Modelling Airline Network Routing and Scheduling under Airport Capacity Constraints

Antony D. Evans*, Andreas Schäfer†, Lynnette Dray‡
Institute for Aviation and the Environment, University of Cambridge, Cambridge, UK

A flight routing and scheduling model is under development that predicts airline routing and scheduling under airport capacity constraints. It consists of several components describing different aspects of the air transport system, including passenger demand, airline competition, flight delay, and airline cost. These models are integrated into a flight routing and scheduling model in which an objective function is defined to maximize airline system profit within a routing network, subject to constraints. This framework allows the relationships between fares, passenger demand, infrastructure capacity constraints, flight delays, flight frequencies, and routing network to be simulated. In this paper the integrated flight routing and scheduling model is first applied to a series of simple theoretical routing networks to illustrate its capabilities. With increasing airport capacity constraints the results show an increase in average fares, a decrease in O-D passenger demand, and a shift in flight routing away from the most constrained airports. The model is then applied to a network of airports in the United States with 2005 population, income and airport capacity inputs. With further development the model is to be applied to forecasting air traffic system growth, including network and schedule changes resulting from increasing delays, in the Aviation Integrated Modelling (AIM) project under development at the University of Cambridge.

I. Introduction

WORLDWIDE demand for air travel has shown significant growth over the past five decades. Between 1960 and 2005 worldwide scheduled passenger air travel grew from 109 billion passenger-km travelled to 3.7 trillion – an average growth rate of over 8% per year [1,2]. Forecasts for future growth are also high – the Intergovernmental Panel on Climate Change (IPCC) forecast a growth rate between 1990 and 2015 of 5% per year [1], which corresponds to that of both the Airbus Global Market Forecast from 2007 to 2026 [3] and the Boeing Current Market Outlook from 2006 to 2026 [4]. By 2050 conservative estimates predict a 30-110% growth in passenger kilometres travelled over 2005 levels [5], while more aggressive estimates predict an increase of an order of magnitude [6]. Associated with such growth in demand for air travel is a growth in air traffic (number of aircraft movements) to serve that demand, as modelled by Hancox and Lowe [7], Bhadra et al. [8,9], and Reynolds et al. [10]. This growth in air traffic is expected to produce a significant environmental impact, as reported by the IPCC [1] and Cairns et al. [11], including air quality and noise impacts, and global climate change.

Growth in air traffic is already constrained by environmental restrictions (particularly noise). An emerging constraint is air traffic system capacity, given widespread local community resistance and environmental restrictions to airport capacity expansion. Airport and airspace capacity already constrain flight operations at many major airports in the United States and Europe. In the United States, average arrival delays at 71 airports were greater than 10 minutes in 2005, with Newark Liberty International Airport experiencing an average arrival delay of 23 min [12]. In Europe, average arrival delays at 42 airports were greater than 10 minutes in 2006, with Istanbul International Ataturk Airport and London Luton Airport experiencing average arrival delays of 19 and 18 minutes respectively [13].

As shown by Reynolds et al. [10], if airport capacity in the United States only grows as described in the US Federal Aviation Administration (FAA) Operational Evolution Plan [14], and airlines continue to increase air traffic to match the projected growth in demand, average arrival delays for the 50 busiest airports in the system would be

*PhD Student, Department of Architecture, Scroope Terrace, Cambridge, CB2 1PX, UK. AIAA Member.
†University Lecturer, Department of Architecture, Scroope Terrace, Cambridge, CB2 1PX, UK.
‡Post-doctoral Researcher, Institute for Aviation and the Environment, Scroope Terrace, Cambridge, CB2 1PX, UK. AIAA Member.

over 2 hours per flight. However, as suggested by Kostiuk et al. [15] and Reynolds et al. [10], these delays are unlikely to occur in reality as airlines would adjust their operations, including minimising the negative impacts of the increasing costs associated with delays and the reduction in passenger demand. The airline response to delays would alter the structure of future air traffic growth, as airlines may use differently sized aircraft, and avoid the most congested airports and routes. In order to generate a plausible forecast of air traffic growth in a capacity constrained system it is therefore essential to quantify how passengers and airlines are likely to respond to capacity constraints.

Airline responses to delay include avoiding congested hubs and gateways (departure routes), using secondary airports, moving flights to off-peak times, broadening the range of departure times, and reducing flight frequency while increasing aircraft size [15]. The effects of each of these responses can be examined using a scenario-based approach as implemented by Long et al. [16], where responses are applied exogenously and their effects examined. However, passenger and airline responses are highly interrelated, with airlines responding to changes in passenger demand by adjusting fares and schedules. In a highly competitive market such as the airline industry in the United States and Europe, airlines are also highly constrained in how they can respond without losing market share (often leading to lower prices and higher flight frequencies than is optimal, as demonstrated by Schipper [17] and Carlsson [18]). Airlines do, however, have the flexibility to adjust their routing networks, and thus change routes along which passengers are flown from their true origin to their ultimate destination.

This paper describes an integrated flight routing and scheduling model to predict changes in passenger demand, fares, flight frequencies, and airline routing under airport capacity constraints. Other airline responses to delays, including moving flights to off-peak times and broadening the range of departure times operated are not examined in this paper. The modeling approach is described in detail in Section II, followed by its application to a series of simple theoretical networks of airports in Section III. These include: a simple hub and spoke network allowing analysis of the effect of capacity constraints on hub and spoke routing; a network with two hubs airports allowing analysis of the effect of capacity constraints on the distribution of traffic between hubs; and a network including a simple multi-airport system allowing analysis of the effect of capacity constraints on the distribution of traffic within a multi-airport system.

Section IV describes the application of the model to a network of airports in the United States with 2005 population, income and airport capacity inputs, allowing the model's predictive capability to be identified and discussed. Conclusions and recommendations for further developments are discussed in Section V.

II. Modeling Approach

The flight routing and scheduling model developed integrates a number of sub-models, including a passenger demand model, an airline competition model, a flight delay model, and an airline cost model. By integrating these sub-models the impact of flight delays on fare, flight frequency, and routing network can be examined. The integration of the sub-models within the flight routing and scheduling model is presented in Figure 1 below, and is described in detail following a description of each sub-model in the following sections.

