

Airline Planning Benchmark Problems

Part I: Characterising Networks and Demand using Limited Data

Kerem Akartunali^a, Natasha Boland^b, Ian Evans^c, Mark Wallace^d, Hamish Waterer^{b,d,*}

^aManagement Science Department, University of Strathclyde, Glasgow G1 1QE, Scotland, United Kingdom

^bSchool of Mathematical and Physical Sciences, University of Newcastle, Callaghan NSW 2308, Australia

^cConstraint Technologies International, Level 7 224 Queen St, Melbourne VIC 3000, Australia

^dFaculty of Information Technology, Monash University, Caulfield VIC 3145, Australia

Abstract

This paper is the first of two papers entitled “Airline Planning Benchmark Problems”, aimed at developing benchmark data that can be used to stimulate innovation in airline planning, in particular, in flight schedule design and fleet assignment. While optimisation has made an enormous contribution to airline planning in general, the area suffers from a lack of standardised data and benchmark problems. Current research typically tackles problems unique to a given carrier, with associated specification and data unavailable to the broader research community. This limits direct comparison of alternative approaches, and creates barriers of entry for the research community. Furthermore, flight schedule design has, to date, been under-represented in the optimisation literature, due in part to the difficulty of obtaining data that adequately reflects passenger choice, and hence schedule revenue. This is Part I of two papers taking first steps to address these issues. It does so by providing a framework and methodology for generating realistic airline demand data, controlled by scalable parameters. First, a characterisation of flight network topologies and network capacity distributions is deduced, based on analysis of airline data. Then a multi-objective optimisation model is proposed to solve the inverse problem of inferring OD-pair demands from passenger loads on arcs. These two elements are combined to yield a methodology for generating realistic flight network topologies and OD-pair demand data, according to specified parameters. This methodology is used to produce 33 benchmark instances exhibiting a range of characteristics. Part II extends this work by partitioning the demand in each market (OD pair) into market segments, each with its own utility function and set of preferences for alternative airline products. The resulting demand data will better reflect recent empirical research on passenger preference, and is expected to facilitate passenger choice modelling in flight schedule optimisation.

Keywords: Airline planning, benchmark data, inverse problems

1. Introduction

There has been relatively little work that has addressed the first stage of the airline planning process, namely, flight schedule design. The many algorithms and techniques reported in the literature for later stages of the airline planning process are difficult to compare because they are evaluated on problem instances representative of a particular airline at a particular date. Each airline operates a different network of airports, a different fleet in terms of the size and mix of aircraft, has different passenger quantities and itineraries, and different crew requirements, bases and rules. Furthermore, the data for these instances is considered confidential by most airlines due to its significant commercial implications. Consequently, obtaining real data is difficult and often requires the researcher to establish a good relationship with an airline partner over many years.

Such issues create a barrier to entry for many prospective researchers and limits potentially fruitful collaboration between research groups.

Some first steps towards addressing these issues are taken in this paper and its sequel Akartunali, Boland, Evans, Wallace, and Waterer (2011) by developing a framework for generating realistic benchmark instances. These instances provide standardised data with which to initiate the airline planning process. Since flight schedule design depends critically on market demand, this initial work has focused on the generation of airline demand benchmark data. The addition of airline resources, such as aircraft and crew, to these benchmarks is planned for the future. By making these instances, and a description of the methodology used to generate them, publicly available it is hoped that research engagement in airline planning will be stimulated in a similar way to what has been so successfully achieved in areas such as vehicle routing, which flourished after the introduction of the Solomon benchmark instances (Solomon, 1987). The DIMACS¹ and ROADEF² challenge instances have

*Corresponding author

Email addresses: kerem.akartunali@strath.ac.uk (Kerem Akartunali), natashia.boland@newcastle.edu.au (Natashia Boland), ian.evans@contecint.com.au (Ian Evans), mark.wallace@infotech.monash.edu.au (Mark Wallace), hamish.waterer@newcastle.edu.au (Hamish Waterer)

¹<http://dimacs.rutgers.edu/Challenges/>

²<http://challenge.roaDEF.org/>

had a similar impact.

As a large body of literature attests, optimisation has been a critical part of airline planning for many decades. See, for example, Klabjan (2005), Bazargan (2004), or Barnhart and Cohn (2004). However, as noted in Klabjan (2005), for the most part, airline schedule planning is a manual process with only a few manuscripts on flight schedule design. Notable among these are two papers, Yan and Tseng (2002) and Yan, Tang, and Lee (2007), on flight scheduling in Taiwan, and that of Lohatepanont and Barnhart (2004), combining flight scheduling with fleet assignment. The authors of this paper believe that the dearth of optimisation research on schedule design is in part due to the difficulty of representing passenger choice, and of collecting adequate data to accurately assess schedule revenue. However, there has been a growing body of both empirical and theoretical research seeking to provide insight into airline passenger decision processes and to develop models of passenger utility. See, for example, Coldren, Koppelman, Kasturirangan, and Mukherjee (2003), Garrow, Jones, and Parker (2007), Koppelman, Coldren, and Parker (2008), Walker (2006), and Wojahn (2002). The insights provided in these papers, combined with an empirical analysis of rich data sets from a wide range of airlines worldwide, including all airlines in the Star and oneworld alliances, has led to the development of a new approach to representing airline demand data, and a methodology for generating realistic demand data sets.

The methodology developed in these two papers is a four step framework. Figure 1 illustrates the four steps in the framework which are:

1. Generate the flight network including passenger load on arcs;
2. Calculate origin-destination (OD) pair demand;
3. Define passenger groups;
4. Allocate OD-pair demand to each passenger group.

The flight network connects the set of airports to be served, and the network topology defines arcs indicating airport pairs between which direct non-stop services are to be offered. Passenger load on an arc indicates the total number of passengers expected to travel on the direct non-stop service over some time period, for example, a day.

This paper presents the methodology behind the first two steps in this framework. The first step generates realistic flight networks and passenger loads with specified characteristics that capture the features of a large fraction of existing airline networks. These networks are scalable so that the effect of different scheduling strategies, and different parameters such as network type and size, or fleet mix, on algorithm performance and solution cost can be readily compared. The second step of the framework solves an inverse problem to determine OD-pair based demand that is compatible with the passenger loads on each arc. This data, sometimes called *market demand*, can be of use in its own right. For example, in performing schedule design, Yan and Tseng (2002) work directly from such data collected from airlines in Taiwan.

The second paper (Akartunali et al., 2011) presents the methodology behind the third and fourth steps in this framework. The

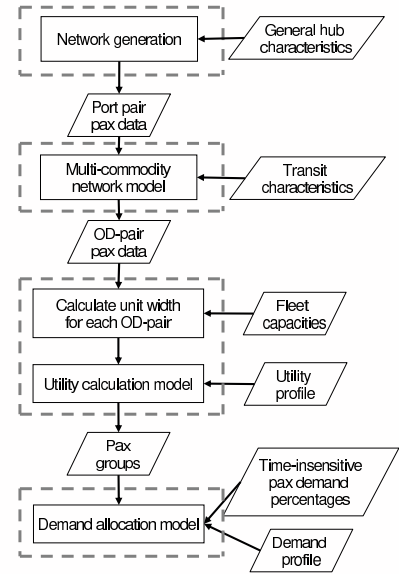


Figure 1: Framework for generating sets of realistic airline planning benchmark problem instances

third step partitions the market demand into *passenger groups*, according to characteristics that differentiate behaviour in terms of airline product selection. Each passenger group has an origin, a destination, a size (number of passengers), a departure time window, and a departure time utility curve indicating the passengers’ willingness to pay for departure in time sub-windows. This data is much richer than simple market demand and can be expected to provide better estimates of schedule revenue in a form that is useful in schedule design optimisation. The integrated airline schedule design and fleet assignment problem studied in a companion paper (Akartunali, Boland, Evans, Wallace, Waterer, and Smith, 2010) demonstrates passenger groups to be a potential alternative to the commonly used spill models (Dumas and Soumis, 2008; Jacobs, Smith, and Johnson, 2008; Barnhart, Farahat, and Lohatepanont, 2009) for estimating passenger flow in an airline network. The fourth step in the framework allocates the previously determined OD-pair demand to each passenger group using a standardised *demand profile*, a generic percentage-wise allocation of passengers throughout a day.

The design of this methodology readily permits the generation of realistic airline data “from scratch” in a way that supports experimentation with key characteristics of that data, as well as providing an approach that other researchers can still use when they have access to partial data. For example, if an existing flight network is already known, and, perhaps, observed passenger loads are also known for that network.

1.1. Terminology, notation, and assumptions

An airline network consists of a set S of *airports* to be served, and a set $A \subseteq S \times S$ of directed *arcs* indicating an ordered pair of airports between which at least one direct non-stop

service is offered. An airline's *fleet* is denoted by the set F of aircraft subtypes, and an aircraft from the fleet $f \in F$ has capacity c_f . The basic time unit used is one day. Let \mathcal{T} denote the length of a day in minutes.

Let $\bar{K} \subset S \times S$ denote the set of ordered potential passenger *origin-destination pairs*, or OD-pairs. For each OD-pair $(o, d) \in \bar{K}$, the *OD-pair demand* D_{od} is the total passenger demand over a day to travel from airport o to airport d . In the case that an OD-pair is not an arc then the only way passengers can complete their travel is to connect to successive arcs by *transiting* at an intermediate airport. The passenger load n_{ij} on arc $(i, j) \in A$ is the number of passengers observed traversing the arc over the course of a day.

1.2. Overview of the paper

Section 2 describes methodology to characterise airline networks. This is the first step in the framework for generating sets of realistic airline planning benchmark problems. Section 3 describes methodology to characterise airline demand using limited data. This is the second step in the framework. A description of the generated benchmark instances is provided in Section 4 and an analysis of the instances is given in Section 5. Section 6 presents some conclusions and a brief description of future work.

2. Characterising airline networks

An airline network's topology depends on factors such as the geographical positions of the airports serviced by the network, the desired operating practices of the airline, the structure of the network of any competitors, and also passenger demand. This paper concentrates on the commonly occurring *hub-and-spoke* topology. This topology consists of a single *hub* airport connected by flight legs to a number of *spoke* airports. The spoke airports are only connected to the hub, that is, no flight legs exist between spoke airports.

Evans, Wallace, and Waterer (2010) analysed data from a wide range of airlines worldwide, including all airlines from Star and oneworld alliances, and found that more than 80% of airports are connected in topologies that resemble hub-and-spoke networks. Thus, analysing such networks is a critical first step. Moreover, more complex topologies such as those consisting of linked hubs, or point-to-point networks, require significantly more analysis.

2.1. Hub-and-spoke networks

The characteristics of hub-and-spoke networks were analysed by Evans et al. (2010) using data collected from the schedules operating at the end of 2007 and early 2008 for a wide range of airlines worldwide, including all airlines in the Star and oneworld alliances. Legs included in the analysis were restricted to those operated by common turbo-fan aircraft. The aircraft capacity on each leg was used as a surrogate for passenger load due to the lack of actual passenger data. This section provides an overview of this characterisation.

Of the 64 hub-and-spoke networks included in the analysis, 41 were classified as short-haul networks as there were no arcs with a great-circle distance greater than 5000km, 10 were classified as long-haul as there existed arcs with a great-circle distance greater than 9500km, and the remaining 13 networks were classified as medium-haul. The statistical analysis of these networks focused on characterising three distributions. The first was the greater-circle length of the network arcs, or the *arc distance*. The second characteristic was the capacities of the aircraft operating within the network, or the *arc capacity*. Finally, the third characteristic was the radial direction of the arcs and their associated capacity, or the *directional capacity*.

The analysis found that the distributions of arc distance for most networks could be clustered into five groups. Each distribution in a particular group was found to be statistically most similar to the other distributions in that group. Two of these groups corresponded to short-haul networks, two to long-haul, and the remaining group to medium-haul. A simple analytical model of an arc distance cumulative distribution function (CDF) was constructed for each group. The CDF was constructed so that it was a good fit to all of the arc distance distributions in that group.

Using a similar analysis, the distributions of arc capacity for most networks could also be clustered into five groups. Similarly, a simple analytical model of an arc capacity CDF was constructed for each group. An analysis of the correlation between arc distance and arc capacity CDFs showed that each arc distance CDF was strongly correlated to one of only two arc capacity CDFs. Short- and medium-haul networks shared the same arc capacity CDF while long-haul networks shared another arc capacity CDF.

To analyse the directional capacity of each network the arcs were partitioned into radial 15-degree sectors. The first sector included spokes radiating from the hub at angles $[0^\circ, 15^\circ)$, where angles were measured anticlockwise from due east, with the remaining sectors continuing in an anticlockwise direction. The capacity of a sector is given by the sum of the capacities of the network arcs in that sector.

The analysis found that there is often a major axis along which most capacity is concentrated, and this is most often closer to an East-West orientation than a North-South orientation. The position of the *greater lobe* of the major axis was defined to be the angle central to the four contiguous sectors with maximum total arc capacity. The position of the *lesser lobe* of the major axis was defined to be the angle central to the four contiguous sectors with maximum total arc capacity subject to the angle being at least 90° from the angle of the greater lobe. These lobes can range from being close to symmetrical to being extremely asymmetrical, that is, in some cases almost all capacity occurs in a lobe towards a single direction from the hub, with only a small amount of capacity being grouped in a lobe in an opposing direction. The capacity of the minor axis was taken as the capacity of all slices that were not contained in the greater or lesser lobes of the major axis. The classification of the directional capacity is illustrated in Figure 2.

