Understanding the Capacity of Airport Runways

Kailin Chen*[1], Daniel J. Graham[1], Richard J. Anderson[1], Anupriya[1], and Prateek Bansal[2]

[1]Transport Strategy Centre, Department of Civil and Environmental Engineering, Imperial College London, United Kingdom
[2]Department of Civil and Environmental Engineering, National University of Singapore, Singapore

Short summary

Understanding the capacity of runway system under different operational conditions is of critical importance to airport operators and planners. The availability of granular data on day-to-day runway operations facilitates the development of models that allow a precise comprehension of runway capacity. However, the exercise is empirically challenging due to statistical biases that emerge via the complex interactions between air traffic control and runway capacity. This paper develops a novel causal statistical framework based on a confounding-adjusted Stochastic Frontier Analysis (SFA) to deliver estimates of runway capacity and its parameters that are robust to such biases. The model captures the key factors and interactions affecting runway capacity in a computationally intensive manner. The performance of the model is demonstrated via benchmarking of the estimated capacities of three major airports around the world.

Keywords: Airport operations; Runway capacity; Empirical estimation; Confounding; Causal statistical modelling; Stochastic frontier analysis.

1 Introduction

Runway capacity, the maximum number of aircraft movements that a runway system in an airport can operate in a given time period, is a primary input to air traffic management and planning (De Neufville et al., 2013). Knowledge of runway capacity supports decisions in planning and operations, including (1) distributing the daily runway demand over the available runway capacity (that is, available slots) (Gilbo, 1993; Cheung et al., 2021), (2) modeling airport delay or even delay propagation within the airport network (Pyrgiotis et al., 2013), and (3) appraising investments in runway capacity expansion (Hansen, 2004), among others. By nature, runway capacity is highly dynamic because it is determined by various time-varying operational factors such as weather conditions, runway configuration, and fleet mix (Ashford et al., 2011). However, even after many endeavors to estimate runway capacity, the literature lacks methods that can robustly quantify its dynamic nature, while being less resource-intensive in terms of time, labour, data, and other monetary requirements (such as expenses for a software license). This study attempts to address this gap by developing a model to precisely estimate runway capacity and the parameters of its associated factors.

Previous models of assessing runway capacity can be grouped into four main categories: (1) table lookup and spreadsheet (FAA, 1983; TRB, 2012), (2) analytical (Blumstein, 1959; Cheung et al., 2017; Mascio et al., 2020), (3) simulation (Bubalo & Daduna, 2011; Kuzminski, 2013; Barrer et al., 2005), and (4) empirical (Gilbo, 1993; Hansen, 2004; Kim & Hansen, 2010; O'Flynn, 2016; Kim et al., 2015). The first two categories of models carry several assumptions on runway operations; such as the absence of airspace constraints and the control of air traffic controllers; that seldom hold true in practical airport operation conditions. Simulation models offer the flexibility to understand the dynamic nature of runway capacity under different operational scenarios. However, the construction of such varying scenarios requires highly granular data on time-varying factors such as air traffic control regulations, which are often difficult to obtain. Moreover, the high cost in time, money and human resources (well-trained programmers) make simulation models remain frequent use only in big and hairy projects (such as detailed airfield design) rather than assessing runway capacity solely. Empirical models provide the ability to understand runway capacity in

the least data-hungry manner. Such models derive estimates of capacity from data on historical throughput that implicitly represent varied operational conditions. Nonetheless, we highlight that estimating capacity from the throughput data is not straightforward as there are external unobserved (unquantifiable) factors such as air traffic control regulations, that are correlated with both capacity and its observed determinants (such as the mix of aircraft served). We note that state-of-the-art empirical approaches in the literature, such as censored regression, fail to adjust for such unobserved sources of confounding, which limits their ability to provide a statistically robust and reproducible characterisation of runway capacity.
To fill these research gaps, we propose an empirical method to estimate runway capacity by employing a causal statistical approach, confounding-adjusted SFA, first developed by Karakaplan & Kutlu (2017); Karakaplan (2022) and further applied in Karakaplan & Kutlu (2019); Ojo & Baiyegunhi (2020); Xu et al. (2022). The proposed model uses historical throughput as the dependent variable and associated operational condition factors as covariates. Demand and delay are introduced in the inefficiency to determine the deviation from throughput to capacity. The adopted SFA delivers estimates of runway capacity by constructing a confounding-robust throughput frontier. The confounding biases originate due to either the correlation between observed operational condition factors and unobserved operational condition factors in random error, the correlation between inefficiency and random error, or both. To the best of our knowledge, this study presents the first application of *causal statistical modelling* in empirical estimation of runway capacity. Further, we apply the proposed model to three major airports around the world by making use of data on their day-to-day operations in 2018 as maintained by the Airport Benchmarking Group (ABG) within the Transport Strategy Centre (TSC) at Imperial College London Airport Benchmarking Group (n.d.). The corresponding weather records are sourced from Weather Underground Weather Underground (n.d.), Visual Crossing Visual Crossing (n.d.) and ECMWF Reanalysis v5 Copernicus Climate Change Service (n.d.). Based on these high-granular and large-scale data, the estimates of parameters for operational factors provide insights into how these factors contribute to changes in runway capacity. Additionally, the reliability of our runway capacity estimates is validated by comparing them against reported data from table lookup and spreadsheet in FAA's Advisory Circular Report 150/5060-5 FAA (1983) and Eurocontrol's Airport Corner Eurocontrol (n.d.); and testing statistically via censored regression Kim & Hansen (2010). Such accurate estimates of runway capacity under specific operational conditions facilitate air traffic controllers a better understanding of the capacity under these conditions and further decisions.