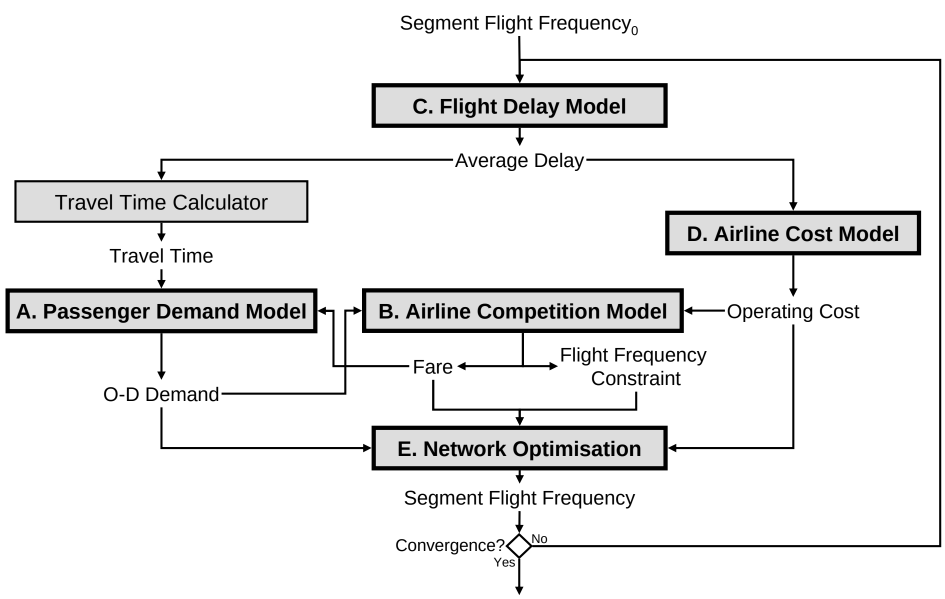

Figure 1. Model integration

A. Modeling Passenger Demand

True origin/ultimate destination (O-D) passenger demand between cities *i* and *j* is modelled using a simple one-equation gravity-type model, as applied by Reynolds et al. [10]:

$$D_{ij} = (I_i I_j)^{\alpha} (P_i P_j)^{\gamma} e^{\delta A_{ij}} e^{\varepsilon B_{ij}} \left(\overline{Fare}_{i,j} + \theta_1 \cdot \overline{T}_{i,j} + C_t \right)^{-\tau} \tag{1}$$

where I is the greater metropolitan area per capita income; P is the greater metropolitan area or equivalent population; A and B are binary variables indicating whether one or both cities in the pair have qualities which might increase visitor numbers (e.g. a major tourist destination or capital city); $\overline{Fare}$ is passenger airfare between the cities averaged over all itineraries; θ_1 is the passenger value of travel time; $\overline{T}$ is the travel time between the cities averaged over all itineraries; C_t covers all other passenger costs associated with the trip, such as getting to and from the airport, and is the same for all airports; and the exponents give the elasticity of demand to each of the explanatory variables (i.e. % change in demand resulting from a % change in each explanatory variable). The expression in brackets represents the generalized cost to a passenger of air travel between the cities, and it is through this expression that it is possible to include the demand-reducing effect of increased fares as well as that of increased travel time. Using demand data for the United States in 2000 [19,20,21], the coefficients (exponents) in equation 1 are estimated separately for short-haul, medium-haul and long-haul journeys. All estimated coefficients are significant at the 95% confidence level, with R^2 values ranging from 0.46 to 0.83 (the lower value is for short-haul routes for which the lack of competing modes in this model formulation has a strong effect).

The model described in equation 1 does not capture some passenger demand effects that are significant in some world regions, such as passenger mode choice. This simplification is, however, necessary for the computational efficiency required of the model for integration in the framework presented in Figure 1.

B. Modeling Airline Competition

As described in Section II.A above, passenger demand is a function of fare and travel time. Fares are defined according to airline economics. In a competitive market airlines typically compete using fare and flight frequency. This means that in order to increase their market share airlines may reduce fares or increase flight frequencies on a given O-D market. However, most airlines will respond in a similar way, subject to their cost constraints. Such price reductions and increases in flight frequency will continue as long as airlines are still making a profit. Airline microeconomics and game theory can be applied to estimate the equilibrium to which average fares and flight frequencies settle.

Schipper [17] and Carlsson [18] describe a formulation of the equilibrium average fare and flight frequency as a function of passenger value of time, airline costs and passenger demand, derived by defining the flight schedule as an address (or spatial) model, and solving a two stage game to maximise airline profit. In the first stage of the game airlines simultaneously choose flight frequencies, and in the second stage, after having observed the other airlines' chosen frequencies, the airlines simultaneously choose fares. This formulation is presented in equations 2 and 3 below. Equation 3 only applies to markets served by more than two airlines.

$$Fare^l = C_p{}^l + \frac{\theta_2}{nx^l} \tag{2}$$

$$nx^l = \sqrt{\frac{\theta_2 D}{C_f{}^l}\left(\frac{n-2}{n}\right)} \tag{3}$$

where $Fare^l$ represents the equilibrium average fare of airline l for the O-D city-pair market examined; $C_p{}^l$ represents the marginal passenger cost of airline l; θ_2 represents the passenger value of schedule delay[§]; n represents the number of airlines serving the market; x^l represents equilibrium flight frequency of airline l on the market, D represents O-D passenger demand on the market; and $C_f{}^l$ represents average cost per flight of airline l.

The formulation described by Schipper [17] and Carlsson [18] was modified to model variable passenger demand as described by equation 1. The results for average fare and flight frequency are presented in equations 4 and 5 below as a function of the same parameters described for equations 1, 2 and 3. K refers to the expression before the generalized cost term in equation 1: $(I_i I_j)^{\alpha} (P_i P_j)^{\gamma} e^{\delta A_{ij}} e^{\varepsilon B_{ij}}$.

[§] Schedule delay refers to the time difference between when a passenger wants to fly, and when the flight which he/she chooses to fly on departs. In this paper passenger value of schedule delay is assumed to be equal to passenger value of travel time.

$$Fare^l = \frac{\left(\tau\theta_1 - \theta_1 + \theta_2\overline{T}nx^l - C_t nx^l - C_p{}^l nx^l\right) - \sqrt{\begin{aligned}&\left(-\tau\theta_1 + \theta_1 - \theta_2\overline{T}nx^l - C_t nx^l + C_p{}^l nx^l\right)^2 + \\ &4nx^l\left(\tau C_p{}^l\theta_1 + \theta_1\theta_2\overline{T} + \theta_1 C_t + \theta_2\overline{T}C_p{}^l nx^l + C_t C_p nx^l\right)\end{aligned}}}{-2nx^l} \quad (4)$$

$$\left[\frac{K\left(Fare^l + \theta_2\overline{T} + C_t\right)^{-\tau}\left(Fare^l - C_p{}^l\right)}{nx^l} - C_f{}^l\right] + x^l\left[\begin{aligned}&-K\tau\left(Fare^l + \theta_2\overline{T} + C_t\right)^{-\tau-1}\left(\frac{\partial Fare^l}{\partial x^l} + \theta_2\frac{\partial\overline{T}}{\partial x^l}\right)\frac{\left(Fare^l - C_p{}^l\right)}{nx^l} + \\ &K\left(Fare^l + \theta_2\overline{T} + C_t\right)^{-\tau}\frac{\partial Fare^l}{\partial x^l}\frac{1}{nx^l} + \frac{K}{\theta_1}\left(Fare^l + \theta_2\overline{T} + C_t\right)^{-\tau}\left(Fare^l - C_p{}^l\right)\frac{\theta_1}{\left(nx^l\right)^2}\end{aligned}\right] = 0 \quad (5)$$

Equation 5 includes the derivative of fare with respect to flight frequency ($\partial Fare^l / \partial x^l$), which can be derived from equation 4, and the derivative of average travel time with respect to flight frequency ($\partial\overline{T}/\partial x^l$), which was found to be small and is therefore ignored. Equation 5 cannot be solved analytically, but can be solved iteratively.