Two parameters are used to measure the distribution of the network's directional capacity. The parameter $R_{\text{minor}, \text{major}}$ is de-

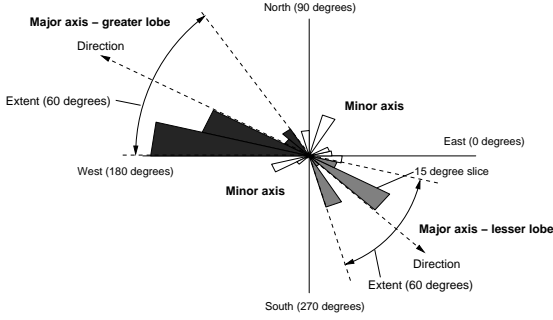


Figure 2: Classification of a hub-and-spoke network's directional capacity (Evans et al., 2010)

finied to be the ratio of the capacity of arcs that are not along the major axis to those that are along the major axis. The parameter $R_{\text{lesser-greater}}$ is defined to be the ratio of the capacity in the lesser lobe of the major axis to the capacity in the greater lobe of the major axis. An analysis of the correlation between each arc distance CDF and the parameters measuring directional capacity showed that networks have a range of geometries and that this range varies for each arc distance CDF. Two pairs of directional capacity parameters that best represented the networks of each arc distance CDF were chosen.

The scheduled time for an aircraft leg between push back from the originating airport gate and arrival at the destination airport gate is known as the *block time*. This time is the sum of the taxi time on departure, flight time, and taxi time on arrival. Taxi times are relatively constant, and the flight time is approximately a linear function of the arc distance plus extra time involved in climbing and descending at lower speed. Prevailing westerly winds mean that flight times for arcs directed west to east are normally less than those for arcs directed from east to west. Arcs are categorised into two groups depending upon whether the travel direction is east-west or west-east. A linear model of block time in minutes as a function of distance in kilometres is fitted to each group.

To model time zone effects, a time zone offset is applied to airports that are at a large east-west distance from the hub. It is assumed that the hub is positioned at the equator at the centre of a one hour time zone. Assuming the mean radius of the Earth is 6371km, the width of each one hour time zone at the equator is 1668km.

In Section 4 it is explained briefly how this analysis is used to generate flight network topologies and arc passenger loads, from given parameters; full details can be found in Evans et al. (2010).

2.2. Airline network data

An airline network topology and passenger loads is insufficient to deduce likely OD-pair demands. Additional exogenous information that can be readily estimated from existing airline network data is required. Two characteristics of such data are the fraction of passenger load arising from single-leg travel, and the fraction of transiting multi-leg passengers arriving at an airport on an incoming arc and connecting to an outgoing arc.

Type	Example	θ
Point-to-point	Australia Domestic	0.77
Point-to-point	Southwest Domestic	0.64
Hub	United to/from LAX	0.58
Hub	United to/from ORD	0.34
Hub	United to/from DEN	0.31
Heavy Hub	Delta to/from ATL	0.19

Table 1: Typical values of θ_{ij} for a variety of airlines and networks (Evans and Waterer, 2011)

Single-leg passenger fractions. Single-leg passenger fraction data was analysed by Evans and Waterer (2011) using data from one of the main Australian domestic carriers and data from the US Department of Transportation Origin and Destination Survey DB1BCoupon data set for Q1 2008 (US DOT, 2008). The range of carriers in the US were selected to include examples that should have quite different characteristics. For example, Southwest is known as a point-to-point carrier, United has several moderate to large hubs, and Delta has a main very large hub in Atlanta. Only purely domestic legs were included in the Australian data, and only legs purely in the lower 48 states were included in the US data (although connections to and from international legs were included). Furthermore, only legs with at least 5% of the passenger capacity of the busiest legs were included to minimize data sampling variation.

Let θ_{ij} denote the expected proportion of single-leg passengers on arc $(i, j) \in A$. Evans and Waterer (2011) suggests that, in general, the values of θ_{ij} and θ_{ji} are likely to be very similar, and depend upon the topology of the network. For example, typical values for θ_{ij} range from around 0.75 for arcs in a point-to-point network, to 0.4 for arcs in a hubbing network, and this value decreases the larger the hub airport the arc is incident to. Table 1 gives typical values of θ_{ij} for a variety of airlines and networks.

Transit passenger fractions. The percentage of passengers transiting through an airport who connect to a given other arc in the airline network is strongly influenced by the geographical location of the airports with respect to each other. To calculate OD-pair passenger demand from the passenger load data, knowledge of these passenger *transit fractions* is needed.

Through value, the desirability to passengers of being able to stay on the same aircraft rather than having to change aircraft at the connecting airport reflects the likely benefit to the airline if the connection can be advertised as a one-stop service. Through values are used in later stages of the airline planning process such as aircraft rotation, or through assignment (see e.g. Clarke, Johnson, Nemhauser, and Zhu, 1997; Barnhart, Boland, Clarke, Johnson, Nemhauser, and Shenoi, 1998; Gopalan and Talluri, 1998; Ahuja, Goodstein, Mukherjee, Orlin, and Sharma, 2001; Sherali, Bish, and Zhu, 2006). While through values are based on flight to flight connections, they are clearly related to the likelihood of a connection being used by passengers, and so are closely related to transit passenger fractions.

Hub Code	Hub City	Airline	Spokes	Paths
ATL	Atlanta	Delta Air Lines	155	10999
CLT	Charlotte	US Airways	102	4908
CVG	Covington	Delta Air Lines	110	3116
DEN	Denver	United Air Lines	126	5434
ORD	Chicago	United Air Lines	114	5601
PHL	Philadelphia	US Airways	82	2584
PHX	Phoenix	US Airways	73	2748
SLC	Salt Lake City	Delta Air Lines	95	2880

Table 2: Hubs used for transit passenger fraction analysis (Evans and Waterer, 2011)

Let Δ_{ij} denote the great-circle distance between any two airports $i, j \in S$, and note that $\Delta_{ii} = 0$. Let Γ_{ij} denote the maximum ratio of the direct distance between airports i and j and the total distance a passenger would travel between these airports. Note that there may not be any flights between the airports i and j , that is, $(i, j) \notin A$, and there may be multiple indirect routes. The value of Γ_{ij} provides a measure of “directness” of travel between airport i and airport j .

Let $\alpha_{ij'}$ denote the fraction of passengers travelling on arc $(i, j) \in A$ and transiting at airport j that transit to arc $(j, i') \in A$. Transiting passengers are likely to connect to arcs (j, i') when airport i' is in some sense geographically “opposite” airport i . Let $A_{ij}^{\text{out}} = \{(j, i') \in A : \Gamma_{ij'} \geq \gamma\}$ denote the set of arcs corresponding to such *outgoing* connections where the threshold parameter γ is the ratio of the direct distance between any two airports and the expected maximum distance a passenger would choose to travel between these airports. Let $\bar{n}_{ji'}$ denote the relative outgoing passenger load, that is,

$$\bar{n}_{ji'} = \frac{n_{ji'}}{\sum_{(j, i'') \in A_{ij}^{\text{out}}} n_{ji''}}, \quad (i, j) \in A, (j, i') \in A_{ij}^{\text{out}}.$$

Intuitively, if either $\bar{n}_{ji'}$ or $\Gamma_{ij'}$ is very low, then it is likely $\alpha_{ij'}$ will be low as geographical considerations may be mitigated by the apparent desirability of the outgoing connection. However, if $\bar{n}_{ji'}$ and $\Gamma_{ij'}$ are both high, then it is likely that $\alpha_{ij'}$ will also be high.

An analysis of the US Department of Transportation Origin and Destination Survey DB1BCoupon data set for Q1 2008 (US DOT, 2008) was performed. Seven large hubs from three different airlines were chosen as listed in Table 2 to ensure that the number of different choices available for connecting passengers in the analysed data was large. The paths column in this table gives the number of different 2-leg path sections connecting through the given hub on flights operated by the given carrier that were included in at least one passenger itinerary. There is no information in this data to indicate whether travel was on week days or weekends.

The use of a threshold parameter γ to restrict the set A_{ij}^{out} of arcs corresponding to the possible outgoing connections of arc $(i, j) \in A$ thereby eliminating rare connections with low distance measures is justified. It was found that approximately 90% of transiting passengers utilise between 30–60% of the

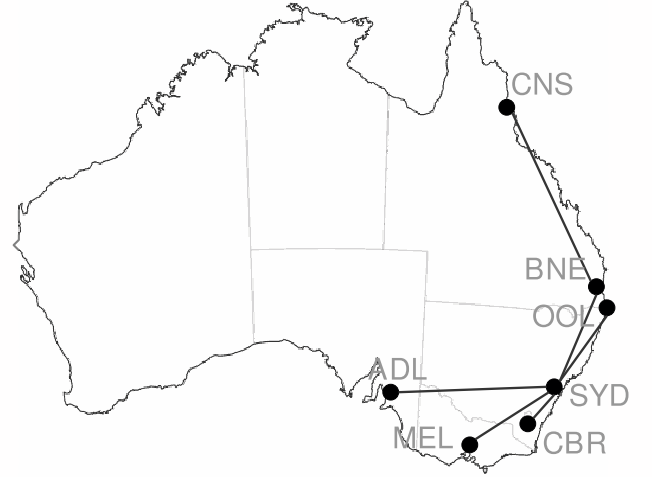


Figure 3: An example hub-and-spoke network

connections existing in the data and that these connections all have distance measures in excess of 0.8. The directness of passenger travel CDF seen in Figure 4 from the ATL airport instance discussed in Section 3.3 shows the fraction of OD-pairs that have at most the given measure of directness, as well as the fraction of combined OD-pair demand, summed over both directions, that experience at most this level of directness.

Various statistical models in which $\alpha_{ij'}$ was related to $\bar{n}_{ji'}$ and $\Gamma_{ij'}$, both independently and as the relative value of some combination of the two, were fitted and tested against the data (Evans and Waterer, 2011). To ensure that incoming arcs with small numbers of connecting passengers did not unduly influence results, only those incoming arcs with at least 200 connecting passengers were included. It was found that the best models had very similar Pearson correlation coefficients, but the best prediction was achieved by the following equation with an average correlation of 0.815.

$$\alpha_{ij'} = \frac{\bar{n}_{ji'}^{0.69} \Gamma_{ij'}^{3.50}}{\sum_{(j, i'') \in A_{ij}^{\text{out}}} \bar{n}_{ji''}^{0.69} \Gamma_{ij''}^{3.50}}, \quad (i, j) \in A, (j, i') \in A_{ij}^{\text{out}} \quad (1)$$

The single-leg passenger fractions and the transit passenger fractions resulting from the above analyses are critical to the deduction of OD-pair demand from observed passenger loads on arcs, a process that is discussed in detail in the next section. To conclude this section an example hub-and-spoke network is introduced. The calculation of transit fractions is illustrated below and the OD-pair demand inferred by our methodology is presented in Section 5.

Example. Consider the hub-and-spoke network given in Figure 3. Table 3 gives the great-circle distance between any pair of airports regardless of whether there exists an arc in the network. Table 4 gives the passenger load on each arc in the network.

The airport SYD is a single hub. Each of the airports ADL, CBR, MEL, and OOL, have a direct connection to only one

Ports	ADL	BNE	CBR	CNS	MEL	OOL	SYD
ADL	-	1621	971	2132	642	1604	1165
BNE	1621	-	956	1392	1381	95	752
CBR	971	956	-	2076	469	892	237
CNS	2132	1392	2076	-	2313	1485	1971
MEL	642	1381	469	2313	-	1329	706
OOL	1604	95	892	1485	1329	-	679
SYD	1165	752	237	1971	706	679	-

Table 3: Great-circle distances in kilometres between any pair of airports in the example network

Ports	ADL	BNE	CBR	CNS	MEL	OOL	SYD
ADL	-	0	0	0	0	0	975
BNE	0	-	0	480	0	0	1477
CBR	0	0	-	0	0	0	486
CNS	0	519	0	-	0	0	0
MEL	0	0	0	0	-	0	776
OOL	0	0	0	0	0	-	222
SYD	1120	1466	538	0	798	214	-

Table 4: Passenger loads for the example network

other airport, namely the hub SYD. Each of these airports and the hub form a simple spoke. The airport CNS and the hub form a spoke with an intermediate stop, namely BNE, which lies approximately on the same flight path.

In the case of a hub-and-spoke network, passengers wishing to travel from one spoke airport to another must transit through the hub. Thus, the percentage of passengers transiting through the hub SYD and connecting to another arc will be high. No passengers arriving at the spoke airports ADL, CBR, CNS, MEL, and OOL, will connect to another arc. There are likely to be a high number of transit passengers at the intermediate stop BNE.

If the distance between spoke airports is much shorter than the total distance to be travelled when connecting via the hub, for example, BNE and OOL, then travellers will not connect between these airports via the hub, instead preferring to use some alternate method of transport.