2 Data and Variables

Data

To estimate the runway capacity, we use large-scale and high-granular operational data for three congested airports provided by the ABG of TSC at Imperial College London. Because of data confidentiality, these three airports are anonymous in this study, denoted as A, B, and C. For each airport, the data provide detailed records for all aircraft movements in 2018, including scheduled and actual arrival and departure time, aircraft type, allocated runway and gate, and number of passengers. Such high-granular data including individual flight records ensure the metrics for operational factors are as close to what happens in actuality as possible Kim et al. (2015). The historical weather data for these airports are sourced from Weather Underground Weather Underground (n.d.), Visual Crossing Visual Crossing (n.d.) and ECMWF Reanalysis v5 Copernicus Climate Change Service (n.d.).
Based on these raw data, we construct panel data for each airport such that cross-sectional unit $i = 1, 2, \dots, N$ is defined as a time interval with 15 minutes length in a week and temporal unit $t = 1, ..., T$ is the weeks in 2018. These time intervals are particularly sampled from the peak hours (10 am - 8 pm) on weekdays in a week since night hours, some holidays, and other off-peak periods are typically times of low traffic demand, which should be excluded for capacity estimation and delay performance evaluation Gelhausen et al. (2013). Therefore, we calculated the relevant variables on the basis of all individual flights that are recorded to be served in the runway system during each quarter-hour.

Variables

In this section, we introduce the dependent variable (throughput), frontier variables (factors determining capacity), and environmental variables (factors determining inefficiency) for the confounding-adjusted SFA, summarized in Table 1. The frontier is the capacity, determined by various frontier variables. These factors are categorized into airport design factors (such as active runway layout and the number of runway exits), aircraft movement characteristics (such as arrival rate and fleet mix), environmental factors (such as visibility, ceiling, precipitation, density altitude, crosswind speed, headwind speed, tailwind speed), and air traffic control factors (such as separation gap, ATM procedures, and air traffic controllers' behaviors) Ashford et al. (2011). Inefficiency, the deviation from the frontier (capacity) to dependent variable (throughput) is explained with two environmental variables, delay, and demand. Delay has a positive effect on inefficiency since the runway system cannot be fully utilized as scheduled when scheduled flights are delayed Diana (2021). Throughput is the minimum value of capacity and demand. When demand is low during off-peak hours or in less congested airports, airport efficiency, and throughput decrease as follows Hansen (2004).

Table 1: Variables for confounding-adjusted SFA

Category	Factor	Definition	Data source
Output	Throughput	The number of aircraft that are served in the runway system	ABG
Frontier variables	Arrival rate (%)	The ratio of actual arrivals to actual departures	ABG
	Fleet mix (%)	The percent of large aircraft plus three times the percent of heavy aircraft	ABG
	Runway layout	An ordinal variable to represent the active runway layout	ABG
	Number of runway exits	The average number of runway exits of this active runway system	ABG
	Separation gap (s)	The average spacing time between leading and trailing aircraft	ABG
	Precipitation (mm)	The product of the condensation of atmospheric water vapor that falls under gravitational pull from cloud	Visual Crossing
	Visibility (miles)	The distance at which an object or light can be clearly discerned	Visual Crossing
	Ceiling (ft)	The height of the lowest clouds that cover more than half of the sky	ECMWF Reanalysis v5
	Density altitude (ft)	The pressure altitude corrected for temperature	Weather underground
	Crosswind speed (mph)	The average crosswind speed	Weather underground
	Headwind speed (mph)	The average headwind speed	Weather underground
	Tailwind speed (mph)	The average tailwind speed	Weather underground
Environmental variables	Delay (min)	The average flight delay for both arrival and departure flights	ABG
	Demand	The number of aircraft that are scheduled to be served	ABG