The formulation described above does not model some important characteristics of airline and passenger behaviour. Most notably, the model maximises profit to solve for a single fare, and therefore does not account for price discrimination (yield management), which is employed extensively in the airline industry. Instead the model estimates average fares and flight frequencies by O-D market. The model also does not distinguish between different passenger routings on the same O-D market. Fares, passenger value of time, and cost elasticities may, however, differ quite significantly for different routes in the same market – particularly between direct flights and connecting flights. Finally, the address model applied also does not capture passenger preferences to fly at certain times of day (particularly the early morning and evening), and therefore ignores the increased demand at these times. These simplifications are, however, necessary for the computational efficiency required of the model for integration in the framework presented in Figure 1.

Equations 4 and 5 are applied to estimate equilibrium average fares and flight frequencies for O-D city pairs. However, flight delays resulting from an increase in flight frequency impact average cost per flight through an increase in block time, and O-D passenger demand through an increase in travel time. Modeling of flight delay is therefore also essential to estimate the effects of airport capacity constraints on passenger demand, fares, and flight frequencies.

C. Modeling Flight Delay

The impact of airport capacity constraints on airline routing and scheduling is modelled using a rapid airport delay model similar to that employed by Reynolds et al. [10], and described in detail by Evans [22]. In this model flight delays, both on the ground and in the air, are estimated as a function of flight frequencies and airport capacity constraints, and are added to gate departure delays due to mechanical failures and late arrivals, which are assumed to remain at current levels (assuming schedule padding increases to maintain schedule reliability). Delays due to airport capacity constraints are estimated using queuing theory, applying the cumulative diagram approach and classical steady state simplifications described by de Neufville and Odoni [23]. Runway departure delays are distributed between the taxiway and the gate according to a taxi-out threshold calculated for each airport from historical delay data. Similarly, delays due to destination airport capacity constraints are distributed between the air and ground according to an airborne holding threshold calculated for each airport from historical delay data, and above which delay is assumed to be propagated upstream to the departure gate. The delays estimated by this model increase passenger travel time (T in equation 1), and increase airline cost per flight $C_f{}^l$ and cost per passenger $C_p{}^l$ in equations 4 and 5.

D. Modeling Airline Costs

Airline costs modelled include aircraft operating costs, aircraft servicing costs, traffic servicing costs, passenger servicing costs, reservation and sales costs, and other indirect and system overhead costs, as described by Belobaba [24]. Aircraft operating costs include fuel and oil costs, crew costs, maintenance costs, aircraft rental, depreciation and amortization costs, and enroute airspace charges. With the exception of fuel and oil costs, these costs can be input directly from US DOT Form41 data [21] for modeling in the United States. Fuel costs are calculated independently as a function of fuel price and aircraft fuel burn in each flight phase: ground idle, taxi, take-off, climb-out, cruise, descent, approach and landing. Fuel burn rates are estimated according to the EUROCONTROL Base of Aircraft Data (BADA) [25] and the ICAO Aircraft Engine Emissions Databank [26].

All other costs are input directly from DOT Form41 data [21] for modeling in the United States. These include aircraft servicing costs, covering handling aircraft on the ground and landing fees; traffic servicing costs, covering the processing of passengers, baggage and cargo at airports; passenger servicing costs, covering meals, flight attendants and in-flight services; reservation and sales costs, covering airline reservations and ticket offices, including travel agency commissions; and other indirect and system overhead costs, covering advertising and publicity expenses and general and administrative expenses.

Airline costs are modelled per flight and per passenger, as required by the airline competition modelled described in Section II.B. Costs per flight include all aircraft operating costs and aircraft servicing costs, with the exception of the proportion of the fuel burn that can be attributed to passengers directly. This fuel burn, along with traffic servicing costs, passenger servicing costs, reservations and sales costs, and other indirect and system overhead costs are modelled per passenger.

E. Modeling Flight Routing and Scheduling

The passenger demand model, the airline competition model, and the flight delay model are integrated into the airline routing and scheduling model, which is based on maximisation of system profit to optimise flight frequencies and routing using large scale linear programming methods, similar to the approach used by Harsha [27]. The airline revenue term is based on passengers flown by itinerary and average fares on those itineraries. The cost terms are based on airline cost per flight multiplied by the flight frequency on a given flight segment, airline cost per passenger multiplied by passengers flown by itinerary, and airline spill cost multiplied by total spill[**]. This objective function is presented in equation 6 below. As described above, passenger demand is a function of fare and travel time, and is estimated by the passenger demand model described in Section II.A. Average fares are a function of airline costs and flight frequency, and are estimated by the airline competition model described in Section II.B. Airline costs are a function of aircraft types flown and block time predicted by the flight delay model, as estimated by the airline cost model described in Section II.D. Flight block times, and consequently passenger travel time, is a function of flight delays, which are estimated by the flight delay model described in Section II.C.

$$\max\left(\sum_{i,j}\sum_{p\in Itin_{i,j}}\overline{Fare}_{i,j}\cdot P_{i,j}^{p}-\sum_{m,n,k}C_{f\,m,n,k}\cdot n_{m,n}x_{m,n,k}-\sum_{i,j}\sum_{p\in P_{i,j}}C_{p\,i,j}\cdot P_{i,j}^{p}-\sum_{i,j}Spill_{i,j}\cdot SpillCost_{i,j}\right) \tag{6}$$

where $\overline{Fare}_{i,j}$ represents the average fare between O-D city pair i and j; $P_{i,j}{}^{p}$ represents passenger demand between O-D city pair i and j, on itinerary p; $C_{f\,m,n,k}$ represents average cost per flight on the flight segment between airports m and n, for aircraft type k; $n_{m,n}$ represents the number of airlines operating between airports m and n, and $x_{m,n,k}$ represents average number of flights per day on the flight segment between airports m and n, using aircraft type k, over all airlines operating on the route; $C_{p\ i,j}$ represents average cost per passenger between O-D city pair i and j; $Spill_{i,j}$ presents total spilled passengers between O-D city pair i and j; and $SpillCost_{i,j}$ is the cost to the airline per spilled passenger between O-D city pair i and j. It is assumed that spill cost per passenger is equal to average fare between the O-D city pair. Cities i and j are served by one or more airports m and n respectively.