Consider the arc (OOL, SYD) and suppose $\gamma = 0.5$. The set $A_{\text{OOL,SYD}}^{\text{out}} = \{(\text{SYD}, \text{ADL}), (\text{SYD}, \text{CBR}), (\text{SYD}, \text{MEL})\}$.

	ADL	CBR	MEL
$\bar{n}_{\text{SYD,XXX}}$	0.4560	0.2191	0.3249
$\Gamma_{\text{OOL,XXX}}$	0.8698	0.9738	0.9596
$\bar{n}_{\text{SYD,XXX}}^{0.69} \Gamma_{\text{OOL,XXX}}^{3.50}$	0.3571	0.3196	0.3985

The transit fraction $\alpha_{\text{OOL,SYD,ADL}}$ can be calculated from the relative passenger load and directness values in the above table.

$$\alpha_{\text{OOL,SYD,ADL}} = \frac{0.3571}{0.3571 + 0.3196 + 0.3985} \approx 0.3321$$

3. Characterising airline demand using limited data

Accurate passenger demand data is vitally important to the design of a good airline schedule and to the subsequent fleet assignment and through assignment problems. There are several dimensions to this demand, starting with the origin and destination of the passengers, moving to their preferred time of flying, and ending with the utility of the various products offered by the different airlines or competing modes of transport.

The existence of demand between an OD pair implies that there is a market that has chosen to fly versus using other modes of transport. Major influences on this decision are the travel time from origin to destination, door-to-door, not just that arising from the airlines' levels of service, and the value a passenger places on their time. Also relevant is the trip's purpose, who is paying, and the length of time the passenger will spend away from home. Leisure passengers in particular are more likely to fly as the cost of the airfare decreases relative to the total cost of the trip (Garrow et al., 2007).

Major influences on OD-pair demand can be categorised as either endogenous factors, such as price and levels of competition between airlines; or exogenous factors, such as population and per capita income levels at either the origin, destination, or both, as well as the distance between the origin and destination, and the level of service offered by the airline (see e.g. Dresner and Windle, 1995; Coldren et al., 2003).

Typically airlines only have data about passenger numbers on flights from historical booking records, and this is used to infer each of the associated dimensions for existing routes. For new routes, the airlines must rely on demand forecasts.

OD-pair demand is not precisely captured even by historical booking records, as there are several cases where passengers' actual origins and destinations are not known to the airline. Examples include where infrequent services or bad connections force passengers to break their trip at an intermediate airport, and when passengers book legs separately, or use another carrier for some legs.

Information about the preferred time of flying is also poorly captured in historical booking records. This demand is *censored*, or constrained, by revenue management initiatives, a lack of capacity at peak times, and portions of the network with a low number of services per day (McGill and Van Ryzin, 1999; Lohatepanont and Barnhart, 2004). Techniques for *uncensoring*, or unconstraining, passenger demand have been developed (Bront, Mendez-Diaz, and Vulcano, 2009; Ratliff, Rao, Narayan, and Yellepeddi, 2008), and data aggregation has been used to overcome sparse data due to the existence of a large number of rare itineraries (McGill and Van Ryzin, 1999). These are discussed further in Part II of this paper (Akartunalı et al., 2011).

Whilst a great deal is known about inferring features of OD demand from observed arc demand in road networks (see, for example, Florian (1976)), nothing of this type has yet been attempted for airlines.

3.1. Calculating arc passenger load data

Airlines often have data on historical or forecast passenger numbers for directly connected airport pairs. However, if arc

passenger data is not known, it can often be obtained from data that is available, and averaging data taken from multiple days can reduce the effects of the frequency of services.

For example, if complete daily flight schedules with passenger numbers for each flight are available, then passenger load can be calculated by averaging the number of passengers over all days and all flights between the two airports in the direction of the arc. If the number of passengers for a flight is not known, then this can be estimated using the capacity of the aircraft assigned to the flight along with an estimate of the average percentage of occupied seats, or *load factor*.

3.2. Calculating OD-pair demand data

Calculating OD-pair demand that is compatible with the observed passenger loads on each arc requires the solution of a type of inverse problem. This problem is modelled as a path-based multi-commodity flow problem and requires the identification of all possible paths passengers may take between each OD-pair. Four objectives are considered. The first is the level of asymmetry in the OD-pair passenger demand. The second and third objectives relate to the mean and standard deviation of the distribution of the expected number of single-leg passengers. The fourth objective is the deviation from the expected number of transiting passengers. As it is not clear what the trade-off is between these objectives, a multi-objective model is considered.

Characterising reasonable OD-paths. There may be many potential paths between an OD-pair $(o, d) \in \bar{K}$ in the flight network. However, not all of them will be considered *reasonable* with respect to the distance travelled from o to d , the time taken, the number of connections required, or the path's subpaths.

The distance travelled on a path between the OD-pair (o, d) is the sum of the great-circle distances of the arcs on the path. A path p is reasonable with respect to the distance travelled if

$$\sum_{a \in \mathcal{A}_p} \Delta_a \leq \gamma \Delta_{od}$$

where \mathcal{A}_p denotes the set of arcs on path p .

The time taken on a path is estimated to be the average waiting time plus the block time. The average waiting time w_{ij} for an arc $(i, j) \in A$ is estimated to be half the expected time between flights. If c_{\max} denotes the largest aircraft capacity, then a lower bound on the number of aircraft operating on the arc $(i, j) \in A$ is $N_{ij} = n_{ij}/c_{\max}$. Thus, the expected waiting time between flights is \mathcal{T}/N_{ij} , and $w_{ij} = 0.5\mathcal{T}/N_{ij}$. Let t_{ij} denote the block time.

The time taken between the OD-pair (o, d) is measured with respect to the time $w_{p_{\min}} + t_{p_{\min}}$ taken to travel the great-circle shortest-distance path p_{\min} . A path p is reasonable with respect to the time taken if

$$\sum_{(i,j) \in \mathcal{A}_p} (w_{ij} + t_{ij}) \leq w_{p_{\min}} + t_{p_{\min}}$$

For the short- and medium-haul networks considered in this paper it is assumed that the maximum number of connections

on a reasonable path is two. That is, the number of arcs on a reasonable path is at most three. Note that limiting the number of connections is a practical consideration. The methodology described in this paper works for an arbitrary number of connections on a path.

A path p is reasonable with respect to its subpaths if all subpaths are reasonable.

Example. Suppose that the arc (MEL, BNE) with 466 passengers per day also exists in the example network. Since (MEL, BNE) is an arc, it is the shortest-distance path for this OD-pair. Consider the path MEL – SYD – BNE and suppose that the biggest aircraft has a capacity for 160 passengers. The expected number of aircraft on each leg is $N_{\text{MEL,SYD}} = 776/160 = 4.85$, $N_{\text{SYD,BNE}} = 752/160 = 4.7$, $N_{\text{MEL,BNE}} = 466/160 = 2.9125$. The average waiting time in minutes for each arc is $w_{\text{MEL,SYD}} = 0.5 \times 1440/4.85 = 148$, $w_{\text{SYD,BNE}} = 0.5 \times 1440/4.7 = 153$, $w_{\text{MEL,BNE}} = 0.5 \times 1440/2.9125 = 247$. The block times in minutes for each arc is $t_{\text{MEL,SYD}} = 65$, $t_{\text{SYD,BNE}} = 68$, $t_{\text{MEL,BNE}} = 112$. Since $148 + 65 + 153 + 68 > 247 + 112$, MEL – SYD – BNE is not a reasonable path.

A path-based multi-objective multi-commodity flow model. Let $\mathcal{P}^k = \{p_1, p_2, \dots, p_{|\mathcal{P}^k|}\}$ denote all reasonable paths between each OD-pair $k \in \bar{K}$. Let the variable x_p denote the number of passengers on path $p \in \mathcal{P}^k$ between OD-pair $k \in \bar{K}$.

As previously mentioned four objective criteria are considered. The first is the level of asymmetry in OD-pair passenger demand. Let the variable ψ_{od} denote the absolute difference in demand between OD-pair $(o, d) \in \bar{K}$ and that of (d, o) .

$$\psi_{od} = \sum_{p \in \mathcal{P}^{(o,d)}} x_p - \sum_{p \in \mathcal{P}^{(d,o)}} x_p, \quad (o, d) \in \bar{K}: o < d \quad (2)$$

Let the parameter η_k denote a normalization factor for the demand for OD-pair $k \in \bar{K}$, so that OD-pair penalties in the optimisation model below are comparable across OD-pairs. The value of η_k is set so as to be a guaranteed “tight” upper bound on the OD-pair demand for k . It is “tight” in the sense that there exists a feasible solution achieving the bound. Let $\Psi_{od} = \max\{\eta_{od}, \eta_{do}\}$.

The second and third objectives relate to the average and standard deviation of the distribution of the single-leg passenger fractions. Specifically, the deviation from a target mean single-leg passenger fraction m_θ and the deviation from a target standard deviation s_θ . A linear L_1 -norm variation on the standard deviation, the average absolute difference of the single-leg passenger fraction from the target average fraction, is used to avoid nonlinear constraints in the model.

$$s_\theta = \frac{1}{|A|} \sum_{(i,j) \in A} |m_\theta - \theta_{ij}|$$

Let the variable ϕ_m denote the deviation of the mean single-leg passenger fraction from the target mean m_θ , and ϕ_s denote the deviation from the linearised standard deviation s_θ .

$$\phi_m = m_\theta - \frac{1}{|A|} \sum_{(i,j) \in A} \frac{1}{n_{ij}} \sum_{p \in \mathcal{P}^{(i,j)}: \mathcal{A}_p = \{(i,j)\}} x_p \quad (3)$$

$$\phi_s = s_\theta - \frac{1}{|A|} \sum_{(i,j) \in A} \bar{\theta}_{ij} \quad (4)$$

$$\bar{\theta}_{ij} \geq \left| m_\theta - \frac{1}{n_{ij}} \sum_{\substack{p \in \mathcal{P}^{(i,j)}: \\ \mathcal{A}_p = \{(i,j)\}}} x_p \right|, \quad (i,j) \in A \quad (5)$$

The variable $\bar{\theta}_{ij}$ bounds the absolute difference of the single-leg passenger fraction of arc $(i,j) \in A$ from the target average fraction.

Finally consider the deviation from the target transit fractions. Let the variable ξ_{ijk} denote the absolute difference between the number of passengers travelling on legs (i,j) and (j,k) , X_{ijk} , and the target $\alpha_{ijk}X_{ij}$, where X_{ij} denotes the number of passengers travelling on leg (i,j) that transit at j .

$$\xi_{ijk} = X_{ijk} - \alpha_{ijk}X_{ij}, \quad (i,j) \in A, (j,k) \in A_{ij}^{\text{out}}: |A_{ij}^{\text{out}}| > 1 \quad (6)$$

$$X_{ijk} = \sum_{\substack{(o,d) \in \bar{K}: \\ d \neq j}} \sum_{\substack{p \in \mathcal{P}^{(o,d)}: \\ \{(i,j),(j,k)\} \subseteq \mathcal{A}_p}} x_p, \quad (i,j) \in A, (j,k) \in A_{ij}^{\text{out}} \quad (7)$$

$$X_{ij} = \sum_{k: (j,k) \in A_{ij}^{\text{out}}} X_{ijk}, \quad (i,j) \in A \quad (8)$$

Let $\Xi_{ijk} = \max\{\alpha_{ijk}n_{ij}, (1 - \alpha_{ijk})n_{ij}\}$.

The path-based multi-objective multi-commodity flow model is formulated as follows.

$$\text{minimize} \begin{cases} \sum_{\substack{(o,d) \in \bar{K}: \\ o < d}} (\psi_{od}/\Psi_{od})^2 \\ \phi_m^2 \\ \phi_s^2 \\ \sum_{(i,j) \in A} \sum_{(j,k) \in A_{ij}^{\text{out}}} (\xi_{ijk}/\Xi_{ijk})^2 \end{cases} \quad (9)$$

$$\text{subject to} \sum_{k \in \bar{K}} \sum_{\substack{p \in \mathcal{P}^k: \\ a \in \mathcal{A}_p}} x_p = n_a, \quad a \in A \quad (10)$$

$$x_p \geq 0, \quad p \in \mathcal{P}^{(o,d)}, (o,d) \in \bar{K} \quad (11)$$

The four objectives (9), asymmetry in passenger flow (2), deviation from the target mean single-leg passenger fraction (3), deviation from the target linearised single-leg passenger fraction standard deviation (4), and deviation from the target transit fractions (6), are measured by minimising the sum of the squares of the individual terms. Quadratic, rather the linear, penalties are chosen in order to reduce the likelihood of outliers. Constraints (10) ensure that the passenger load on each arc is met exactly. Constraints (11) ensure nonnegativity of the x_p variables.