However, some frontier variables, such as the behaviors of air traffic controllers and the regulation rules they decided on are difficult to observe and estimate in empirical data. Therefore, the complex correlations between these unobserved frontier variables and other observed frontier/environmental variables result in confounding biases. In Figure 1, we use the level of air traffic control (ATC) to represent the air traffic controllers' preferences and regulation rules. For example, first-come-first-serve (FCFS) is a common discipline to serve the aircraft in the runway system, while quite often, air traffic controllers loosen (decrease) the level of air traffic control and process a sequence of arrivals first and insert departures without disturbing the arrivals flow, which increases the runway capacity De Neufville et al. (2013). Therefore, when the level of ATC is tight (increase), the runway capacity decreases and arrivals would be served as a priority. The downward bias observed in the effect of arrival rate on runway capacity is due to confounded with the effect of ATC on runway capacity. The fleet mix also has a negative effect on runway capacity, while the increase in ATC might cause either a decrease or an increase in the fleet mix depending on the behaviors of air traffic controllers. Therefore, the effect of fleet mix on runway capacity would have either an upward or downward bias. The positive effect of delay on inefficiency is explained by non fully utilized runway system when scheduled flights are delayed. The inefficiency also increases when ATC becomes strict. Therefore, the estimated causal effect of delay on inefficiency experiences an upward bias when the unobserved ATC cannot be handled properly.

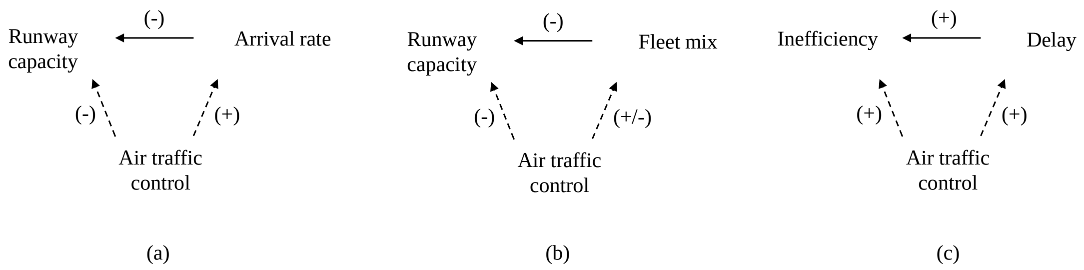

Figure 1: Causal relationships between variables

3 Methodology

To address the confounding biases discussed above, we employ a confounding-adjusted SFA, which allows the existence of a correlation between frontier variables/environmental variables and random error Karakaplan & Kutlu (2017); Karakaplan (2022).

$$y_{it} \quad = \quad c_{it} - u_{it} + v_{it}, \qquad where\ c_{it} = \boldsymbol{x_{fit}\alpha} \tag{1}$$

$$u_{it} \quad = \quad exp(\boldsymbol{x_{uit}\beta})u_i \tag{2}$$

$$\boldsymbol{x_{it}^{en}} \quad = \quad \boldsymbol{z_{it}\gamma} + \boldsymbol{\epsilon_{it}} \tag{3}$$

The confounding-adjusted SFA under a panel data specification is developed in Equation 1, where unit $i = 1, ..., N$ is the set of all 15 minutes length time intervals in a week; time $t = 1, ...T$ is the set of weeks. y_{it} is observed throughput for time interval i at week t. The deviation of throughput y_{it} from latent capacity c_{it} is the sum of negative inefficiency $-u_{it}$ and random error v_{it}. As we review in Section 2, runway capacity is determined by multiple airport operational condition factors. Therefore the latent runway capacity c_{it} is expressed as a function of a vector of airport operational condition factors, frontier variables, $\boldsymbol{x_{fit}}$ and α is a vector of unknown parameters to be estimated. v_{it} follows a normal distribution with time-invariant variance $v_{it} \sim N(0, \sigma_v^2)$. u_{it} is explained by a vector of environmental variables $\boldsymbol{x_{uit}}$. Delay and demand are introduced as environmental variables to determine inefficiency. u_i follows a non-negative normal distribution $u_i \sim N_+(\mu, \sigma_u^2)$. $\boldsymbol{x_{it}^{en}}$ represents a vector of frontier variables and environmental variables that are confounding with error term; $\boldsymbol{z_{it}}$ is a vector of instrumental variables for $\boldsymbol{x_{it}^{en}}$. The confounding biases are introduced by the correlation between ϵ_{it} and v_{it}, as Equation 4.