The objective function is constrained by a system of linear equations describing airline routing and scheduling requirements and limitations, including total O-D demand $D_{i,j}$ between O-D city pair i and j estimated by the demand model (equation 1), average O-D fares $\overline{Fare}_{i,j}$ and total O-D flight frequencies ($n_{i,j}\times x_{i,j}$) between O-D city pair i and j estimated by the airline competition model. A seat constraint also limits the number of passengers served on each flight segment to be less than or equal to the number of seats available; a simplified balance constraint limits the number of flights of each aircraft type departing an airport in a day to equal the number of flights of that aircraft type arriving, and vice versa; and a fleet constraint limits the total hours operated by each aircraft type to be less than or equal to that available in the existing fleet. The fleet constraint may be relaxed in order to estimate what fleet requirements exist, or may be integrated with an aircraft stock model. In this paper the fleet constraint is not applied. It is also assumed that airlines have the option to route passengers directly between their origin and destination, or through a single hub. Multiple connections are not modelled.

[**] Spill is the total number of passengers who want to fly but cannot obtain a reservation due to insufficient capacity provided by the airlines.

Because the models integrated into the flight routing and scheduling model are not all linear, making the objective function and constraints non-linear, the flight routing and scheduling model is solved by iteration in the following variables: O-D passenger demand $D_{i,j}$, average O-D fares $\overline{Fare}_{i,j}$, flight frequencies $x_{n,m}$, average flight arrival delays, and average cost per flight $C_{f\,i,j}$ and per passenger $C_{p\,i,j}$. The iteration procedure is presented in Figure 1. As described by the figure, given an initial estimate of segment flight frequencies $x_{n,m,0}$, average flight delays at each airport are estimated using the flight delay model described in Section II.C. This allows average travel time and average costs per flight and per passenger to be estimated using a simple travel time calculator and the airline cost model described in Section II.D. O-D passenger demand $D_{i,j}$, average O-D fares $\overline{Fare}_{i,j}$, and the minimum flight frequency required for competition can then be estimated using the demand model and airline competition model described in Sections II.A and II.B. The outputs of these models, and the costs per flight and per passenger estimated by the airline cost model allow the network optimization to be run, solving the segment flight frequencies required to maximise the profit function described by equation 6. These flight frequencies can then be used to re-estimate average flight delay. The iteration is repeated until the system flight frequency converges to within 1 flight per day. The converged results yield modelled passenger demand, flight frequency, average fares, and average delays for the system modelled.

Because of the simplifications in each of the sub-models integrated into the flight routing and scheduling model, and the lack of any constraints specific to any one airline, this model is not capable of predicting or advising any one airline's response to capacity constraints. Instead, the model is intended to predict general system response well into the future – as far as 2030 – given alternative airport capacity growth scenarios.

III. Application to Simple Theoretical Networks

The flight routing and scheduling model described in Section II was applied to a series of simple routing networks to illustrate its capabilities and facilitate discussion of the model functionality. Three simple routing networks are modelled:

- Three spoke airports equidistant from a central hub airport, as illustrated in Figure 2(a). This network allows analysis of the effect of capacity constraints at a hub airport on the profit maximising routing through the hub, causing a shift away from a pure hub and spoke network.
- Three spoke airports surrounding two hub airports, as illustrated in Figure 2(b). This network allows analysis of the effect of capacity constraints at a hub airport on the distribution of traffic between that hub and an alternative hub.
- Four spoke airports, two of which serve the same market (forming a multi-airport system), equidistant from a central hub airport, as illustrated in Figure 2(c). This network allows analysis of the effect of capacity constraints at one airport within a multi-airport system on the distribution of traffic between the airports in the system.

The application and results of the application of the model to each of these theoretical networks is described in detail below.

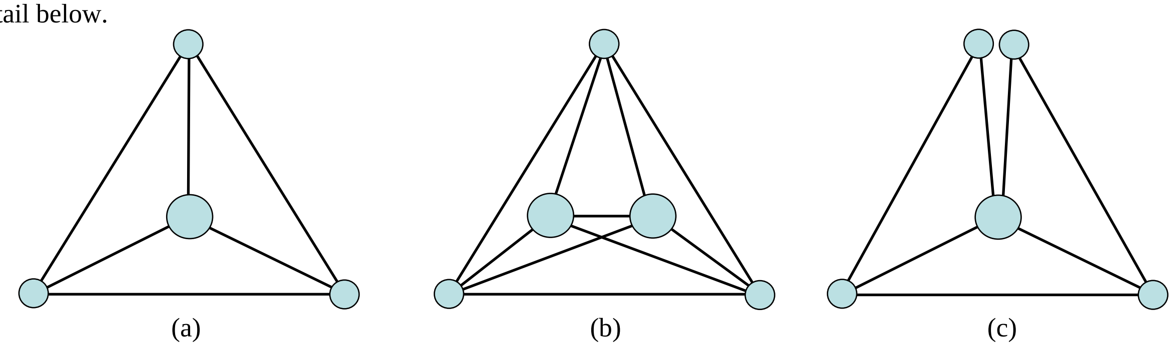

Figure 2. Simple theoretical networks modelled – a) a simple hub and spoke routing network; b) a simple hub and spoke routing network with two hub airports; and c) a simple hub and spoke routing network with one spoke served by a multi-airport system of two airports.

A. Simple Hub and Spoke Routing Network

The input data to the model is presented in Table 1. These inputs are hypothetical, but are based on typical values for the air transport system in the United States in 2005. Non-network traffic accounts for all flights from airports outside the network modeled. Revenues and costs from this non-network traffic are not included in the profit maximization within the flight routing and scheduling model, but the traffic is modeled at each airport to ensure that realistic delays are simulated. In a real network, however, all flights would impact airline profit. Hypothetical populations and incomes were selected to generate the unconstrained demand between the cities presented in the table. Three aircraft types are modeled – a small aircraft type (applying aircraft performance data for a Boeing B737-300), a medium aircraft type (applying aircraft performance data for a Boeing 767-300), and a large aircraft type (applying aircraft performance data for a Boeing 747-400). Aircraft cost data is extracted directly from the DOT Form41 data [21] for 2005. Passenger value of time and cost elasticities are derived from Reynolds et al. [10]. Average gate departure delay due to mechanical failures and late arrivals is extracted from the FAA ASPM database [12] for the 50 busiest airports in the US air traffic system.

Table 1. Input parameters for analysis of simple hub and spoke routing network.

Input Parameters		
Non-network traffic [flts/day]	Hub Airport	2,000
	Spoke Airport	500
Unconstrained demand [pax/yr]	Hub City – Spoke City	200,000
	Spoke City – Spoke City	100,000
Fuel price [2005 US$/kg]		1.73
Aircraft operating cost (excl. fuel) [2005 US$/hr]	Small	2,123
	Medium	2,929
	Large	4,597
Volume related cost [2005 US$/pax]		21.39
Passenger value of schedule delay [2005 US$/hr]		34.24
Passenger value of travel time [2005 US$/hr]		34.24
Passenger Cost Elasticity	Short haul (< 500 miles)	-0.39
	Medium haul (500 to 1000 miles)	-1.76
	Long haul (> 1000 miles)	-2.07
Extra cost term C_t [2005 US$]		50
Hub-spoke stage length [miles]		500
Number of airlines serving each O-D market		5
Average gate departure delay due to mechanicals and late arrivals [min]		4
Average connection time [min]		40

The model described in Section II was run applying the inputs described in Table 1. Three scenarios were solved: applying unconstrained airport capacities at all airports; applying a medium capacity constraint at the hub airport; and applying a severe capacity constraint at the hub airport. Results are presented for each of the scenarios in Table 2 and Figure 3 below.