The feasible set of solutions to this problem is nonempty. For each feasible solution there is a corresponding set $K \subseteq \bar{K}$ of OD-pairs with nonzero passenger demand. There are infinitely many efficient solutions in decision space and, correspondingly, infinitely many nondominated points in objective

space. It is not clear what the trade-off is between the four objective functions. To keep the exposition of our methodology as simple as possible, the objectives were normalised, assigned equal weight, and the following single quadratic objective that is simply the sum of these terms was then minimized.

$$\begin{aligned} \text{minimize} \quad & \sum_{\substack{(o,d) \in \bar{K}: \\ o < d}} \frac{(\psi_{od}/\Psi_{od})^2}{|\{(o,d) \in \bar{K}: o < d\}|} \\ & + \phi_m^2 + \phi_s^2 \\ & + \sum_{(i,j) \in A} \sum_{(j,k) \in A_{ij}^{\text{out}}} \frac{(\xi_{ijk}/\Xi_{ijk})^2}{|\{(i,j,k): (i,j) \in A, (j,k) \in A_{ij}^{\text{out}}\}|} \end{aligned} \quad (12)$$

Each decision maker will place a different importance on each of the objectives and so by choosing different weights they can identify different efficient solutions as the best solution for their needs.

3.3. Validating the OD-pair demand model

To validate that the OD-pair demand model described above can infer realistic OD-pair demands, real network passenger data was needed. A “pure” short-haul hub-and-spoke instance was extracted from the US Department of Transportation Origin and Destination Survey DB1BCoupon data set for 2008 Q1 (US DOT, 2008) by including only those portions of passenger itineraries that included travel solely within the lower 48 states, and involved travel to or from the given hub, namely the airport ATL in Atlanta. All carriers and fare classes in the coupon data were included. An itinerary was considered to have a break at the end of a given leg if either a break was marked in the coupon data, or the following leg returned to the origin of the given leg, or the following leg did not travel to or from the hub, or the given leg was the last leg in the itinerary.

The data associated with this instance was analysed to identify suitable choices for the methodology’s controlling parameters. Figure 4 shows CDFs of arc passenger load and OD-pair demand as a fraction of the total number of passengers in the network, directness of passenger travel, relative asymmetry of OD-pair demand, single-leg passenger fractions, and transit passenger fractions.

The arc passenger load CDF shows the fraction of network arcs that have at most the given passenger load as a fraction of the total number of passengers in the network. Similarly, the OD-pair demand CDF shows the fraction of OD-pairs that have at most the given demand as a fraction of the total number of passengers in the network.

The directness of passenger travel CDF shows the fraction of OD-pairs that have at most the given measure of directness, as well as the fraction of combined OD-pair demand, summed over both directions, that experience at most this level of directness. The methodology requires a minimum measure of directness γ to be specified. This parameter is the minimum ratio of the direct distance between airports to the expected maximum distance a passenger would choose to travel between these airports and restricts the possible connections a multi-leg passenger can make. The parameter γ was set to be 0.5 which elim-

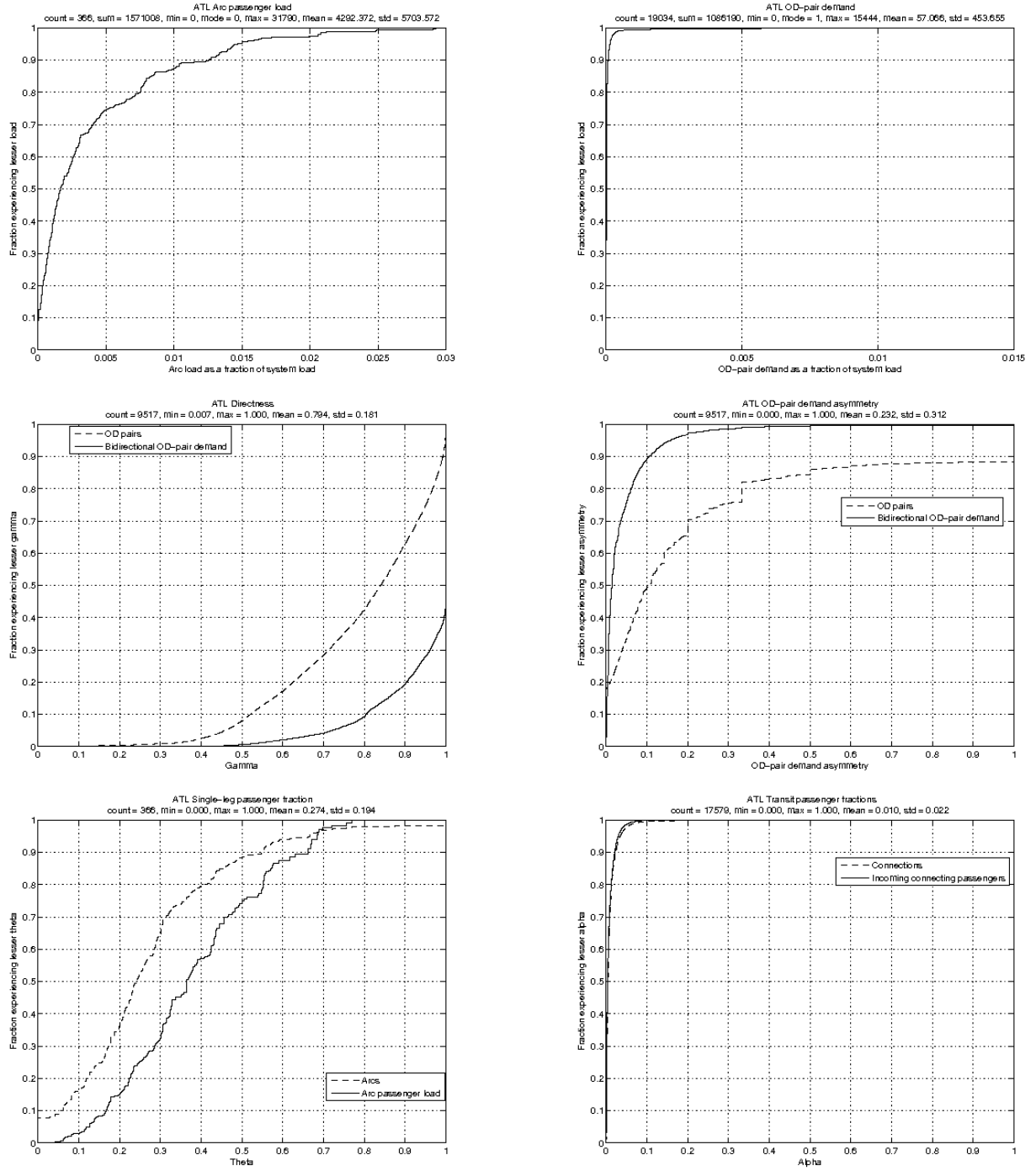


Figure 4: CDFs of arc passenger load, OD-pair demand, directness of passenger travel, relative asymmetry of OD-pair demand, single-leg passenger fractions, and transit passenger fractions, for the extracted hub-and-spoke ATL instance

inates approximately 8% of the possible OD-pair paths while only affecting approximately 1% of passengers.

The CDF of the relative asymmetry of OD-pair demands confirms that symmetric OD-pair passenger demand is a reasonable ideal. Approximately 75% of passengers travel on 35% of the OD-pair paths and experience a relative asymmetry of at most 5%, while approximately 90% of passengers travel on 50% of the OD-pair paths and experience a relative asymmetry of at most 10%.

The single-leg passenger fraction CDF shows the fraction of network arcs that have at most the given fraction of single-leg passengers, as well as the fraction of the passengers that experience at most this ratio of single-leg passengers to multi-leg passengers on an arc. The mean and linearised standard deviation of the distribution of single-leg passenger fractions was found to be $(m_\theta, s_\theta) = (0.274, 0.143)$.

Finally, the CDF of the transit passenger fractions shows the fraction of multi-leg connections that have at most the given transit fraction, as well as the fraction of incoming connecting passengers that experience at most this transit fraction.

The distributions of OD-pair demands, single-leg passenger loads, and transiting passenger loads inferred by the OD-pair demand model were found to have very similar characteristics to those of the real data. The averages of the distributions for the real and model inferred data for the three characteristics were found to coincide, while the relative error between the standard deviations of the distributions of the real and inferred data were 10.4%, 7.6%, and 6.7%, respectively. This level of variation is acceptable given that similar levels of variation are likely to be seen between the distributions of real network data extracted from the US Department of Transportation Origin and Destination Survey DB1BCoupon data sets for different quarters.

4. Benchmark instances

The benchmark instances consist of thirty single-hub and three two-hub networks. Tables 5 and 6 provide the parameters used to generate the networks. In these tables the hub name is preceded by either an “s”, “m”, or “l”, indicating whether the instance is considered to be a short-, medium-, or long-haul network, respectively.

To generate a single-hub network having a given number of spokes, the network characteristics identified in Section 2.1 are used. Recall that five possible arc distance CDFs and five possible arc capacity CDFs were identified. Which of these is used for an instance is indicated in Table 5, by index. Details of the CDF corresponding to each index can be found in Evans et al. (2010). For each spoke, first the length of the spoke is sampled from the given arc distance CDF, and then its capacity is sampled from the given capacity CDF. Then the directions of all spokes are determined, so as to match the lobe characteristics described in Section 2.1, and quantified by the $R_{minor-major}$ and $R_{lesser-greater}$ values given in Table 5. Target and actual values are given for these directional capacity parameters. For each combination of arc distance CDF, arc capacity CDF, and target directional capacity parameters, three networks of different

Hub name	Distance CDF	Capacity CDF	$R_{minor-major}$		$R_{lesser-greater}$		Spoke ports
			Target	Actual	Target	Actual	
sHAA	1	4	0.4	0.43	0.4	0.43	24
sHBA			0.4	0.44	0.4	0.43	72
sHCA			0.4	0.43	0.4	0.42	120
sHDA	1	4	0.8	0.77	0.4	0.41	24
sHEA			0.8	0.79	0.4	0.43	72
sHFA			0.8	0.81	0.4	0.44	120
sHGA	2	4	0.2	0.23	0.4	0.43	24
sHHA			0.2	0.22	0.4	0.42	72
sHIA			0.2	0.23	0.4	0.42	120
sHJA	2	4	0.4	0.43	0.8	0.81	24
sHKA			0.4	0.42	0.8	0.8	72
sHLA			0.4	0.42	0.8	0.8	120
mHMA	3	4	0.2	0.23	0.2	0.23	12
mHNA			0.2	0.23	0.2	0.23	24
mHOA			0.2	0.22	0.2	0.22	60
mHPA	3	4	0.2	0.23	0.8	0.81	12
mHQA			0.2	0.23	0.8	0.81	24
mHRA			0.2	0.21	0.8	0.79	60
lHSA	4	5	0.1	0.12	0.5	0.55	12
lHTA			0.1	0.13	0.5	0.53	24
lHUA			0.1	0.12	0.5	0.51	60
lHVA	4	5	0.5	0.53	0.8	0.78	12
lHWA			0.5	0.52	0.8	0.8	24
lHXA			0.5	0.5	0.8	0.78	60
lHYA	5	5	0.2	0.22	0.1	0.14	12
lHZA			0.2	0.22	0.1	0.12	24
lH1A			0.2	0.23	0.1	0.13	72
lH2A	5	5	0.1	0.11	0.7	0.67	12
lH3A			0.1	0.14	0.7	0.72	24
lH4A			0.1	0.13	0.7	0.72	72

Table 5: Parameters for generating single-hub benchmark instances

Hub name	Distance CDF	Capacity CDF	$R_{minor-major}$		$R_{lesser-greater}$		Spoke ports	Spoke capacity		Inter-hub capacity		Multi-hub spoke ports	Inter-hub distance
			Target	Actual	Target	Actual		Target	Actual	Target	Actual		
sHAB	2	4	0.2	0.19	0.75	0.75	18	0.4	0.39	0.013	0.013	14	1400
sHBB			0.1	0.07	0.55	0.65	20	0.59	0.60				
sHCB	2	4	0.2	0.25	0.75	0.73	54	0.40	0.42	0.013	0.012	42	1400
sHDB			0.1	0.15	0.55	0.46	60	0.59	0.57				
sHEB	2	4	0.2	0.15	0.75	0.76	90	0.40	0.38	0.013	0.012	70	1400
sHFB			0.1	0.18	0.55	0.5	100	0.59	0.61				

Table 6: Parameters for generating two-hub benchmark instances

sizes were generated in order to allow for the investigation of scaling effects.

The two-hub instances were generated by “gluing” two single-hub instances together in an acceptable way. Table 6 contains the additional parameters needed in order to generate the two-hub networks. The spoke capacity and inter-hub capacity columns provide the fraction of the total number of passengers in the network that are observed on each hub’s spoke arcs and the inter-hub arc. The multi-hub spoke airports column provides the number of spoke airports that the two hubs have in common. The inter-hub distance is the number of kilometres that the second hub lies to east of the first.

In generating OD-pair demands, the nominal value used as the target mean and linearised standard deviation of the distribution of single-leg passenger fractions was $(m_\theta, s_\theta) = (0.4, 0.2)$ for short-haul networks and $(m_\theta, s_\theta) = (0.6, 0.2)$ for both medium- and long-haul. The minimum ratio of the direct distance between airports to the expected maximum distance a passenger would choose to travel between these airports used for all instances was $\gamma = 0.5$.

