

The Estimates of Fixed and Random Effects in the Hierarchical Linear Models

In Table 4, the fixed-effects estimates represent the estimated means for the random intercept and slope, respectively.

Most of the fixed effects are significant at a 95% confidence level except enroute miles flown in summer 2012 ($p > 0.05$). Taking the example of summer 2007, block delays decreased 0.0029 minutes on average for one nautical mile change in the enroute miles flown, while holding other predictors in the model constant. In summer 2012 and 2013, the sign of the estimates for enroute miles flown did not significantly change and block delays increased 0.0015 and 0.0037 minutes, respectively, for one nautical mile flown, holding other variables constant.

Although speed was significant in the three samples, its magnitude decreased after the implementation of new procedures (optimized profile descent, area navigation approaches and departures) and airspace redesign.[6] In summer 2007, block delays decreased 0.04 minutes for one nautical mile change in speed, compared with 0.0231 minutes in summer 2012 and 0.0198 minutes in summer 2013, holding other factors constant.

It is important to note that surface operation delays have the greatest impact on the variation of block delays given the magnitude of the estimates for taxi-out and taxi-in delays. In the case of the DC Metroplex, this can be explained by the lack of available gates at peak times in summer 2007 and ramp congestion prior to runway and terminal building enhancements at the three airports that were under way in summer 2012 and 2013. This explains why block delays increased 1.28 minutes on average for a one-minute change in taxi-in time in summer 2007 compared with 0.68 and 0.84 minutes, respectively, in summer 2012 and 2013, holding all factors constant. As for taxi-out operations, the implementation of tarmac delays rules[7] in April 2010 induced airlines to defer departure or cancel flight departures at times of airport congestion or poor weather conditions in order to avoid hefty penalties ($27,500 per passenger).

Table 4: Solutions for Fixed Effects

	Summer 2007					Summer 2012					Summer 2013				
Effect	Estimate	Standard Error	DF	t value	Pr > |t|	Estimate	Standard Error	DF	t value	Pr > |t|	Estimate	Standard Error	DF	t value	Pr > |t|
Intercept	17.2138	2.9363	274	5.86	<.0001	8.3413	1.8992	274	4.39	<.0001	4.8150	1.3960	235	3.45	0.0007
Enroute Miles Flown	-0.0029	0.0009	276	-3.33	0.001	0.0015	0.0007	38.4	1.96	0.0569	0.0037	0.0005	3.89	7.51	0.0019
Speed	-0.0403	0.0065	274	-6.16	<.0001	-0.0231	0.0043	270	-5.36	<.0001	-0.0198	0.0032	269	-6.23	<.0001
NAS Delays	0.0820	0.0064	274	12.81	<.0001	0.0388	0.0042	270	9.31	<.0001	0.0313	0.0049	275	6.43	<.0001
Taxi-Out Delays	0.1652	0.0199	274	8.29	<.0001	0.1172	0.0230	274	5.09	<.0001	0.1194	0.0195	272	6.12	<.0001
Taxi-In Delays	1.2784	0.0717	275	17.82	<.0001	0.6793	0.0691	250	9.83	<.0001	0.8380	0.8380	22.8	19.17	<.0001