As can be seen in Figure 3a, the routing network that maximises airline profit in the unconstrained scenario is a pure hub and spoke network, with 11 flights per day between the spoke airports and the hub, and no direct flights between the spoke airports. All the passenger demand between the spoke cities connects through the hub airport. Because of this connection the cost per passenger between the spoke cities is higher than from the hub to the spoke, and therefore the average fare between the spoke cities is higher than to the hub city ($113 versus $90). Delays are 4 minutes at both the hub and spoke airports, corresponding to the average gate departure delay due to mechanicals and late arrivals only, as airport capacity is unconstrained. O-D passenger demand is greater from the spoke cities to the hub than between the spoke cities because of the input population and income data.

Table 2. Results for simple hub and spoke routing network under varying hub airport capacity.

	Unconstrained		Medium Capacity Constraint at Hub		Severe Capacity Constraint at Hub	
	Hub – Spoke	**Spoke – Spoke**	**Hub – Spoke**	**Spoke – Spoke**	**Hub – Spoke**	**Spoke – Spoke**
Airport Capacity [ac/hr]	unconstrained	unconstrained	100	unconstrained	85	unconstrained
O-D Pax Demand [pax/yr]	250,000	100,000	211,000 (-39,000)	85,000 (-15,000)	194,000 (-56,000)	124,000 (+24,000)
Flight Frequency [flts/day]	11	0	7 (-4)	1 (+1)	4 (-7)	2 (+2)
Avg. Arrival Delay [min]	4	4	106	4	172	4
Avg. Fare [2005 US$]	90	113	131 (+41)	131 (+18)	169 (+79)	91 (-22)

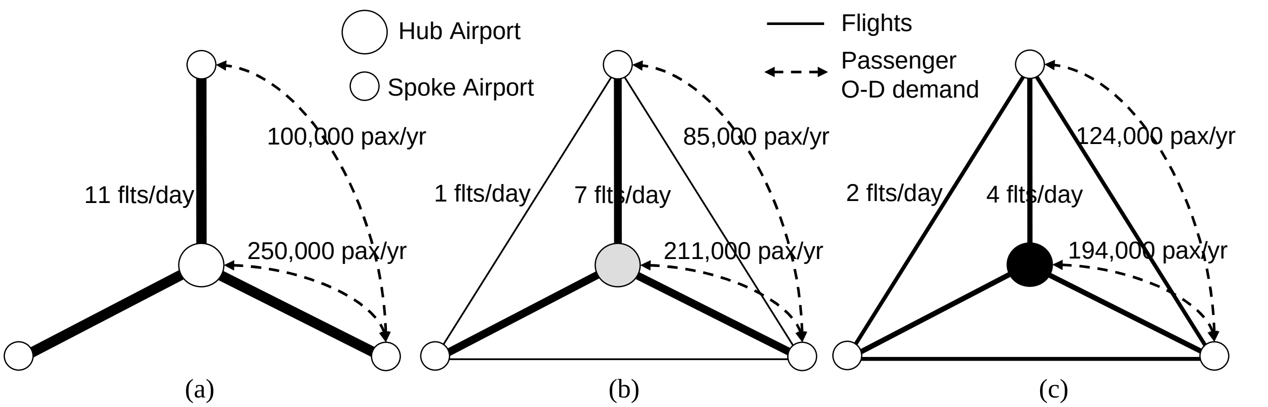

Figure 3. Results for simple hub and spoke routing network – a) unconstrained airport capacities; b) medium capacity constraint at hub; c) severe capacity constraint at hub.

As can be seen in Figure 3b, the routing network that maximises airline profit under a medium capacity constraint at the hub airport shifts to a partial hub and spoke network of 7 flights per day between the hub and spoke airports, but with a single direct flight per day between the spoke airports. Considering firstly the hub to spoke results, O-D passenger demand between the hub and spoke cities reduces from the 250,000 passengers per annum predicted in the unconstrained scenario, to 211,000 in the capacity constrained scenario. This reduction in O-D passenger demand results from an increase in travel time – caused by an average arrival delay of 106 minutes at the hub airport – and an increase in average fare by $41 to $131. Average fare increases because of an increase in costs from the hub to the spoke airport resulting from the increased fuel burn associated with the portion of the 106 minute average arrival delay incurred on active taxiways and in airborne holding at the hub. Because of the reduced O-D passenger demand on the route, the number of aircraft operated is reduced by 4 flights per day to 7 flights per day.

Considering the results for the traffic between the spoke airports, O-D passenger demand between the spoke cities reduces from the 100,000 passengers per annum predicted in the unconstrained scenario, to 85,000 in the capacity constrained scenario. This reduction in O-D passenger demand is for the same reasons as the reduction in demand between the hub and spoke cities – an increase in travel time caused by the average arrival delay of 106 minutes at the hub airport, and an increase in average fare by $18 to $131. The increase in travel time is still experienced by the majority of passengers travelling between the spoke cities despite the introduction of a direct flight between the spoke cities, because the majority of passengers still connect through the hub airport, where the 106 minutes delay is incurred. Only the 4 minutes of average delay due to mechanicals and late arrivals is incurred at the spoke airports as they remain unconstrained. Average fare between the spoke cities also increases for the same reason as between the hub and spoke cities – although the increased fuel burn is incurred at the hub airport, fare is based on average costs per passenger, and the majority of passengers still pass through the hub airport. However, it is profitable for the airlines to operate a single flight per day on the direct routes between the spoke airports.

As can be seen in Figure 3c the routing network that maximises airline profit under a severely capacity constrained hub airport shifts to a pure point to point network with 4 flights per day between the hub and spoke airports, and 2 direct flights per day between the spoke airports. Technically this would require some of the 5 airlines serving each market to withdraw, but this effect is not modeled. Again considering firstly the hub to spoke results, O-D passenger demand between the hub and spoke cities reduces further to 194,000 passengers per annum in the severely capacity constrained scenario. This effect is a more extreme case of that in the medium capacity constrained scenario, with a significant increase in travel time, caused by a massive average arrival delay of 172 minutes at the hub airport, and an increase in average fare to \$169 – \$79 higher than that in the unconstrained scenario. Average fare increases because of the increased costs of the fuel burn associated with the portion of the 172 minute average arrival delay incurred on active taxiways and in airborne holding at the hub. Because of the further reduced O-D passenger demand on the route, the number of aircraft operated is more than halved to only 4 flights per day.