The complete set of instances and supporting material is available at the URL http://www.infotech.monash.edu.au/~wallace/airline_benchmarks/ along with the references Evans et al. (2010) and Evans and Waterer (2011).

5. Analysis of instances

Summary statistics are presented for all of the benchmark instances. More detailed results are presented for three selected single-hub instances, one two-hub instance, and a fictitious Australian carrier called Emu Airlines that operates a point-to-point network. For these instances, network and directional capacity diagrams, and CDFs of various network and passenger characteristics.

The arc passenger load CDFs show the fraction of network arcs that have at most the given passenger load as a fraction of the total number of passengers in the network. Similarly, OD-pair demand CDFs show the fraction of OD-pairs that have at most the given demand as a fraction of the total number of passengers in the network.

The directness of passenger travel CDFs show the fraction of OD-pairs that have at most the given measure of directness, as well as the fraction of combined OD-pair demand, summed over both directions, that experience at most this level of directness. Similarly, the relative asymmetry of OD-pair demand

Ports	ADL	BNE	CBR	CNS	MEL	OOL	SYD
ADL	-	158	51	36	0	64	666
BNE	197	-	185	183	237	0	529
CBR	52	173	-	33	0	59	170
CNS	50	189	50	-	46	0	185
MEL	0	238	0	40	-	77	421
OOL	68	0	62	0	77	-	15
SYD	753	600	191	188	438	15	-

Table 7: Passenger demand between OD-pairs in the example network

CDFs show the fraction of OD-pairs that have at most the given level of asymmetry, as well as the fraction of combined OD-pair demand that experience at most this level of asymmetry.

The single-leg passenger fractions CDFs show the fraction of network arcs that have at most the given fraction of single-leg passengers, as well as the fraction of the passengers that experience at most this ratio of single-leg passengers to multi-leg passengers on an arc. The transit passenger fraction CDFs show the fraction of multi-leg connections that have at most the given transit fraction, as well as the fraction of incoming connecting passengers that experience at most this transit fraction.

5.1. Example instance

The nominal value used as the target mean and linearised standard deviation of the distribution of single-leg passenger fractions was $(m_\theta, s_\theta) = (0.4, 0.2)$. The minimum ratio of the direct distance between airports to the expected maximum distance a passenger would choose to travel between these airports used was $\gamma = 0.5$. Table 7 provides the passenger demand between OD-pairs in the example network. The small amount of asymmetry seen in the OD-pair demands is not unrealistic. The deviations from the target mean and linearised standard deviation of the distribution of single-leg passenger fractions, as well as the target transit passenger fractions, were negligible.

5.2. Benchmark instances

Summary statistics for the thirty single-hub and three two-hub networks are presented in Table 8. These tables summarise the number of spoke airports, network arcs, OD-pairs with nonzero demand, and passengers in the network. Statistics on the distribution of the observed number of passengers on an arc, the duration of a leg on an arc, the great-circle distance of an arc, the

Hub name	Spoke ports	Arcs	OD pairs	Pax count	avg	Arc pax load			Block time			Arc distance			Origin degree			OD-pair demand			Transit pax %							
						sidev	min	max	avg	sidev	min	max	avg	sidev	min	max	avg	sidev	min	max	avg	sidev	min	max				
SHAA	24	48	474	71410	1487.71	1340.70	124	4860	137.81	46.49	60	270	1338.83	454.88	249	3079	18.96	2.16	14	23	116.47	398.25	1	3524	0.60	0.21	0.27	1.00
SHBA	72	144	3823	106474	739.40	821.65	28	4860	135.35	54.42	55	300	1304.17	531.70	178	3444	52.41	9.37	13	72	20.35	108.95	0	2620	0.60	0.10	0.36	0.86
SHCA	120	240	10624	224898	937.08	911.61	68	4860	132.71	53.11	55	330	1268.03	520.55	178	3846	87.97	12.52	34	120	16.00	122.08	0	3779	0.60	0.13	0.21	0.93
SHDA	24	48	446	71410	1487.71	1340.70	124	4860	137.81	46.49	60	270	1338.83	454.88	249	3079	17.84	2.88	13	24	119.22	372.80	1	3509	0.60	0.18	0.28	1.00
SHEA	72	144	3840	106474	739.40	821.65	28	4860	135.35	54.42	55	300	1304.17	531.70	178	3444	52.66	8.94	17	72	20.33	111.01	0	2769	0.60	0.11	0.36	0.85
SHFA	120	240	10877	224898	937.08	911.61	68	4860	132.71	53.11	55	330	1268.03	520.55	178	3846	90.03	11.73	31	120	15.59	118.56	0	3668	0.60	0.13	0.23	0.92
SHGA	24	48	480	53230	1108.96	1189.57	83	4860	161.04	73.21	70	325	1663.79	717.68	423	3782	19.20	1.94	16	23	81.28	253.43	1	2111	0.60	0.15	0.29	1.00
SHHA	72	144	3795	104278	724.15	808.63	28	4860	163.23	70.57	70	325	1693.36	691.72	409	3782	52.00	8.73	19	72	20.40	122.22	0	3478	0.60	0.11	0.28	0.84
SHIA	120	240	9766	173132	721.38	789.65	36	4860	161.56	72.02	60	330	1670.18	706.43	244	3846	80.86	16.21	14	120	13.13	95.79	0	3233	0.60	0.10	0.31	0.82
SHJA	24	48	492	53230	1108.96	1189.57	83	4860	161.04	73.21	70	325	1663.79	717.68	423	3782	19.68	2.24	15	23	79.51	242.49	1	2057	0.60	0.17	0.34	1.00
SHKA	72	144	3833	104278	724.15	808.63	28	4860	163.23	70.57	70	325	1693.36	691.72	409	3782	52.60	9.40	17	72	20.04	113.64	0	3151	0.60	0.11	0.35	0.85
SHLA	120	240	10175	173132	721.38	789.65	36	4860	161.56	72.02	60	330	1670.18	706.43	244	3846	84.53	16.93	9	120	12.67	97.73	0	3514	0.60	0.09	0.28	0.72
mHMA	12	24	116	17564	731.83	810.76	50	2640	278.33	140.98	70	555	3295.33	1378.39	401	6857	8.92	2.09	6	12	132.10	398.48	1	2338	0.40	0.25	0.11	0.90
mHNA	36	72	1051	56430	783.75	706.39	55	2640	257.85	154.65	65	560	3009.56	1513.15	300	6927	28.49	3.47	19	36	44.11	178.43	0	1685	0.40	0.12	0.17	0.65
mHOA	60	120	2352	51246	427.05	479.20	15	2640	252.96	146.21	65	620	2941.65	1432.26	300	7748	38.59	10.47	1	60	18.32	103.59	0	1807	0.38	0.16	0.00	0.74
mHPA	12	24	128	17564	731.83	810.76	50	2640	278.33	140.98	70	555	3295.00	1377.96	401	6853	9.85	1.51	8	12	117.73	343.90	1	2048	0.40	0.14	0.22	0.64
mHOA	36	72	1092	56430	783.75	706.39	55	2640	257.85	154.65	65	560	3009.97	1513.66	300	6927	29.51	3.04	21	36	43.00	180.77	1	1835	0.40	0.11	0.28	0.74
mHRA	60	120	2372	51246	427.05	479.20	15	2640	252.96	146.21	65	620	2941.67	1432.26	300	7747	38.92	10.86	1	60	18.14	104.34	0	1978	0.38	0.11	0.00	0.64
IHSA	12	24	128	20062	835.92	403.54	253	1472	460.42	231.24	85	740	5831.17	2263.04	619	9295	9.85	1.46	8	12	127.50	230.35	2	1066	0.40	0.09	0.28	0.62
IHTA	36	72	976	48332	671.28	332.73	106	1472	473.40	229.20	90	840	6008.22	2246.04	664	10635	26.38	4.50	21	36	40.85	135.61	1	1090	0.40	0.16	0.21	0.99
IHTA	60	120	2610	56192	468.27	256.74	155	1472	479.75	217.37	90	865	6090.78	2127.48	731	10990	42.82	5.87	33	60	17.56	74.54	0	1155	0.40	0.08	0.22	0.56
IHVA	12	24	138	20062	835.92	403.54	253	1472	460.62	231.32	85	740	5831.75	2263.32	619	9295	10.62	1.15	9	12	117.90	221.18	2	1083	0.40	0.08	0.27	0.54
IHTA	36	72	1047	48332	671.28	332.73	106	1472	471.11	226.16	90	815	5974.58	2215.32	664	10336	28.52	3.43	23	36	37.61	123.31	0	995	0.40	0.09	0.29	0.69
IHXA	60	120	2773	56192	468.27	256.74	155	1472	478.50	215.94	90	885	6074.42	2113.71	731	11226	45.51	5.35	37	60	16.45	70.21	0	1110	0.40	0.06	0.25	0.51
IHYA	12	24	100	14160	590.00	327.26	212	1472	584.79	144.67	360	825	7555.58	1406.59	4597	10452	7.69	2.61	5	12	119.29	230.40	1	1286	0.40	0.24	0.13	0.88
IHTA	36	72	715	19526	271.19	243.00	46	1472	582.22	124.35	235	840	7524.83	1199.59	2819	10687	19.35	6.01	13	35	23.24	91.65	0	1292	0.40	0.26	0.12	1.00
IHLA	60	120	1955	55908	465.90	258.45	155	1472	587.67	139.90	240	905	7593.20	1354.79	2851	11538	32.13	10.04	19	60	24.01	99.70	0	1280	0.40	0.26	0.13	0.99
IHTA	12	24	106	14160	590.00	327.26	212	1472	586.88	147.54	360	825	7584.50	1433.49	4597	10437	8.15	1.46	6	12	109.43	198.33	4	1199	0.40	0.11	0.18	0.54
IHTA	36	72	744	19526	271.19	243.00	46	1472	582.71	124.81	235	845	7529.25	1202.46	2819	10701	20.11	5.55	13	36	21.64	79.19	1	1193	0.40	0.10	0.19	0.67
IHTA	60	120	2131	55908	465.90	258.45	155	1472	586.04	137.30	240	905	7569.65	1328.56	2851	11532	34.95	6.32	27	60	21.35	80.79	0	1073	0.40	0.08	0.24	0.54
Hub name	Spoke ports	Arcs	OD pairs	Pax count	avg	Arc pax load			Block time			Arc distance			Origin degree			OD-pair demand			Transit pax %							
SHAB	22	74	486	94290	1274.19	1030.61	167	4860	175.00	75.59	65	325	1854.41	742.25	300	3795	20.25	2.03	17	23	133.13	283.11	1	3463	0.60	0.23	0.24	1.00
SHCB	70	226	3849	184406	815.96	888.14	25	4860	168.65	80.25	60	375	1767.53	786.96	290	4459	53.67	9.79	17	71	30.73	95.47	0	1949	0.59	0.24	0.06	0.97
SHCB	118	378	10159	287234	759.88	825.92	35	4860	169.83	78.24	55	405	1783.70	767.31	216	4835	85.17	17.83	52	114	19.00	84.02	0	2157	0.59	0.22	0.00	0.98

Table 8: Summary statistics for single-hub (top) and two-hub (bottom) benchmark instances

number of unique passenger itineraries (paths) for an OD-pair, the demand for an OD-pair, and the percentage of passengers transiting at a airport, are also given. The instances are grouped into threes. Each group of three networks were generated using the same parameters except the number of spokes which was varied to provide networks of different sizes.

Single-hub benchmark instances. Three single-hub instances are presented. The instance HBA is a short-haul network with 72 spokes. Figure 5 show the network and directional capacity diagrams for this instance. The hub is quite asymmetrical. Seventy percent of the arc capacity is concentrated on the major axis which has an east-west orientation. Although the lobes of the major axis are diametrically opposed the east-orientated greater lobe has more than twice the capacity of the lesser lobe.

The instance HRA is a medium-haul network with 60 spokes. Figure 6 show the network and directional capacity diagrams for this instance. The hub is relatively symmetric. More than 80% of the arc capacity is concentrated on the major axis which has an east-west orientation. The lobes of the major axis are diametrically opposed with the lesser lobe having nearly 80% of the capacity of the greater lobe. The greater lobe is orientated east.

The instance HXA is a long-haul network with 60 spokes. Figure 7 show the network and directional capacity diagrams for this instance. The hub is relatively symmetric. Two-thirds of the arc capacity is concentrated on the major axis which has an east-west orientation. The lobes of the major axis are diametrically opposed with the lesser lobe having nearly 80% of the capacity of the greater lobe. The greater lobe is orientated east.

Figures 8–10 show the CDFs of the various network and passenger characteristics. There is nothing evident in these plots to suggest that the methodology is lacking. The level of asymmetry in OD-pair demands, as well as the deviations from the target mean and linearised standard deviation of the distribution of single-leg passenger fractions, and the target transit passenger fractions, were negligible in every instance.

Two-hub benchmark instances. The instance HCB-HDB is a network with two hubs. Figure 11 shows the network and diagram for this instance. The network has significant asymmetry. The hub HDB is located 1400km east of the hub HCB and just over 1% of the network's passengers are observed using this arc. Forty-two of the 70 spokes are shared by the two hubs. Forty-two percent of the network's passengers are observed travelling to spokes from HCB, while 57% are observed travelling from HDB. Eighty percent of the arc capacity of HCB is concentrated on its major axis with the lesser lobe having almost three quarters of the capacity of the greater lobe. More than 80% of the arc capacity of HDB is concentrated on its major axis with the greater lobe having slightly more than twice the capacity of the lesser lobe.