Table 5: Solutions for Random Effects

	Levels		Summer 2007					Summer 2012					Summer 2013				
Effect	apt	imc	Estimate	Standard Error	DF	t value	Pr > |t|	Estimate	Standard Error	DF	t value	Pr > |t|	Estimate	Standard Error	DF	t value	Pr > |t|
Intercept		0	0	.	.	.	.	0	.	.	.	.	0	.	.	.	.
Airborne Delays		0	0.5319	0.06675	266	7.97	<.0001	0.7526	0.09407	10	8	<.0001	0.9826	0.06987	128	14.06	<.0001
Intercept	BWI	0	0	.	.	.	.	0	.	.	.	.	0.05113	0.07158	1	0.71	0.6052
Intercept	DCA	0	0	.	.	.	.	0	.	.	.	.	-0.0626	0.07689	1	-0.81	0.565
Intercept	IAD	0	0	.	.	.	.	0	.	.	.	.	-0.01064	0.07697	1	-0.14	0..9125
Airborne Delays	BWI	0	0	.	.	.	.	0.0669	0.0773	4.31	0.87	0.4323	0	.	.	.	.
Airborne Delays	DCA	0	0	.	.	.	.	-0.09182	0.08018	4.58	-1.15	0.3084	0	.	.	.	.
Airborne Delays	IAD	0	0	.	.	.	.	0.04451	0.0863	5.14	0.52	0.6274	0	.	.	.	.
Intercept		1	0	.	.	.	.	0	.	.	.	.	0	.	.	.	.
Airborne Delays		1	0.5324	0.05308	268	10.03	<.0001	0.7593	0.08628	7.1	8.8	<.0001	0.9226	0.05629	120	16.39	<.0001
Intercept	BWI	1	0	.	.	.	.	0	.	.	.	.	0.07898	0.07607	1	1.04	0.4881
Intercept	DCA	1	0	.	.	.	.	0	.	.	.	.	-0.02519	0.0786	1	-0.32	0.8025
Intercept	IAD	1	0	.	.	.	.	0	.	.	.	.	-0.03168	0.07747	1	-0.41	0.7529
Airborne Delays	BWI	1	0	.	.	.	.	0.163	0.07816	4.47	2.08	0.0982	0	.	.	.	.
Airborne Delays	CDA	1	0	.	.	.	.	-0.1197	0.07935	4.52	-1.51	0.1978	0	.	.	.	.
Airborne Delays	IAD	1	0	.	.	.	.	-0.02346	0.085	5.03	-0.28	0.7935	0	.	.	.	.

In Table 5, the random-effects coefficients represent the estimated deviation from the mean intercept and slope for IMC and each airport in IMC. The assumption is that airborne delays exhibit more correlation with the other factors at specific airports and meteorological conditions. According to OPSNET data, the number of traffic management initiatives (TMI) that includes miles-in-trail and minutes-in-trail, airborne holding was higher at the three airports in summer 2012 (1,267) than in summer 2007 (1,154) and summer 2013 (1,038). Minutes-in-trail describes the number of minutes while miles-in-trail describes the number of miles required between aircraft departing an airport, over a fix (a point in space that guides aircraft along a flight path), through a sector, or on a specific route. Therefore, additional spacing in time or distance was likely to have an impact of block delays, even though the enroute miles flown may have not varied drastically. That may explain why the variable "enroute miles flown" was not significant at a 95% confidence level in summer 2012.

Table 5 shows that the random effects of airborne delays are significant at a 95% confidence level whether in IMC or VMC. Regardless of the type of meteorological conditions, only the summer 2012 random-effects estimates of airborne delays at each airport were different from zero. However, the summer 2012 random-effects estimates were not significant at a 95% confidence level. It is also important to point out that the magnitude of the random-effects estimates of airborne delays in summer 2007 and 2012 were not significantly different in both meteorological conditions. In summer 2013, the magnitude of airborne delays was greater in VMC than IMC. As operations were on the rise, traffic management initiatives associated with Time-Based Flow Management between the ZDC (Washington Air Route Traffic Control Center) and adjacent centers may have increased the incidence of airborne delays to regulate the flow of traffic into the three airports.

In Table 6, the F statistic in the "Type 3 Tests of Fixed Effects" is the square of the t statistic used in the test of the independent variables. Both statistics test the null hypothesis that the slope assigned to the dependent variables equals 0.

Table 6: Type 3 Tests of Fixed Effects

	Summer 2007			Summer 2012			Summer 2013		
Effect	**Den DF**	**F Value**	**Pr > F**	**Den DF**	**F Value**	**Pr > F**	**Den DF**	**F Value**	**Pr > F**
Enroute Miles Flown	276	11.07	0.001	38.4	3.86	0.0569	3.89	56.41	0.0019
Speed	274	37.95	<.0001	270	28.74	<.0001	269	38.86	<.0001
NAS Delays	274	164.14	<.0001	270	86.61	<.0001	275	41.32	<.0001
Taxi-Out Delays	274	68.75	<.0001	274	25.94	<.0001	272	37.47	<.0001
Taxi-In Delays	275	317.52	<.0001	250	96.66	<.0001	22.8	367.56	<.0001

The slopes in Table 6 are the same as those in Table 4. The significant level ($p<0.0001$) indicates that there is evidence the slopes are not equal to zero and, therefore, significant at a 95% confidence level.