Considering the results for the traffic between the spoke airports, a significantly different result to the medium capacity constrained scenario is observed. O-D passenger demand between the spoke cities increases beyond the values predicted by the unconstrained scenario – to 124,000 passengers per annum. This is because all passengers are flown directly between the spoke airports, with none connecting through the hub. None of the O-D passengers between the spoke airports therefore experience the 172 minutes of delay at the hub to increase their travel time. Travel time is in fact reduced relative to even the unconstrained scenario, as the direct flight time between the spoke airports is less that the combined flight time to and from the hub airport, and passengers do not experience the connection time at the hub. Average fare is reduced to below the values in the unconstrained scenario for the same reason – there is no increase in fuel burn incurred for passengers flying between the spoke airports, as all passengers fly on the direct flights between the spoke airports, which only experience the 4 minutes of average delay due to mechanicals and late arrivals. 2 flights per day are required to serve all the O-D demand between the spoke cities.

The simple theoretical network modelled in this example shows how the routing network selected by an airline to maximise its profits can change significantly as airports become capacity constrained – particularly a hub airport in this case. It should be noted, however, that the capacity constraint at the hub airport must be very severe (an average arrival delay of 172 minutes is well above any average arrival delay currently experienced in any airspace system globally) for airlines to shift to a fully point to point network because of the costs associated with delays only. Airlines may shift to this kind of network for other reasons, however, such as increasing demand between spoke cities.

B. Distribution of Traffic between Hubs

The second theoretical routing network modelled – a system of three spoke airports surrounding two hub airports – illustrates the capability of the model to distribute traffic between hub airports with different capacity constraints. The model takes the same inputs as the simple hub and spoke routing network described above in Table 1. The hub airports are located 300 miles apart and relative to the spoke airports as shown in Figure 4. Two scenarios were solved – an unconstrained scenario in which both hub airports are unconstrained; and a constrained scenario, in which one of the hub airports is capacity constrained. Results for both scenarios are presented in Figure 4.

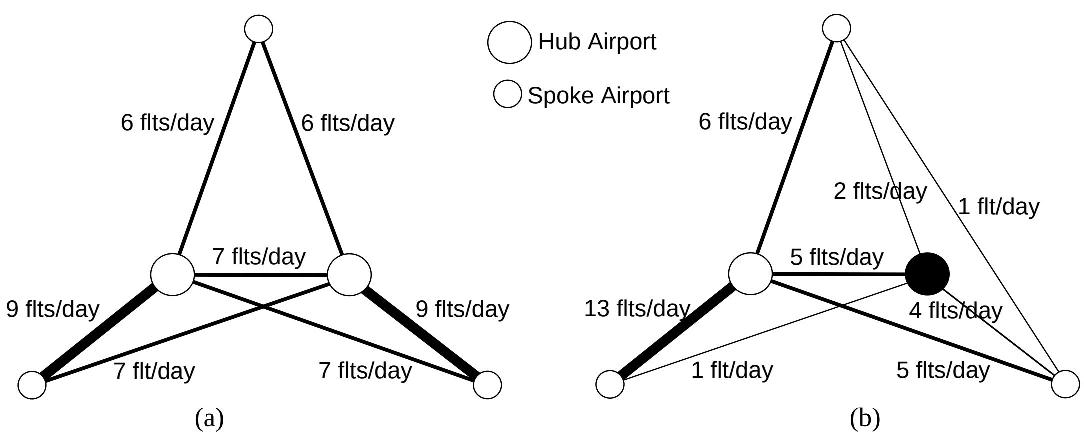

Figure 4. Results for distribution of traffic between two hubs – a) both of which are unconstrained; and b) one of which is capacity constrained.

In the results for the unconstrained scenario, presented in Figure 4a, the traffic between the spoke airports is distributed symmetrically between the two hub airports in a pure hub and spoke network. An equal number of flights are scheduled to each hub when the spoke airport is equidistant from both hub airports (the top spoke airport in Figure 4a), while the majority of flights are scheduled to the closer hub when one hub is closer than the other (the two lower spoke airports in Figure 4a). This latter effect is due to the lower costs associated with flying the shorter distance to the closer hub.

In the results presented in Figure 4b, where one of the hubs (the right hub) is capacity constrained, while the other hub remains unconstrained, there is a shift of traffic from the constrained hub to the unconstrained hub. Thus, even though the unconstrained hub is further from one of the spoke airports (the lower right spoke airport), all the connecting traffic from this airport is flown through the unconstrained hub. This is because of the increased fuel burn costs associated with the delays at the capacity constrained airport, and the increased travel time through the airport, which would reduce passenger demand. There are still direct flights from the spoke airports to the capacity constrained airport serving the O-D passenger demand between these cities. Another further effect of the capacity constraint at one hub is the introduction of a direct flight between two of the spoke cities (the spoke cities furthest from the unconstrained hub), allowing O-D traffic between these cities to by-pass both hubs. This effect is caused by the increased costs associated with travel between these two spoke cities – either the delayed route through the constrained hub, or the longer route through the unconstrained hub. These increased costs make a direct flight more attractive to the airline economically.

This example illustrates the capability of the model to model airline decisions to distribute traffic between hubs, based on the delays experienced at each hub. It is noted, however, that a number of factors that affect an airline's decision in selecting hubs for connecting flights are not captured in this model, including existing airline presence at the hub airport and incentives provided by the airport authority. The model may however indicate at what delay levels it becomes economically attractive for airlines to switch operations to alternative hubs.

C. Distribution of Traffic in a Multi-Airport System

The third and final theoretical routing network modelled – a simple hub and spoke routing network with one spoke served by a multi-airport system of two airports – illustrates the distribution of traffic between airports in a multi-airport system with different capacity constraints. The model takes the same inputs as the simple hub and spoke routing network described above in Table 1. Three scenarios are compared – an unconstrained scenario in which all spoke airports are unconstrained; a partially constrained scenario in which one of the airports in the multi-airport system is capacity constrained; and a second constrained scenario in which both airports in the multi-airport system are capacity constrained, but to different degrees. Results for all three scenarios are presented in Figure 5.