Figure 12 shows the CDFs of various network and passenger characteristics. There is nothing evident in these plots to suggest that the methodology is lacking. The deviations from

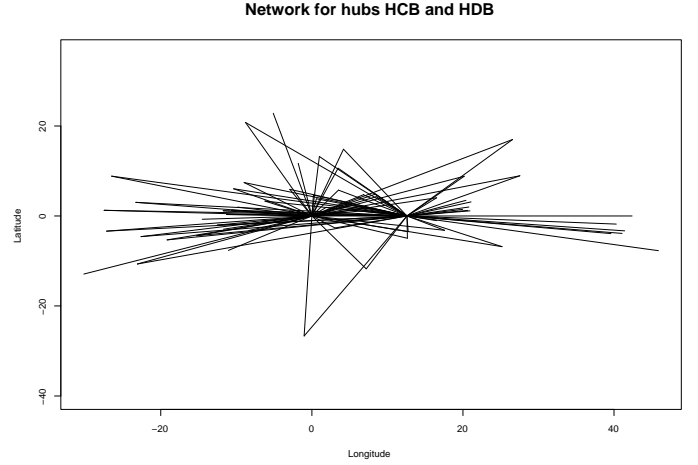


Figure 11: Network diagram for the HCB-HDB instance

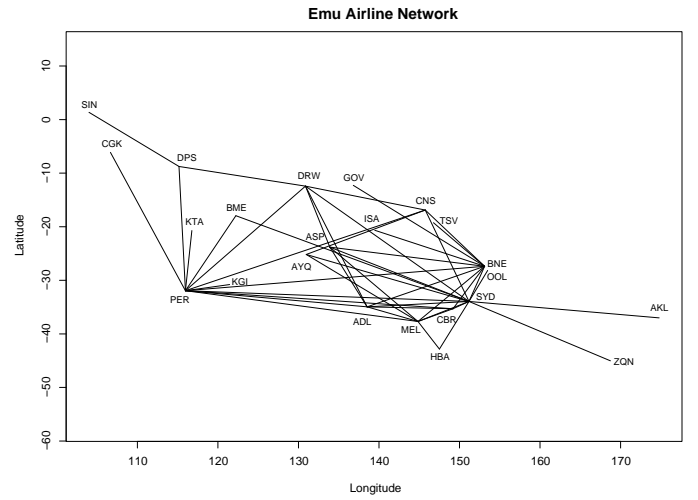


Figure 13: Point-to-point network for the Emu Airlines instance

the target mean and linearised standard deviation of the distribution of single-leg passenger fractions were negligible. The level of asymmetry in the OD-pair demands and the deviations from the target transit passenger fractions are small.

5.3. Point-to-point Emu Airlines instance

Emu Airlines is a fictitious Australian carrier that operates the point-to-point network shown in Figure 13. The observed passenger numbers on each arc used in this instance are based on proprietary data made available by an industrial partner.

The nominal value used as the target mean and linearised standard deviation of the distribution of single-leg passenger fractions was $(m_\theta, s_\theta) = (0.75, 0.2)$. The minimum ratio of the direct distance between airports to the expected maximum distance a passenger would choose to travel between these airports was $\gamma = 0.5$.

Figure 14 shows the CDFs of various network and passenger characteristics. There is nothing evident in these plots to suggest that the methodology is lacking. The level of asymmetry in the OD-pair demands is small. The deviations from

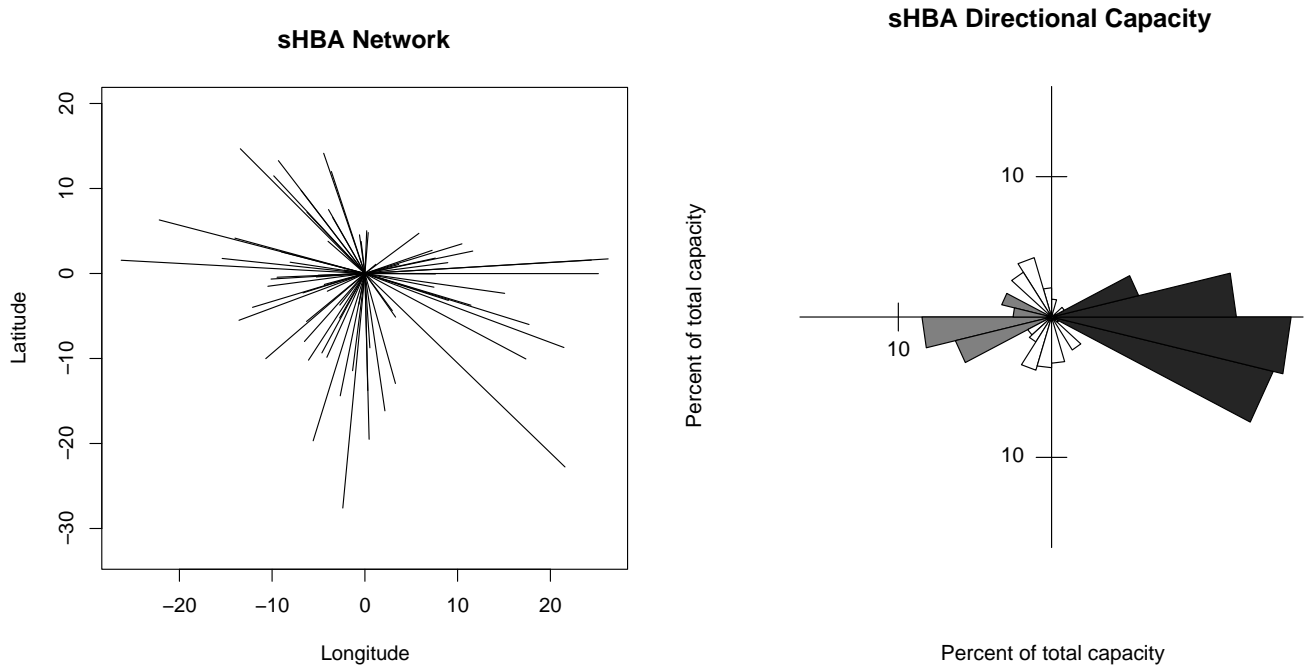


Figure 5: Network and directional capacity diagrams for the short-haul HBA instance

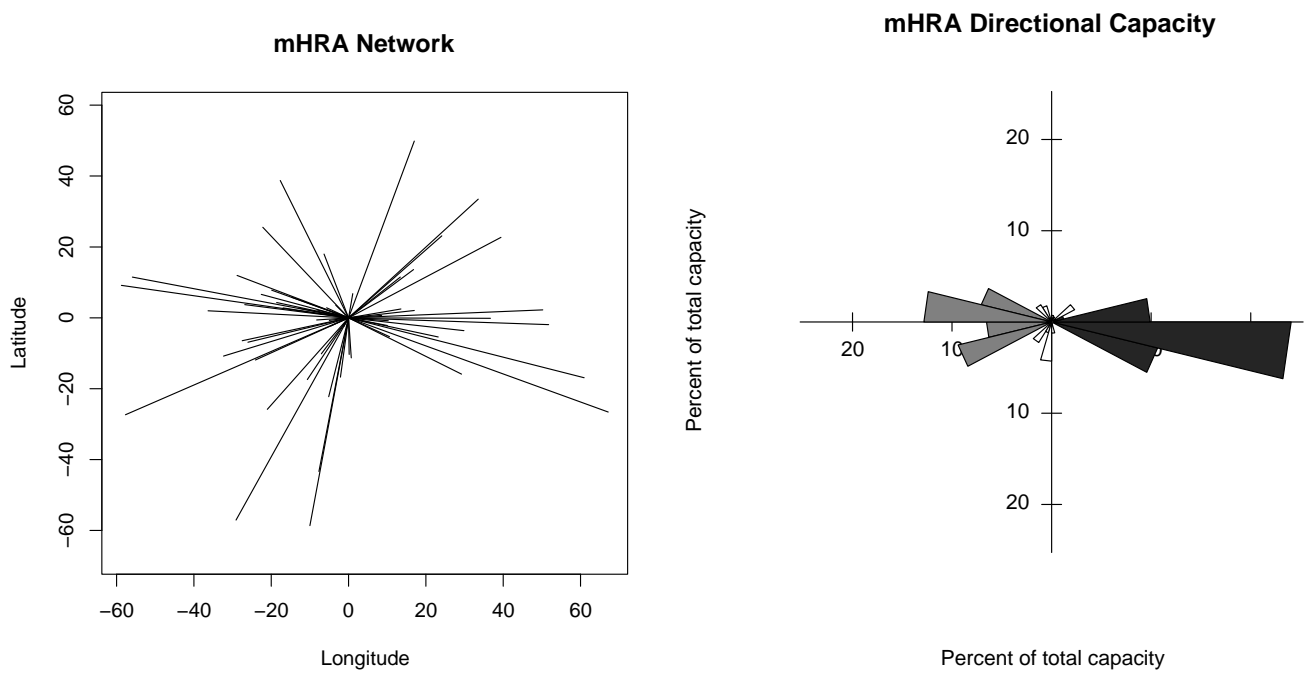


Figure 6: Network and directional capacity diagrams for the medium-haul HRA instance

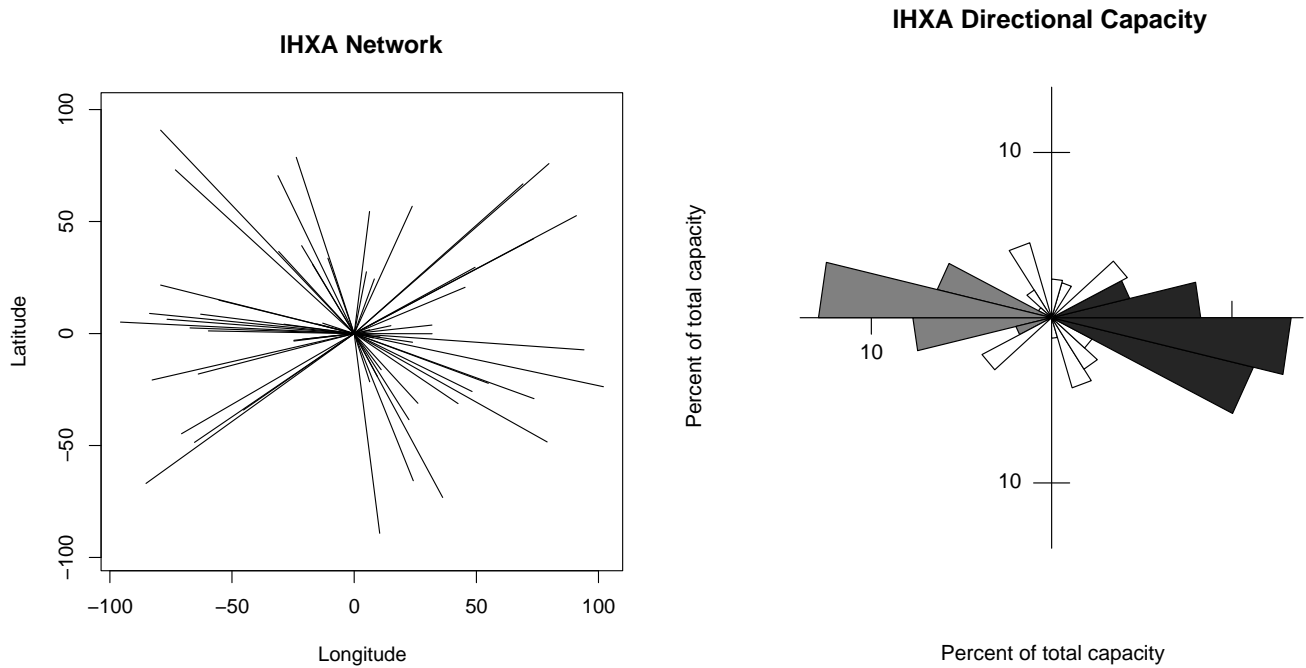


Figure 7: Network and directional capacity diagrams for the long-haul HXA instance

the target mean and linearised standard deviation of the distribution of single-leg passenger fractions, as well as the target transit passenger fractions, were negligible.

6. Conclusions and future work

This paper is the first of two papers entitled “Airline Planning Benchmark Problems” that present a four step framework for generating realistic airline planning benchmark problem instances. These instances are a result of analysing rich data sets from a wide range of airlines worldwide, including all airlines in the Star and oneworld alliances. The methodology behind the first two steps in the framework, namely, characterising airline networks and OD-pair demand using limited data, were presented in this paper. The methodology of the second two steps, namely, characterising passenger groups and the allocating the OD-pair demand, is presented in the second paper (Akartunali et al., 2011).