$$\begin{bmatrix} \widetilde{\epsilon}_{it} \\ v_{it} \end{bmatrix} = \begin{bmatrix} \Omega^{-1/2}\widetilde{\epsilon}_{it} \\ v_{it} \end{bmatrix} \sim N(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} I_m & \sigma_v\rho \\ \sigma_v\rho' & \sigma_v^2 \end{bmatrix}) \tag{4}$$

where Ω is the variance-covariance matrix of $\boldsymbol{\epsilon_{it}}$, σ_v^2 is the variance of v_{it}, and ρ is a vector of correlation between $\widetilde{\epsilon}_{it}$ and v_{it}. By adopting Cholesky decomposition of the variance-covariance matrix of $\begin{bmatrix} \widetilde{\epsilon}_{it} \\ v_{it} \end{bmatrix}$, we have

$$\begin{bmatrix} \widetilde{\epsilon}_{it} \\ v_{it} \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ \sigma_v\rho' & \sigma_v\sqrt{1-\rho'\rho} \end{bmatrix} \begin{bmatrix} \widetilde{\epsilon}_{it} \\ \widetilde{w}_{it} \end{bmatrix} \tag{5}$$

where $\widetilde{\epsilon}_{it} \sim N(0, 1)$ and $\widetilde{w}_{it} \sim N(0, 1)$ are independent. Therefore, Equation 1 is expressed as

$$y_{it} \quad = \boldsymbol{x_{fit}\alpha} - u_{it} + w_{it} + \sigma_v\rho'\widetilde{\epsilon}_{it} \quad = \boldsymbol{x_{fit}\alpha} + (\boldsymbol{x_{it}^{en}} - \boldsymbol{z_{it}\gamma})'\eta + e_{it} \tag{6}$$

where $e_{it} = w_{it} - u_{it}$, $w_{it} = \sigma_v\sqrt{1-\rho'\rho}\widetilde{w}_{it}$ and $\eta = \sigma_v\Omega^{-1/2}\rho$. Therefore, e_{it} is conditionally independent from frontier variables $\boldsymbol{x_{fit}}$ given $\boldsymbol{x_{it}^{en}}$ and $\boldsymbol{z_{it}}$. Therefore, the log-likelihood function for each panel i, constructing by all T_i time periods for unit i, is given by Equation 7.

$$lnL_i = lnL_{i,y|x} + lnL_{i,x} \tag{7}$$

$$L_{i,y|x} = -\frac{1}{2}(T_i ln(2\pi\sigma_w^2) + \frac{e_i' e_i}{\sigma_w^2} + (\frac{\mu^2}{\sigma_u^2} - \frac{\mu_{i*}^2}{\sigma_{i*}^2})) + ln(\frac{\sigma_{i*}\Phi(\frac{\mu_{i*}}{\sigma_{i*}})}{\sigma_u \Phi(\frac{\mu}{\sigma_u})}) \quad (8)$$

$$lnL_{i,x} = -\frac{1}{2}\sum_{t=1}^{T_i}(ln(|2\pi\Omega|) + \epsilon_{it}'\Omega^{-1}\epsilon_{it}) \quad (9)$$

where $\mu_{i*} = \frac{\sigma_w^2\mu - \sigma_u^2 e_i' h_i}{\sigma_u^2 h_i' h_i + \sigma_w^2}$; $\sigma_{i*}^2 = \frac{\sigma_w^2\sigma_u^2}{\sigma_u^2 h_i' h_i + \sigma_w^2}$; $\sigma_w = \sigma_v\sqrt{1-\rho'\rho}$; $h_i = (h_{i1}, h_{i2}, ..., h_{iT_i})$; $h_{it}^2 = exp(x_{uit}'\beta)$; $e_{it} = y_{it} - x_{fit}'\alpha - \epsilon_{it}'\eta$; $\epsilon_{it} = x_{it}^{en} - z_{it}\gamma$; and Φ is the standard normal cumulative distribution function. The formula to predict inefficiency is $EFF_i = exp(-u_i)$. Based on the $\hat{\alpha}$ obtained from Maximum Likelihood Estimation, the expected runway capacity estimation is $\boldsymbol{x_{fit}\hat{\alpha}}$.