Multiple Regression Outputs

Table 7 shows the multiple regression estimates. The shaded cells highlight the factors that are not significant at a 95% confidence level. Compared with the fixed-effects estimates in Table 4, only miles flown and speed were not significant at a 95% confidence level in summer 2012 (Table 7). While IMC is significant overall during the three sampled time periods, we do not know if there is any difference by airport.

Table 7: The Multiple Regression Estimates

Variable	Summer 2007 Parameter Estimate	Summer 2007 Standard Error	Summer 2007 t Value	Summer 2007 Pr > |t|	Summer 2012 Parameter Estimate	Summer 2012 Standard Error	Summer 2012 t Value	Summer 2012 Pr > |t|	Summer 2013 Parameter Estimate	Summer 2013 Standard Error	Summer 2013 t Value	Summer 2013 Pr > |t|
Intercept	17.5050	3.4044	5.14	0.001	2.0080	2.5373	0.79	0.4294	3.4100	1.9137	1.78	0.0759
enroute_mls_flown	0.0022	0.0008	2.84	0.0049	-0.0006	0.0005	-1.09	0.2749	0.0007	0.0005	1.42	0.1567
speed	-0.0446	0.0076	-5.88	<.0001	-0.0052	0.0057	-0.91	0.3634	-0.0108	0.0044	-2.47	0.0143
NAS_del	0.0942	0.0073	13	<.0001	0.0621	0.0053	11.74	<.0001	0.0669	0.0059	11.33	<.0001
txout_del	0.1684	0.0232	7.27	<.0001	0.1732	0.0291	5.94	<.0001	0.2526	0.0244	10.37	<.0001
txin_del	1.4436	0.0804	17.95	<.0001	0.8826	0.0733	12.04	<.0001	0.8115	0.0548	14.8	<.0001
imc	0.5466	0.1500	3.64	0.0003	0.5619	0.1246	4.51	<.0001	0.5108	0.1158	4.41	<.0001

Not significant at 95% confidence level.

FINAL REMARKS AND IMPLICATIONS

Hierarchical linear models, including fixed and random effects, can provide a better picture of performance at a construct such as the metroplex. Metroplexes play a significant role in the Next Generation Air Transportation System: They represent large metropolitan areas where the close proximity of airports is likely to create conflicting approaches and departure paths, thus reducing access potential to larger and general aviation airports. This paper used the example of the District of Columbia Metroplex where new procedures have been implemented.

At the metroplex level, enroute miles flown, speed, the number of NAS-related delays, and taxi-out and taxi-in delays have significant fixed effects on the variation of actual block times when compared with airlines' flight plans. The study indicates that instrument meteorological conditions have a significant impact overall.

As NextGen capabilities and procedures are deployed into the NAS, it will be of interest for aviation practitioners to assess whether the greater utilization of satellite navigation, as well as the implementation of performance-based navigation, will have an impact on the variation of block delays in metroplexes. While pilots' flexibility to choose trajectories and data-sharing in the cockpit are important to reduce excess miles flown, the study also suggests that attention should also be paid to surface operations' efficiency in the forms of taxi-in and taxi-out times to reduce block delays.

Endnotes

1. Created in 1935, RTCA is an organization of aviation experts and practitioners working to improve flight performance standards.

2. RTCA, NextGen Mid-Term Implementation Task Force Report, September 9, 2009, p. xiii. The document was retrieved in September 2013 at the following website: http://www.faa.gov/nextgen/media/nextgen_progress_report.pdf.

3. Multilevel analysis is also called "random coefficient model" (de Leeuw and Kreft, 1986; Longford 1993), variance component model (Longford, 1987), hierarchical linear model (Raudenbush and Bryk, 1986 and 1988), as well as mixed effects model (Littell et al. 1996). Also refer to Hox (2002:11).

4. The definitions of the causes of delay were retrieved at the website of the U.S. Department of Transportation, Bureau of Transportation Statistics, whose link is http://www.rita.dot.gov/bts/help/aviation/html/understanding.html#q4.

5. See SAS/Stats® 9.2, User's Guide, Second Edition, retrieved at http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_mixed_sect034.htm.