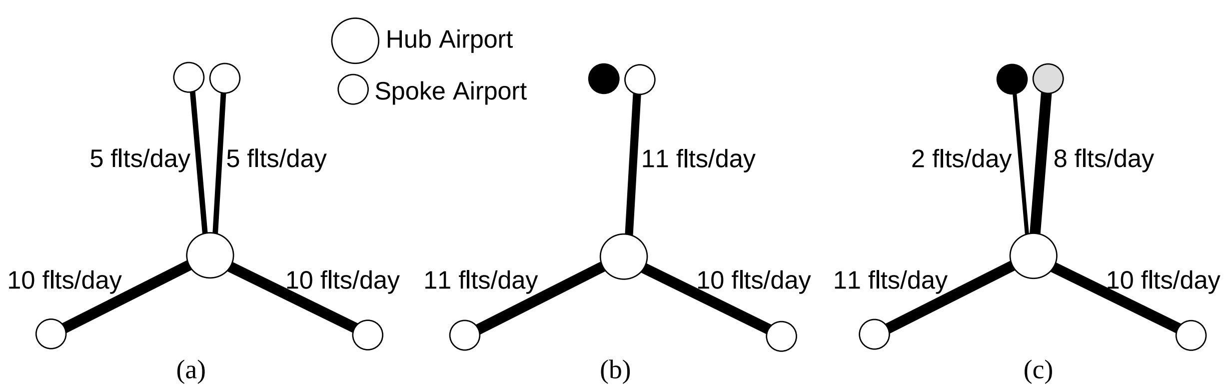

Figure 5. Results for distribution of traffic in a multi-airport system of two airports – a) both of which are unconstrained; b) one of which is capacity constrained; and c) both of which are capacity constrained, but to different degrees.

In the results for the unconstrained scenario, presented in Figure 5a, the traffic between the airports in the multi-airport system is distributed equally between the two airports, with the rest of the system forming a pure hub and spoke network. Traffic is equally distributed because the costs at each airport are identical.

In the results presented in Figure 5b, where one of the airports in the multi-airport system is capacity constrained (the left airport) while the other remains unconstrained, all the flights from the hub airport are routed to the unconstrained airport. The constrained airport does not serve any traffic in the network. Instead the system forms a pure hub and spoke network with the unconstrained airport in the multi-airport system forming a spoke in the system. All O-D passenger demand from the multi-airport city is routed through the unconstrained airport. This is again because of the increased fuel burn costs associated with the delays at the capacity constrained airport, and the increased travel time through the airport, which would reduce passenger demand. It is noted that because each airport has exogenously specified non-network traffic, there are still high delays at the constrained airport, even though no traffic operates there in the network modeled.

In the results presented in Figure 5c, where both of the airports in the multi-airport system are capacity constrained, although one (the left airport) is slightly more constrained than the other, the flights from the hub airport to the multi-airport city are distributed between the two airports. The more constrained airport receives less traffic than the less constrained airport because of the higher fuel burn costs associated with delays, and the increased travel time through the airport, which would reduce passenger demand. All the traffic is not routed to the less constrained airport because, with this increase in traffic delays would increase, increasing costs. The routing shown in Figure 5c is the optimal distribution of flights between the airports, yielding maximum profits.

This example illustrates the capability of the model to model airline decisions to distribute traffic between airports in a multi-airport system, based on the delays experienced at each airport. It is noted that a number of factors not captured in this model also affect an airline's choice of airports in a multi-airport system, such as proximity to urban areas, accessibility, facilities, and existing airline presence at an airport.

IV. Application to Actual Network

The flight routing and scheduling model described in Section II was also applied to a network of 16 airports serving the 10 cities with greatest passenger demand in the United States in 2005[††], allowing the model's predictive capability to be identified. The model inputs presented in Table 1 were applied, with the exception of non-network traffic, unconstrained demand, stage lengths, and average gate departure delay due to mechanicals and late arrivals, which were modelled according to 2005 data from the DOT T100 database [28] and FAA ASPM database [12]. The model was run and the converged results compared to actual 2005 data [20,21,28]. The model deviations from the observed values are presented in Table 3 below.

Table 3. Results for model application to actual network.

	Deviation from Observations 2005
Mean O-D pax demand by O-D market	6% high (max 59%)
Total O-D system pax demand	12% high
Mean O-D fare by O-D market	16% low (max 93%)
Total system revenue	11% low
Mean flight frequency by segment	34% low (max 760%)
Total system flight frequency	61% low

As shown in Table 3, O-D passenger demand is over predicted by 6% on average per O-D market (with a maximum of 59%), and 12% for the whole system. This result is primarily because of an under prediction of average fares by 16% on average per O-D market (with maximum of 93%), which is caused by an under prediction of airline costs. These results combine to under predict system revenue by 11%.

Airline costs are under predicted because the model does not capture all constraints on airline operations. This includes particularly constraints on airline selection of aircraft types, such as requirements to operate the same aircraft types on multiple routes in order to provide flexibility to swap equipment or crews as needed, and to reduce maintenance costs (requiring maintenance facilities for only a few aircraft types). These constraints significantly limit airline fleet choice. This effect is also the cause of the significant under prediction of flight frequency in Table

[††] Cities (and airports) modeled include: New York City (JFK, EWR and LGA), Chicago (ORD and MDW), Atlanta (ATL), Washington DC (IAD, DCA), Los Angeles (LAX), Dallas/Fort Worth (DFW and DAL), Houston (IAH and HOU), Detroit (DTW), Phoenix (PHX), and Seattle (SEA).

3, by 34% on average per route (with a maximum of 760%) and by 61% for the system as a whole, despite the over prediction of passenger demand. The under prediction of flight frequency is because the majority of aircraft selected by the model are medium sized aircraft. This aircraft size class offers lower costs per passenger kilometre than small aircraft, and demand is high enough to operate them at high load factor. However, the majority of aircraft operated in the 2005 US fleet are small, such as the Boeing B737 series aircraft and Airbus A319/320/321 series aircraft. These aircraft are selected by airlines because of their flexibility to operate on many different routes.

Some flight frequency results show very high deviations from observed traffic. This is particularly the case on flight segments between airports within multi-airport systems. This includes the segment with maximum deviation (760%) which is between Dallas/Fort Worth International airport (DFW) and Newark Liberty International airport (EWR), both of which are part of multi-airport systems. These high deviations reflect the factors not captured in the model which have a significant impact on airline choice of airports in a multi-airport system, including particularly existing airline presence at an airport. DFW and EWR are established hubs for American and Continental Airlines respectively. In the modelled results the majority of the demand between Dallas/Fort Worth and New York City is instead routed through the other airports in the respective multi-airport systems.

The results presented indicate that other airline constraints, including specifically fleet constraints and existing airline presence at airports, need to be included in the model for more accurate prediction of air traffic in the US air transport system. Further development of the model is therefore required.

V. Conclusions

A model is under development that predicts airline routing and scheduling under airport capacity constraints by integrating models of passenger demand, airline competition, flight delay, aircraft cost and flight routing and scheduling. The model is applied to a series of theoretical routing networks, and to a network of airports in the United States with 2005 population, income and airport capacity inputs.

The application of the model to the theoretical routing networks illustrates the capability to simulate airline responses to airport capacity constraints through adjusting their schedules and routing networks to maximise profit in a competitive environment. These responses include altering the traditional hub-and-spoke network structure, distributing traffic between alternative hubs and alternative airports within a multi-airport system.

The model has also been applied to an actual routing network of 10 cities and 16 airports in the United States in 2005. The model is found to under predict flight frequencies as it selects larger aircraft than are typically operated. This is because the model does not capture the airline choice of aircraft to provide flexibility over routes and crews, and to reduce maintenance costs.