4 Results and discussion

Estimation results

In this section, we apply confounding-adjusted SFA to Airport A, B, and C respectively. Table 2 provides insights into how these airport operational factors contribute to the change of runway capacity and environmental factors affect inefficiency. For those exogenous factors, their parameters are in line with the intuition. The positive effect of the number of runway exits shown in airport A is in accordance with the intuition that sufficient runway exits shorten runway occupancy time and thus increase runway capacity. However, insignificant and even counter-intuitive results shown in other airports might be due to little variations in the number of runway exits for those airports with a symmetric parallel runway layout. Higher visibility and ceiling mean a greater flexible flight rule and more runway capacity. They are less significant in some airports because these European airports prefer to declare a more robust capacity, usually closer to the capacity under instrument meteorological condition (IMC), and thus capacity experiences less change when visibility and ceiling decrease Gulding et al. (2010). The results show that the increase in precipitation, crosswind speed, headwind speed, and even tailwind speed causes a significant reduction in runway capacity. Although headwind shortens runway occupancy time for both takeoffs and landings, such improvement is offset by a long time traveling through the terminal maneuvering area with a low speed due to the headwind.

In addition to these exogenous variables, three variables (arrival rate, fleet mix, and delay) are identified as endogenous. The endogeneity of these three variables is detected as statistically significant in terms of both joint significance η and individual significance η_1, η_2, η_3. Therefore, the proposed model corrects these endogenous variables with their associated lagged difference as instruments $\Delta x_{it}(x_{it} - x_{it-1})$ and $\Delta x_{it-1}(x_{it-1} - x_{it-2})$ Arellano & Bover (1995). The performance of Model EN is reliable since these instruments pass the tests for exclusion (exogeneity) and inclusion restriction (relevance), which are the J test for over-identifying instruments (less than critical value $\chi^2_{3,0.95} = 7.81$) and reduced from regression (greater than 10 based on the rule of thumb) respectively. The results show that arrival rate and fleet mix have negative effects on runway capacity, and delay has a positive effect on the inefficiency term, which is in accordance with intuitive signs.

Validation results

To further assess the reliability of the proposed confounding-adjusted SFA, we aim to validate our capacity estimates by simply comparing them with estimates from other empirical methods and testing them via a statistical model. Table 3 displays hourly capacity estimates under visual meteorological condition (VMC) from different empirical sources and methods, including table lookup FAA (1983), spreadsheet FAA (1983), Airport Corner in Eurocontrol Eurocontrol (n.d.), and our proposed confounding-adjusted SFA. To make this comparison practical, the operational condition across methods should be as consistent as possible, although the required inputs (operational condition factors) in each method are different.

The estimates from confounding-adjusted SFA in airport B and C appear to be low compared with the direct estimates from table lookup and spreadsheet methods while being within a reasonable range with 90 % of those estimates (in parentheses). Such a 10% reduction is suggested to make estimates comparable with the actual flow during peak hours since estimates from the previous

Table 2: Estimation Results

	Airport A	Airport B	Airport C
Dep.var:throughput			
constant	24.401***	20.806***	23.751***
	(0.280)	(0.068)	(0.312)
arrival rate	-0.247***	-0.448***	-0.513***
	(0.055)	(0.033)	(0.047)
fleet mix	0.074	-0.225***	-0.156**
	(0.061)	(0.042)	(0.049)
runway layout	0.614***	0.136***	0.409***
	(0.046)	(0.031)	(0.032)
number of runway exits	0.617***	-0.019	-0.052*
	(0.046)	(0.026)	(0.030)
separation gap	-0.302**	-0.567***	-0.070
	(0.093)	(0.076)	(0.082)
precipitation	-0.080*	-0.011	-0.103**
	(0.034)	(0.026)	(0.033)
visibility	-0.015	0.070*	0.106**
	(0.041)	(0.029)	(0.034)
ceiling	-0.010	0.068*	-0.044
	(0.042)	(0.029)	(0.032)
density altitude	0.464***	0.192***	0.323***
	(0.040)	(0.028)	(0.034)
crosswind speed	-0.115***	-0.069**	-0.080**
	(0.033)	(0.027)	(0.031)
headwind speed	-0.149***	-0.114***	-0.099**
	(0.036)	(0.029)	(0.032)
tailwind speed	-0.047*	-0.054*	-0.100**
	(0.034)	(0.027)	(0.031)
$Dep.var{:}ln\sigma_u^2$			
constant	3.456***	0.312*	3.671***
	(0.126)	(0.149)	(0.127)
demand	-0.504***	-0.529***	-0.542***
	(0.28)	(0.140)	(0.027)
delay	0.032***	0.221***	0.029**
	(0.009)	(0.038)	(0.009)
$Dep.var{:}ln\sigma_w^2$			
constant	2.632***	2.241***	2.458***
	(0.012)	(0.012)	(0.012)
η_1(arrival rate)	-0.458***	-0.102*	-0.190*
	(0.089)	(0.056)	(0.075)
η_2(fleet mix)	0.235*	-0.051	-0.226**
	(0.095)	(0.063)	(0.09)
η_3(delay)	-0.263***	-0.212***	-0.322***
	(0.063)	(0.054)	(0.062)
η (joint endogeneity)	x2=48.76	x2=19.61	x2=40.09
	p=0.000	p=0.000	p=0.000
exogeneity	0.55	3.21	0.13
relevance	34668.57	52558.43	37160.94
observations	14000	13999	13999
log likelihood	-85190.00	-77971.76	-82539.24
mean technical efficiency	0.0430	0.4419	0.0316
median technical efficiency	0.060	0.4295	0.0031