6. Federal Aviation Administration, NextGen Performance Snapshots, "Honoring the Past While Flying into the Future," retrieved at http://www.faa.gov/nextgen/snapshots/stories/?slide=16.

7. See 14 Code of Federal Regulation (CFR) 259.4 for the tarmac delay contingency plans.

References

Beaubien, J. Matthew, William R. Hamman, Robert W. Holt, and Deborah A. Boehm-Davis. "The Application of Hierarchical Linear Modeling (HLM) Techniques in Commercial Aviation Research." *Proceedings of the 11th Annual Symposium on Aviation Psychology*. The Ohio State University Press, 2001.

Bryk, Anthony S. and Stephen W. Raudenbush. "Toward a More Appropriate Conceptualization of Research On School Effects: A Three-Level Hierarchical Linear Model." *American Journal of Education* (1988): 65-108.

Bryk, Anthony S. and Stephen W. Raudenbush. *Hierarchical Linear Models, Applications and Data Analysis Methods, Second Edition*. Sage Publications, Thousand Oaks, CA, 2001.

Castelli, Lorenzo, Raffaele Pesenti, and Walter Ukovich. "An Airline-Based Multilevel Analysis of Airfare Elasticity for Passenger Demand." Brent Bowen, Sveinn Gudmunsonn, and Tae Oum, eds. *The Conference Proceedings of the 2003 Air Transportation Research Society (ATRS) World Conference, Volume 27, University of Nebraska Omaha*.

Chung, Yi-Shi and Jinn-Tsai Wong. "Exploring the Impact of Introducing Scheduled Flights on Aviation Markets." *Transportation Research Record: Journal of the Transportation Research Board* 2214, (2011): 41-49.

De Leeuw, Jan and Ita Kreft. "Random Coefficient Models for Multilevel Analysis." *Journal of Educational and Behavioral Statistics* 11(1), (1986): 57-85.

Federal Aviation Administration, Aviation System Performance Metrics (ASPM), https://aspm.faa.gov.

Fidell, Sanford, Vincent Mestre, Paul Schomer, Bernard Berry, Truls Gjestland, Michel Vallet, and Timothy Reid. "A First-Principles Model for Estimating the Prevalence of Annoyance with Aircraft Noise Exposure." *The Journal of the Acoustical Society of America* 130, (2011): 791-806.

Gudmundsson, Sveinn. "Management Emphasis and Performance in the Airline Industry: An Exploratory Multilevel Analysis." *Transportation Research Part E: Logistics and Transportation Review* 40(6), (2004): 443-463.

Haines, M. M., S. A. Stansfeld, J. Head, and R. F. S. Job. "Multilevel Modeling of Aircraft Noise on Performance Tests in Schools Around Heathrow Airport in London." *Journal of Epidemiology and Community Health* 56(2), (2002): 139-144.

Hox, Joop. *Multilevel Analysis: Techniques and Applications, Second Edition*. Routledge, New York, NY, 2010.

Innovata Flight Schedules, http://www.innovata-llc.com.

Littell, R. C., G. A. Milliken, W. W. Stroup, and R. D. Wolfinger. *SAS System for Mixed Models*. SAS Institute, Cary, NC, 1996.

Longford, Nicholas T. *Random Coefficient Models*. Clarendon Press, New York, NY, 1993.