The model will be further developed to more realistically simulate airline choice of aircraft types and of airports within multi-airport systems, which will improve its forecasting of air traffic growth in the future. This will allow its application in the Aviation Integrated Modelling (AIM) project under development at the University of Cambridge to forecasting air traffic growth, including network and schedule changes resulting from capacity constraints.

Acknowledgements

This work was funded through the AIM grant from the UK Engineering and Physical Sciences Research Council (EPSRC) and the Natural Environment Research Council (NERC). Their support is gratefully acknowledged. The authors would also like to thank colleagues in the AIM group and the wider Institute for Aviation and Environment (IAE) at the University of Cambridge, and Professors Cynthia Barnhart and Amedeo Odoni at the Massachusetts Institute of Technology for helpful discussions, and especially María Vera-Morales for supplying aircraft fuel burn rates.

References

[1] United Nations Intergovernmental Panel on Climate Change, 1999. “IPCC Special Report – Aviation and the Global Atmosphere, Working Group I and III,” IPCC, Geneva, Switzerland.
[2] International Civil Aviation Organisation, 2006. “World Total Traffic 1995 – 2006”, www.icaodata.com.
[3] Airbus, “Global Market Forecast 2007-2026”, http://www.airbus.com/fileadmin/documents/gmf/PDF_dl/00-all-gmf_2007.pdf, 2007 [Cited 27 June 2008]
[4] Boeing, “Current Market Outlook 2007”, http://www.boeing.com/commercial/cmo/pdf/Boeing_Current_Market_Outlook_2007.pdf [Cited 27 June 2008].

[5] Berghof R. et al., 2005, CONSAVE 2050 – Executive Summary, Competitive and Sustainable Growth, http://www.dlr.de/consave/CONSAVE%202050%20Executive%20Summary.pdf
[6] Schäfer A., 2006, "Long-Term Trends in Global Passenger Mobility," The Bridge, The National Academies Press, Washington, D.C., Vol. 36, pp. 24-32.
[7] Hancox R., S. Lowe, 2000. "Aviation Emissions and Evaluation of Reduction Options (AERO) – Air Transport Demand and Traffic Model (ADEM), Appendix J – Forecasting Aircraft Movements", MVA Limited.
[8] Bhadra D., J. Gentry, B. Hogan, M. Wells, 2003. "CAASD's Future Air Traffic Estimator: A Micro-Econometric Approach", 13th Annual Federal Forecasters Conference, BLS Training Center.
[9] Bhadra D. 2003. "Choice of Aircraft Fleets in the US NAS: Findings from a Multinomial Logit Analysis", 3rd Annual Technical Forum of the ATIO/AIAA, Denver, CO.
[10] Reynolds T., S. Barrett, L. Dray, A. Evans, M. Köhler, M. Vera-Morales, A. Schäfer, Z. Wadud, R. Britter, H. Hallam, R. Hunsley, 2007. "Modelling Environmental & Economic Impacts of Aviation: Introducing the Aviation Integrated Modelling Tool", 7th AIAA Aviation Technology, Integration and Operations Conference, Belfast, 18-20 September 2007.
[11] Cairns S., C. Newson, B. Boardman, J. Anable, 2006, "Predict and Decide – Aviation, Climate Change and UK Policy, Final report," Environmental Change Institute, University of Oxford, Oxford, UK.
[12] United States Federal Aviation Administration, Operations and Performance Data, Aviation System Performance Metrics, http://www.apo.data.faa.gov/aspm/ASPMframe.asp
[13] EUROCONTROL, 2007, "Trends in Air Traffic, Volume 2, A Matter of Time: Air Traffic Delay in Europe", EATM CODA, EUROCONTROL, https://extranet.eurocontrol.int/http://prisme-web.hq.corp.eurocontrol.int/ecoda/coda/public/standard_page/public_application.html.
[14] USDOT, FAA, Mitre CAASD, 2004, "Airport Capacity Benchmarking Report 2004", Office of the Administrator, United States Department of Transport, 800 Independence Ave. S.W., Washington, DC, 20591, http://www.faa.gov/about/office_org/headquarters_offices/ato/publications/bench/2004download.htm, September 2004.
[15] Kostiuk P.F., D. Lee, D. Long, 2000. "Closed Loop Forecasting of Air Traffic Demand and Delay," 3rd USA/Europe Air Traffic Management R&D Seminar, Napoli, Italy.
[16] Long D., E. Wingrove, D. Lee, J. Gribko, R. Hemm, P. Kostiuk, 1999. "A Method for Evaluating Air Carrier Operational Strategies and Forecasting Air Traffic with Flight Delay," Logistics Management Institute, NS902S1, McLean, VA.
[17] Schipper, Y., P. Rietveld, P. Nijkamp, 2003. "Airline deregulation and external costs: a welfare analysis", Transportation Research Part B 37 (2003) 699-718.
[18] Carlsson, F., 2002. "Price and Frequency Choice under Monopoly and Competition in Aviation Markets," Working Paper in Economics no. 71, Department of Economics, Göteborg University, Gothenburg, Germany.
[19] US Census Bureau, Census 2000, 2000, www.census.gov/main/www/cen2000.html [cited 10 September 2007].
[20] US Department of Transportation, "DB1B Survey," Bureau of Transportation Statistics, www.transtats.bts.gov [cited 10 Sept. 2007].
[21] US Department of Transportation, 2005, "Air Carrier Financial Reports (Form 41 Financial Data)," US Department of Transport, Research and Innovative Technology Administration, Bureau of Transportation Statistics, Washington DC, USA.
[22] Evans, A.D., 2008, "Rapid Modelling of Airport Delay", 12th Air Transport Research Society World Conference, Paper 202, Athens, Greece, 6-9 July 2008.
[23] de Neufville, R., A. Odoni, 2003. "Airport Systems Planning, Design, and Management," McGraw Hill Companies, Inc., ISBN 0-07-138477-4.
[24] Belobaba P.P., 2006, "Operating Costs and Productivity Measures," Course 16.75J / 1.234J Airline Management, http://ocw.mit.edu/OcwWeb/Aeronautics-and-Astronautics/16-75JSpring-2006/CourseHome/index.htm, Massachusetts Institute of Technology, Cambridge MA, USA.
[25] EUROCONTROL, "Base of Aircraft Data (BADA)", Version 3.6, July 2004.
[26] International Civil Aviation Organisation, 2007, "ICAO Aircraft Engine Emissions Databank," http://www.caa.co.uk/default.aspx?catid=702&pagetype=90, July 2007 [cited 10 June 2008].
[27] Harsha, P., 2005. "Auctions for Airport Landing Slots – The Bidder Problem", Research Oriented Paper, Massachusetts Institute of Technology, Cambridge, MA, USA.
[28] US Department of Transportation, "T100 Traffic & Financial Data," Bureau of Transportation Statistics, www.transtats.bts.gov [cited 10 Sept. 2007].