Notes: Standard errors in parenthesis. Asterisks indicate significance at 0.1% (***), 1 % (**) and 5% (*)levels. All inputs are demeaned.

two empirical methods are calculated based on minimum separation gap TRB (2012). The large deviation in airport A is largely caused by mis-specifying runway layout in the first two methods without considering the runway incursion in actual runway operation that arrivals crossing the departure runway for taxing in. The performance of our proposed method is also validated in airport B and C by comparing it with the declared capacity from Eurocontrol's Airport Corner. Such declared capacity for the whole system is calculated by summing up the declared capacity for each runway as provided, which is valid when runways are independent, such as airports B and C. However, in airport A, the dependent parallel runways and incursion problem make actual operation cannot attain the maximum movement of each runway. Our estimate from the proposed method is still reasonable and acceptable since the percentile 99 of the throughput per hour observed in 2019 is only 108 from Eurocontrol's performance review report Commission (1997). Other slight differences, within the 10% or even less variance range, might be due to different specifications in operational condition factors across methods or measurement errors.

Table 3: Validation via comparison

Airport	Runway layout	Capacity estimation method	Fleet mix (%)	Maximum capacity (VMC, hourly)
A	Dual independent parallel runway	Table lookup	121-180	189 (170)
		Spreadsheet	121-180	187 (168)
		Eurocontrol	–	148
		Confounding-adjusted SFA	149	98
B	Independent parallel runway	Table lookup	121-180	103 (93)
		Spreadsheet	121-180	95 (86)
		Eurocontrol	–	88
		Confounding-adjusted SFA	174	83
C	Independent parallel runway	Table lookup	81-120	111 (100)
		Spreadsheet	81-120	110 (99)
		Eurocontrol	–	90
		Confounding-adjusted SFA	116	95

Although the reliability of our estimates has been validated via simple comparison with other empirical methods, it is just a point estimate for the capacity under normal condition. Our proposed method is able to provide the estimates of runway capacity for each 15 minutes time interval under different runway operational conditions. Therefore, we further adopt censored regression to validate the proposed method statistically Kim & Hansen (2010).

$$y_i^* = \beta_1 x_i + \epsilon_i \tag{10}$$

$$y_i = \begin{cases} y_i^* & y_i^* \leq c_i \\ c_i & y_i^* > c_i \end{cases} \tag{11}$$

In this model, y_i^* is the latent capacity for a 15 minutes time interval i, y_i is the throughput (observed capacity), and c_i is the upper censoring limit, demand. x_i is the estimated capacity from the confounding-adjusted SFA. ϵ_i follows the independent identically distributed normal distribution with mean 0 and variance σ_0^2. If estimates from our proposed method are reliable, $\hat{\beta}_1 \to 1$ could be yielded from the estimation of censored regression. Therefore, we calculate t-statistic $\frac{\hat{\beta}-\beta}{SE(\beta)}$ for the coefficients with hypothesis $H_0 : \beta_1 = 1$ against $H_1 : \beta_1 \neq 1$. Table 4 shows the results of censored regression based on the whole data set with 14000 time intervals. The estimates for β_1 are reasonably and acceptably close to the expected value in practice. Although the null hypothesis is rejected in terms of the t-statistic value, this is always a statistical problem that the null hypothesis will be rejected when the sample size is extremely large. Overall, the estimates from confounding-adjusted SFA are found to be reliable and compared favorably with the other capacity estimation method.