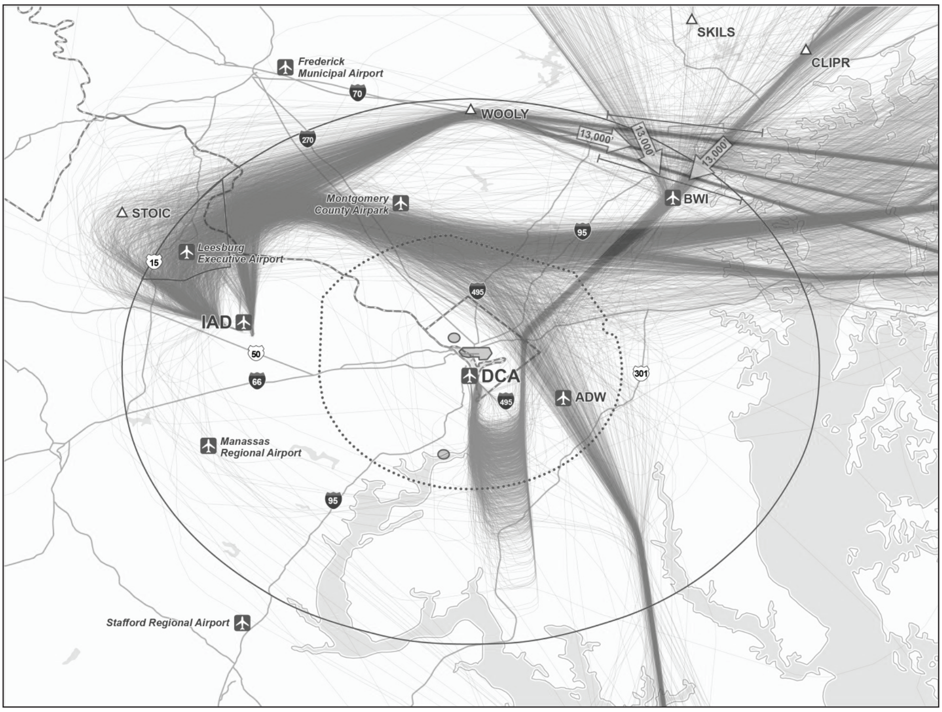
Longford, Nicholas T. "A Fast Scoring Algorithm for Maximum Likelihood Estimation in Unbalanced Mixed Models with Nested Random Effects." *Biometrika* 74(4), (1987): 817-827.

Miranda, Marie Lynn, Rebecca Anthopolos, and Douglas Hastings. "A Geospatial Analysis of the Effects of Aviation Gasoline on Childhood Blood Lead Levels." *Environmental Health Perspectives* 119(10), (2011): 1513-1516.

Raudenbush, Stephen and Anthony S. Bryk. "A Hierarchical Model for Studying School Effects." *Sociology of Education* (1986): 1-17.

Rozenblat, Céline, Guy Melançon, Romain Bourqui, and David Auber. "Comparing Multilevel Clustering Methods on Weighted Graphs: The Case of Worldwide Air Passenger Traffic 2000–2004." Céline Rozenblat, Guy Melançon eds. *Methods for Multilevel Analysis and Visualisation of Geographical Networks*. Springer, Dordrecht, Netherlands, 2013.

APPENDIX: Arrival and Departure Flows at the DC Metroplex



source: ATAC

ACRONYMS

AIC	Akaike Information Criterion
AICC	Akaike Information Criterion corrected for finite sample sizes
ASPM	Aviation System Performance Metrics
ATC	Air Traffic Control
BIC	Bayesian Information Criterion
BTS	Bureau of Transportation Statistics
BWI	Baltimore/Washington International Thurgood Marshall Airport
DCA	Washington Reagan National Airport
FAA	Federal Aviation Administration
IAD	Washington Dulles International Airport
IMC	Instrument Meteorological Conditions
NAS	National Airspace System
NextGen	Next Generation Air Transportation System
OPD	Optimized Profile Descent
OPSNET	Operations Network
PBN	Performance-Based Navigation
RNAV	Area Navigation
RNP	Required Navigation Procedure
RTCA	Radio Technical Commission for Aeronautics
TBFM	Time-Based Flow Management
TFMS	Traffic Flow Management System
VMC	Visual Meteorological Conditions
ZDC	Washington Air Route Traffic Control Center

Tony Diana *is the division manager, NextGen Performance at the U.S. Federal Aviation Administration. He received his doctorate in policy analysis and quantitative management from the University of Maryland Baltimore County (UMBC). Diana is involved in the measurement of operational outcomes resulting from the implementation of NextGen initiatives at U.S. airports, metroplexes and airspaces. Prior to that position, he was deputy division manager, forecasting and performance analysis in the Office of Aviation Policy and Plans of the FAA where he managed the Aviation System Performance Metrics data warehouse. At the Maryland Aviation Administration, Diana was involved in performance measurement and route development. His main interests are performance evaluation and benchmarking, as well as the study of delay.*

NOTE: This research does not reflect the official opinion of the Federal Aviation Administration.