The thirty single-hub and three two-hub instances provide standardised data that includes OD-pair passenger demand data which is critical for the first step in the airline planning process, namely, flight schedule design. It is hoped that the availability of these instances, and a description of the methodology used to generate them, will not only make research in airline planning accessible to researchers from outside this area, but will also stimulate existing research by providing data that facilitates the accurate and repeatable comparison of the many different algorithms and techniques for the airline planning process reported in the literature.

Future work includes extending the characterisation of airline networks to include other topologies, such as linked hubs and point-to-point networks, and to generate sets of benchmarks

for such networks, as well as incorporating airline resources, such as aircraft and crew, into the benchmarks.

Acknowledgements

The authors are very grateful to Ian Evans and Alan Dormer (CTI Pty Ltd) for their ongoing support and guidance on a variety of practical airline-related issues and for numerous technical suggestions and insightful feedback that improved the content and exposition of this work. This research is supported by the Australian Research Council, under Linkage Projects LP00668076 and LP0883855, and by CTI Pty Ltd.

References

- Ahuja, R., Goodstein, J., Mukherjee, A., Orlin, J., Sharma, D., 2001. A very large-scale neighborhood search algorithm for the combined through and fleet assignment model. Working paper 4388-01, MIT Sloan School of Management.
- Akartunali, K., Boland, N., Evans, I., Wallace, M., Waterer, H., 2011. Airline planning benchmark problems. Part II: Passenger groups, utility and demand allocation. Tech. rep., University of Newcastle.
- Akartunali, K., Boland, N., Evans, I., Wallace, M., Waterer, H., Smith, O., 2010. Airline schedule design: Network design optimization and heuristics. Working paper, The University of Melbourne.
- Barnhart, C., Boland, N., Clarke, L., Johnson, E., Nemhauser, G., Shenoi, R., 1998. Flight string models for aircraft fleet and routing. *Transportation Science* 32, 208–220.
- Barnhart, C., Cohn, A., 2004. Airline schedule planning: accomplishments and opportunities. *Manufacturing and Service Operations Management* 6 (1), 3–22.
- Barnhart, C., Farahat, A., Lohatepanont, M., 2009. Airline fleet assignment with enhanced revenue modeling. *Operations Research* 57 (1), 231–244.
- Bazargan, M., 2004. *Airline Operations and Scheduling*. Ashgate Publishing Ltd.

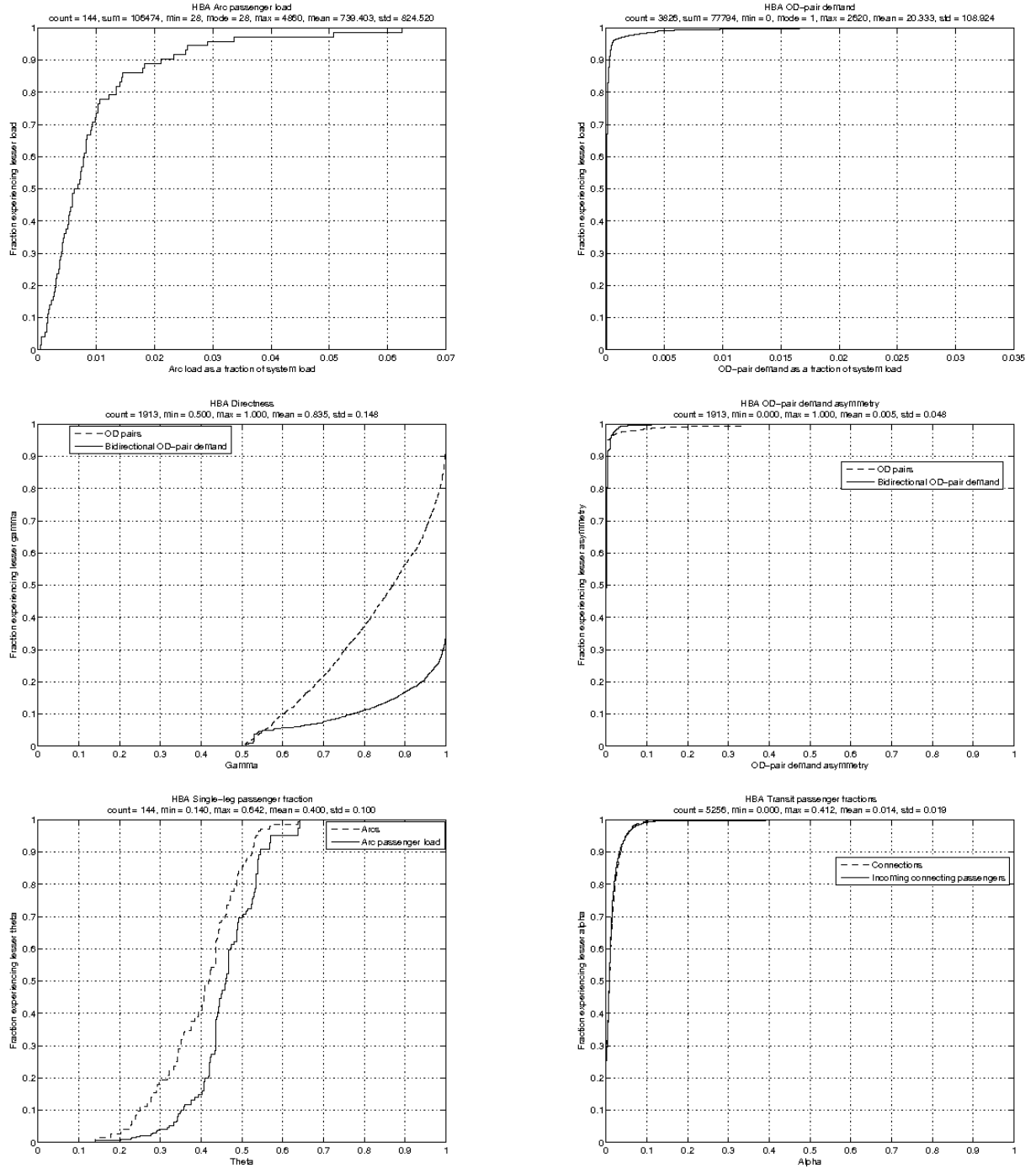


Figure 8: CDFs of arc passenger load, OD-pair demand, directness of passenger travel, relative asymmetry of OD-pair demand, single-leg passenger fractions, and transit passenger fractions, for the short-haul HBA instance

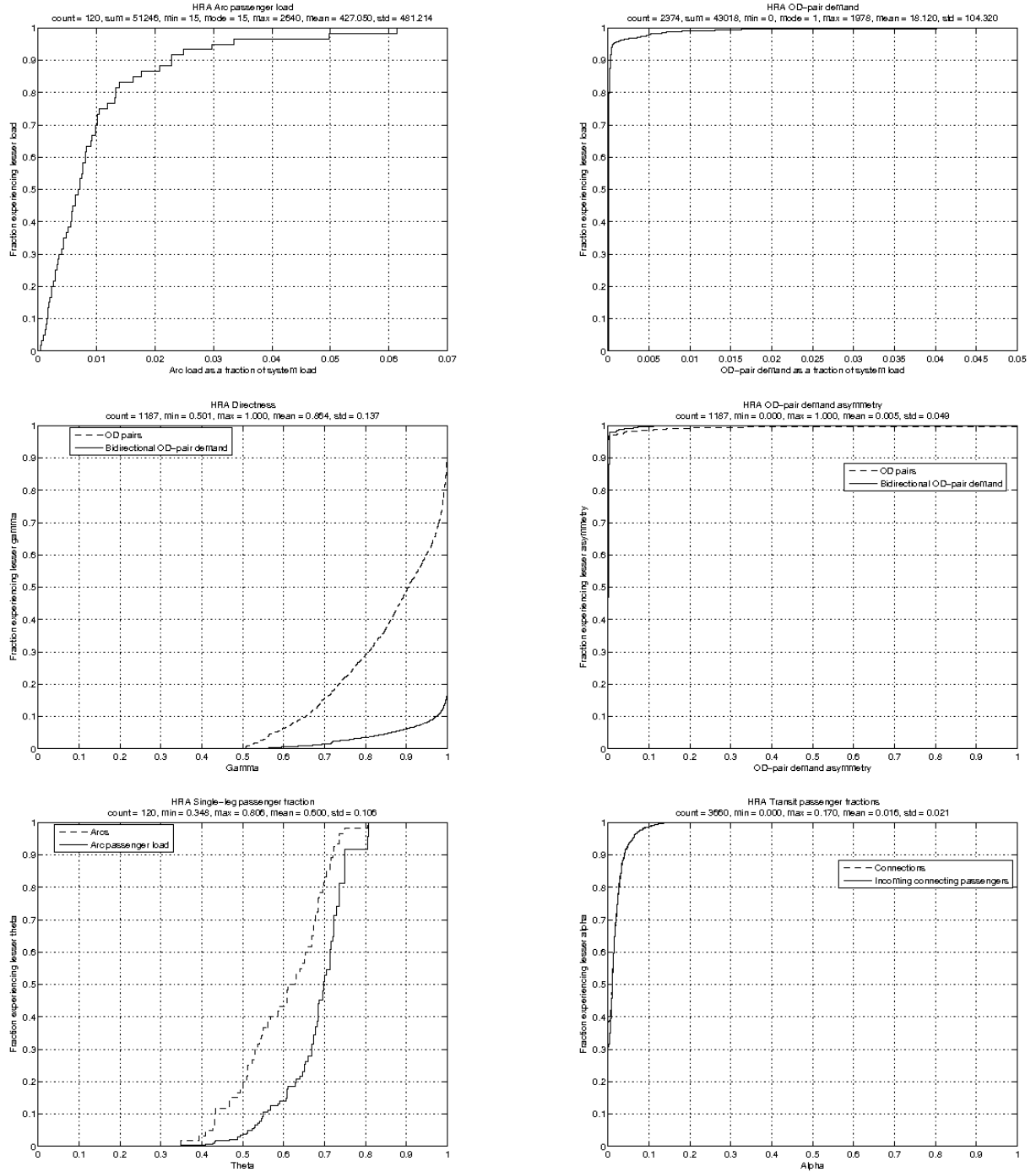


Figure 9: CDFs of arc passenger load, OD-pair demand, directness of passenger travel, relative asymmetry of OD-pair demand, single-leg passenger fractions, and transit passenger fractions, for the medium-haul HRA instance

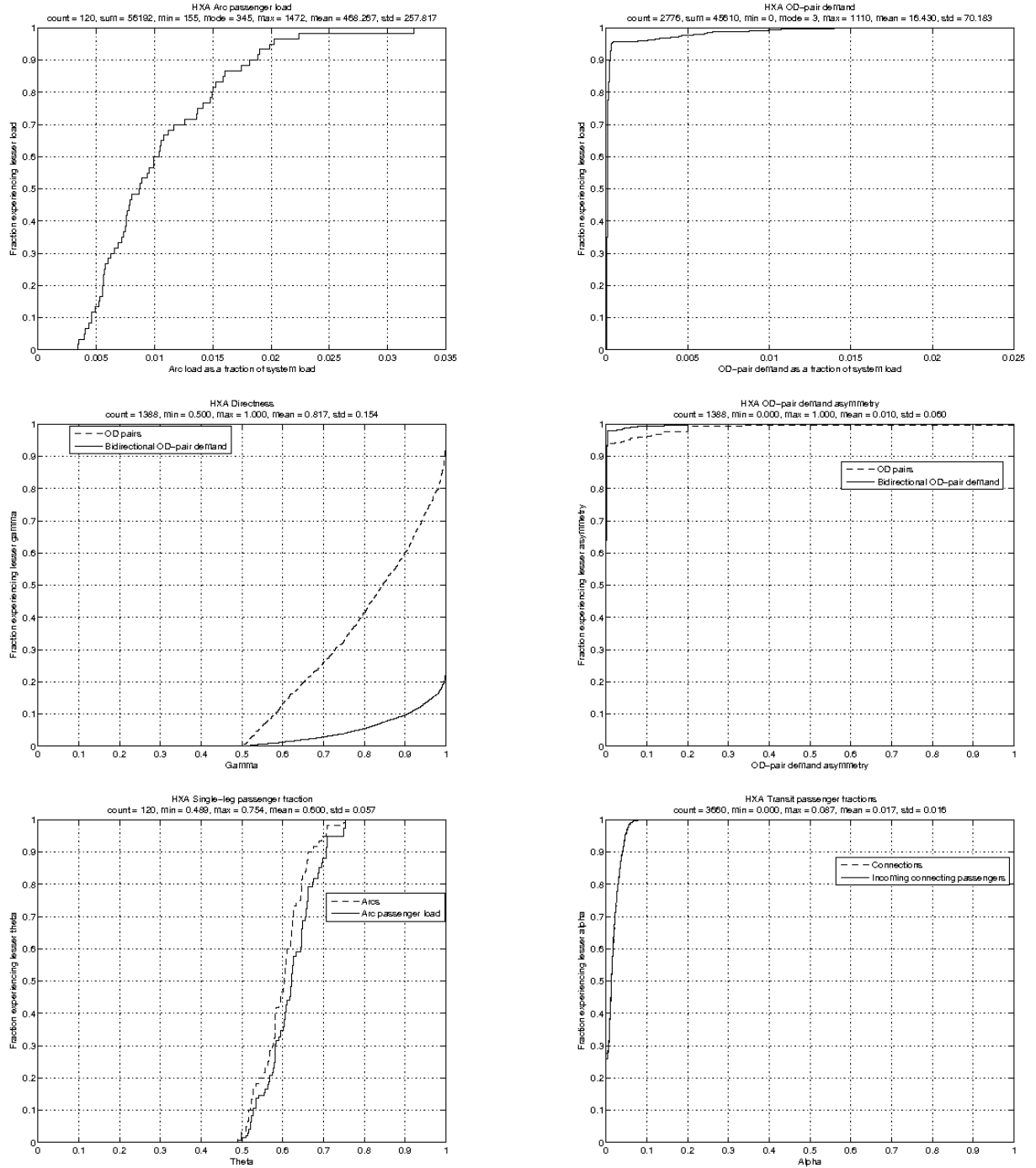


Figure 10: CDFs of arc passenger load, OD-pair demand, directness of passenger travel, relative asymmetry of OD-pair demand, single-leg passenger fractions, and transit passenger fractions, for the long-haul HXA instance