Table 4: Validation via censored regression

	Airport A			Airport B			Airport C		
	Estimate	Error	t-statistic	Estimate	Error	t-Statistic	Estimate	Error	t-Statistic
β_1	0.9239	0.0034	-22.38235	1.0683	0.0031	22.03226	0.8861	0.0033	-34.51515
σ_0	6.9568	0.0605	–	5.2932	0.0463	–	6.8202	0.0571	–

5 Conclusion

The contribution of this study is bifold, (1) first developing the idea of empirically estimating the runway capacity based on the SFA framework along with available large-scale and high-granular data, and (2) further demonstrating the added value by handling the confounding problems properly and obtaining unbiased estimates of runway capacity and the parameters of their operational condition factors via a confounding-adjusted SFA. This confounding-adjusted SFA is applied to three airports' day-to-day operation data respectively. The estimation results show the parameters for those exogenous variables are in line with the intuition. The endogeneities for those endogenous variables are statistically significant, and their parameters after instrument correction via the proposed method are also in accordance with the intuitive signs. Moreover, the runway capacity estimates from our proposed method are validated by comparing them with estimates from other empirical methods and testing via a statistical model. The capacity estimates from our proposed method are within 10% or less variance range and even nearly the same as the estimates from other methods in terms of the point estimate under the normal operational condition. The results from the statistical test also demonstrate the estimated capacity is reasonably and acceptably close to the true capacity in practice. Therefore, such unbiased estimates of parameters for associated operational condition factors facilitate airport operators' better understanding of how these factors contribute to the dynamic change of runway capacity. Based on unbiased parameters, runway capacity during short time intervals is estimated accurately which allows air traffic controllers to manage the daily demand on the runway by allocating optimal runway capacity effectively and modeling airport delay and delay propagation within airport networks. Moreover, the proposed method developed to estimate runway capacity is also capable to evaluate the latent capacity in other transport modes.

Acknowledgements

The authors thank Airport Benchmarking Group (ABG) in Transport Strategy Centre (TSC) for providing the data used in the analysis.

References

Airport Benchmarking Group. (n.d.). *Airport benchmarking at the transport strategy centre.* `https://www.imperial.ac.uk/transport-studies/transport-strategy-centre/applied-research/benchmarking-groups/airports`.

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of econometrics*, *68*(1), 29–51. doi: 10.1016/0304-4076(94)01642-D

Ashford, N. J., Mumayiz, S., & Wright, P. H. (2011). *Airport engineering: planning, design, and development of 21st century airports.* John Wiley & Sons.

Barrer, J., Kuzminski, P., & Swedish, W. (2005). Analyzing the runway capacity of complex airports. In *Aiaa 5th atio and16th lighter-than-air sys tech. and balloon systems conferences* (p. 7354). doi: 10.2514/6.2005-7354

Blumstein, A. (1959). The landing capacity of a runway. *Operations Research*, *7*(6), 752–763. doi: 10.1287/opre.7.6.752

Bubalo, B., & Daduna, J. R. (2011). Airport capacity and demand calculations by simulation—the case of berlin-brandenburg international airport. *NETNOMICS: Economic Research and Electronic Networking*, *12*(3), 161–181. doi: 10.1007/s11066-011-9065-6

Cheung, W. L., Desai, J., Prakash, R., & Lim, H. S. (2017). An improved macroscopic analytical model for estimating runway capacity. In *Iie annual conference. proceedings* (pp. 567–572).

Cheung, W. L., Piplani, R., Alam, S., & Bernard-Peyre, L. (2021). Dynamic capacity and variable runway configurations in airport slot allocation. *Computers & Industrial Engineering*, *159*, 107480. doi: 10.1016/j.cie.2021.107480

Commission, P. R. (1997). *Airport capacity imbalance* (Tech. Rep.).

Copernicus Climate Change Service. (n.d.). *Era5 hourly data on pressure levels from 1959 to present.* https://cds.climate.copernicus.eu.

De Neufville, R., Odoni, A. R., Belobaba, P. P., & Reynolds, T. G. (2013). *Airport systems: Planning, design, and management.* McGraw-Hill Education. Retrieved from https://www.accessengineeringlibrary.com/content/book/9780071770583

Diana, T. (2021). Doing more with less: An assessment of capacity utilisation using stochastic frontier and spectral analysis models in the case of atlanta hartsfield-jackson international airport. *Journal of Airport Management*, *16*(1), 87–104. Retrieved from https://www.ingentaconnect.com/contentone/hsp/cam/2021/00000016/00000001/art00009?crawler=true&mimetype=application/pdf&casa_token=MRTuYR_c3NoAAAAA:6qw3n_T3ZCQV5cWJnDA04kgqXUhhwNowO7vWZHhnJ_migrlaPcGPQv4aaIS1-js4a3AK40G2U8X37ZCW

Eurocontrol. (n.d.). *Public airport corner homepage data services.* https://ext.eurocontrol.int/airport_corner_public/.

FAA. (1983). Airport capacity and delay. *Federal Aviation Administration, Advisory Circular AC 150/5060-5*. Retrieved from https://www.faa.gov/documentLibrary/media/Advisory_Circular/150_5060_5.pdf

Gelhausen, M. C., Berster, P., & Wilken, D. (2013). Do airport capacity constraints have a serious impact on the future development of air traffic? *Journal of Air Transport Management*, *28*, 3–13. doi: 10.1016/j.jairtraman.2012.12.004

Gilbo, E. P. (1993). Airport capacity: Representation, estimation, optimization. *IEEE Transactions on control systems technology*, *1*(3), 144-154. doi: 10.1109/87.251882

Gulding, J., Knorr, D., Rose, M., Chen, X., Enaud, P., & Hegendoerfer, H. (2010). Us/europe comparison of atm-related operational performance. *Air Traffic Control Quarterly*, *18*(1), 5–27.

Hansen, M. (2004). Post-deployment analysis of capacity and delay impacts of an airport enhancement: case of a new runway at detroit. *Air Traffic Control Quarterly*, *12*(4), 339-365. doi: 10.2514/atcq.12.4.339

Karakaplan, M. U. (2022). Panel stochastic frontier models with endogeneity. *The Stata Journal*, *22*(3), 643–663. doi: 10.1177/1536867X221124539

Karakaplan, M. U., & Kutlu, L. (2017). Endogeneity in panel stochastic frontier models: an application to the japanese cotton spinning industry. *Applied Economics*, *49*(59), 5935–5939. doi: 10.1080/00036846.2017.1363861

Karakaplan, M. U., & Kutlu, L. (2019). School district consolidation policies: endogenous cost inefficiency and saving reversals. *Empirical Economics*, *56*(5), 1729–1768. doi: 10.1007/s00181-017-1398-z

Kim, A., & Hansen, M. (2010). Validation of runway capacity models. *Transportation research record*, *2177*(1), 69–77. doi: 10.3141/2177-09

Kim, A., Rokib, S., & Liu, Y. (2015). Refinements to a procedure for estimating airfield capacity. *Transportation Research Record*, *2501*(1), 18–24.

Kuzminski, P. C. (2013). An improved runwaysimulator—simulation for runway system capacity estimation. In *2013 integrated communications, navigation and surveillance conference (icns)* (pp. 1–11). doi: 10.1109/ICNSurv.2013.6548575

Mascio, P. D., Rappoli, G., & Moretti, L. (2020). Analytical method for calculating sustainable airport capacity. *Sustainability*, *12*(21), 9239. doi: 10.3390/su12219239

Ojo, T., & Baiyegunhi, L. (2020). Impact of climate change adaptation strategies on rice productivity in south-west, nigeria: An endogeneity corrected stochastic frontier model. *Science of The Total Environment*, *745*, 141151. doi: 10.1016/j.scitotenv.2020.141151

O'Flynn, S. (2016). Airport capacity assessment methodology—acam manual. *Network Manager, EUROCONTROL, Tech. Rep*, *1*. Retrieved from `https://www.eurocontrol.int/publication/airport-capacity-assessment-methodology-acam-manual`

Pyrgiotis, N., Malone, K. M., & Odoni, A. (2013). Modelling delay propagation within an airport network. *Transportation Research Part C: Emerging Technologies*, *27*, 60–75. doi: 10.1016/j.trc.2011.05.017

TRB. (2012). Evaluating airfield capacity. *Transportation Research Board, National Academy of Sciences*. doi: 10.17226/22674

Visual Crossing. (n.d.). *Visual crossing historical weather data & weather forecast data.* `https://visualcrossing.com/weather-data`.

Weather Underground. (n.d.). *Weather underground weather history and data archive.* `https://www.wunderground.com/history`.

Xu, M., Tan, R., & He, X. (2022). How does economic agglomeration affect energy efficiency in china?: Evidence from endogenous stochastic frontier approach. *Energy Economics*, *108*, 105901. doi: 10.1016/j.eneco.2022.105901