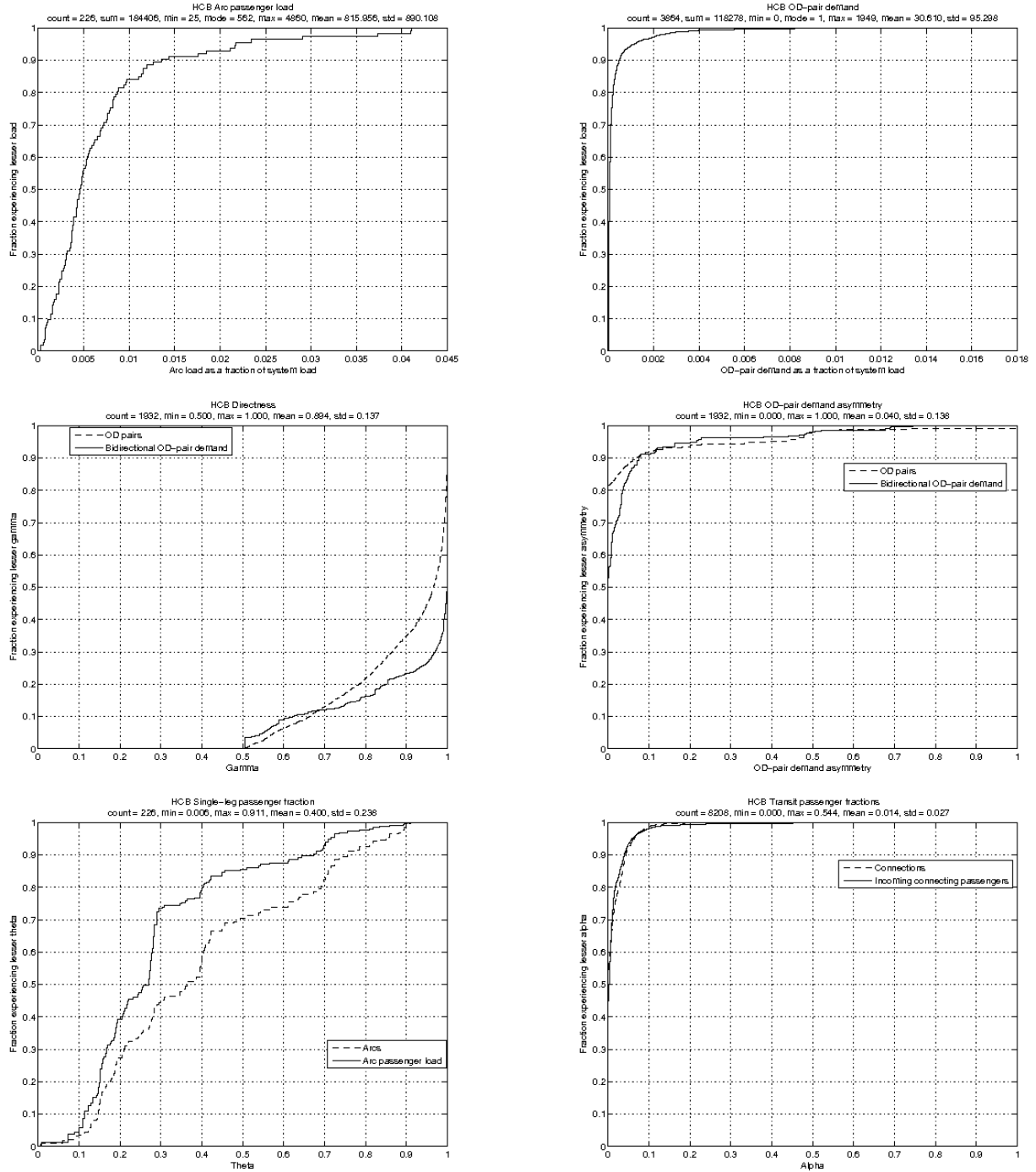


Figure 12: CDFs of arc passenger load, OD-pair demand, directness of passenger travel, relative asymmetry of OD-pair demand, single-leg passenger fractions, and transit passenger fractions, for the HCB-HDB instance

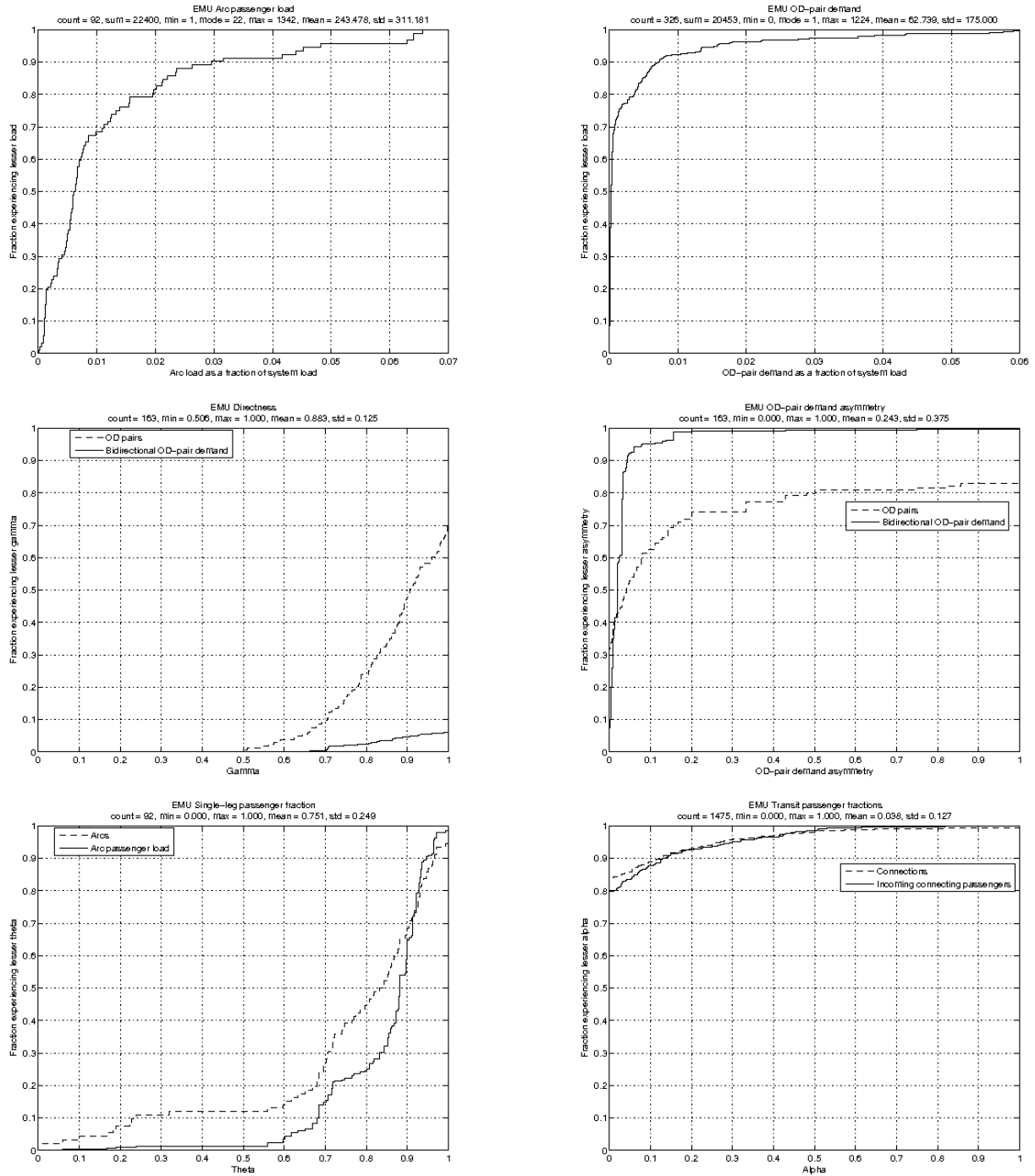


Figure 14: CDFs of arc passenger load, OD-pair demand, directness of passenger travel, relative asymmetry of OD-pair demand, single-leg passenger fractions, and transit passenger fractions, for the Emu Airlines instance

- Bront, J. J. M., Mendez-Diaz, I., Vulcano, G., 2009. A column generation algorithm for choice-based network revenue management. *Operations Research* 57 (3), 769–784.
- Clarke, L., Johnson, E., Nemhauser, G., Zhu, Z., 1997. The aircraft rotation problem. *Annals of Operations Research* 69, 33–46.
- Coldren, G., Koppelman, F., Kasturirangan, K., Mukherjee, A., 2003. Modeling aggregate air-travel itinerary shares: logit model development at a major US airline. *Journal of Air Transport Management* 9, 361–369.
- Dresner, M., Windle, R., 1995. Are US air carriers to be feared? Implication of hubbing to North Atlantic competition. *Transport Policy* 2 (3), 195–202.
- Dumas, J., Soumis, F., 2008. Passenger flow model for airline networks. *Transportation Science* 42 (2), 197–207.
- Evans, I., Wallace, M., Waterer, H., 2010. Characteristics of airline hubs. Tech. rep., Constraint Technologies International, available from http://www.infotech.monash.edu.au/~wallace/airline_benchmarks/.
- Evans, I., Waterer, H., 2011. Airline transit passenger percentage calculations. Tech. rep., Constraint Technologies International, available from http://www.infotech.monash.edu.au/~wallace/airline_benchmarks/.
- Florian, M., 1976. *Traffic Equilibrium Methods*. Springer-Verlag.
- Garrow, L., Jones, S., Parker, R., 2007. How much airline customers are willing to pay: An analysis of price sensitivity in online distribution channels. *Journal of Revenue and Pricing Management* 5, 271–290.
- Gopalan, R., Talluri, K., 1998. Mathematical models in airline schedule planning: A survey. *Annals of Operations Research* 76, 155–185.
- Jacobs, T., Smith, B., Johnson, E., 2008. Incorporating network flow effects into the airline fleet assignment process. *Transportation Science* 42 (4), 514–529.
- Klabjan, D., 2005. Large-scale models in the airline industry. In: Desaulniers, G., Desrosiers, J., Solomon, M. M. (Eds.), *Column Generation*. Springer, pp. 163–196.
- Koppelman, F., Coldren, G., Parker, R., 2008. Schedule delay impacts on air-travel itinerary demand. *Transportation Research Part B* 42, 263–273.
- Lohatepanont, M., Barnhart, C., 2004. Airline schedule planning: Integrated models and algorithms for schedule design and fleet assignment. *Transportation Science* 38 (1), 19–32.
- McGill, J. I., Van Ryzin, G., J., 1999. Revenue management: Research overview and prospects. *Transportation Science* 33 (2), 233–256.
- Ratliff, R. M., Rao, B. V., Narayan, C. P., Yellepeddi, K., 2008. A multi-flight recapture heuristic for estimating unconstrained demand from airline bookings. *Journal of Revenue and Pricing Management* 7 (2), 153–171.
- Sherali, H., Bish, E., Zhu, X., 2006. Airline fleet assignment concepts, models, and algorithms. *European Journal of Operational Research* 172 (1), 1–30.
- Solomon, M. M., 1987. Algorithms for the vehicle routing and scheduling problems with time window constraints. *Operations Research* 35, 254–265.
- US DOT, 2008. Research and Innovative Technology Administration (RITA), U.S. Department of Transportation (US DOT), Origin and Destination Survey DB1BCoupon 2008 Q1 data set. http://www.transtats.bts.gov/DL_SelectFields.asp?Table_ID=289&DB_Short_Name=Origin%20and%20Destination%20Survey.
- Walker, J., 2006. Time of day in airline passenger demand, presentation at the INFORMS Annual Meeting, Pittsburgh PA.
- Wojahn, O., 2002. The impact of passengers' preferences regarding time and service quality on airline network structure. *Journal of Transport Economics and Policy* 36, 139–162.
- Yan, S., Tang, C.-H., Lee, M.-C., 2007. A flight scheduling model for taiwan airlines under market competitions. *Omega* 35 (1), 61–74.
- Yan, S., Tseng, C.-H., 2002. A passenger demand model for airline flight scheduling and fleet routing. *Computers and Operations Research* 29, 1559–1581.