

# An Algorithm for Calculating Consistent Itinerary Flows

ROBERT L. PHILLIPS, DEAN W. BOYD and THOMAS A. GROSSMAN, JR.

*Decision Focus Incorporated, Los Altos, California 94022*

*As airlines have moved toward more concentrated "hub-and-spoke" schedules, connecting passengers have become an increasingly important part of their loads and revenues. Yet, many optimization approaches to such airline planning problems as revenue management and aircraft assignment have ignored the "cross-leg" effects of seating capacity on connecting traffic and the associated revenue. This paper presents a simple approach for calculating consistent on-board itinerary flows and the corresponding loads, revenues and marginal seat values given a set of underlying demands and an assignment of seating capacities to legs. An application of this approach to a large-scale airline network is presented and the implication of the results for decision-making discussed. Although the approach is presented in the context of airline planning, it is applicable to any transportation system in which connecting traffic plays a significant part as well as to hotel and automobile rentals.*

## 1. INTRODUCTION

Airlines have long recognized that the development of efficient connections is an important factor in filling their aircraft. By designing route structures based upon the "hub-and-spoke" philosophy, a single leg may serve hundreds of passenger markets. For example, a leg from Chicago to Houston may carry passengers traveling from Grand Rapids to Houston as well as passengers traveling from Chicago to Lubbock. In a well-designed hub-and-spoke system, the number of markets being served by an airline through either direct or connecting service will be vastly greater than the number of legs it offers. Consider the case of  $n$  legs arriving at a hub, each of which is timed to connect with  $n$  legs departing the hub. Assuming that all of the possible connections are valid, the airline can serve  $n^2 + 2n$  markets with only  $2n$  legs. Thus, the economic incentive to develop such systems in order to best utilize the aircraft fleet is extremely strong. This observation has, of course, been borne out by the proliferation of highly concentrated "hub-and-spoke" systems in the wake of deregulation.

A natural consequence of a hub-and-spoke system is that each leg carries a mix of passengers from many different markets, each paying different fares. It also means that the marginal revenue resulting

from accommodating more passengers on a given leg (through, for example, assigning a larger aircraft) is difficult to estimate and is likely not to be equal to the average revenue. Indeed, both the load and the revenue on a given leg are likely to be dependent both on the capacity (i.e., number of seats) offered on that leg as well as the capacities offered on all legs connecting with that leg.

In fact, the situation is even more complex. Consider a passenger who wishes to travel from station A to station C. Assume his chosen itinerary is Flight 1 from station A to station B connecting with Flight 2 from station B to station C. To be accommodated on this itinerary, there needs to be a seat available on Flight 1 and a seat available on Flight 2. As just discussed, however, the availability of such seats will depend on the demands for connecting traffic and the seating capacities offered on all legs connecting with 1 and with 2. Therefore, the load on Flight 1 depends not only on the capacity of Flight 2 but also on the capacities offered for all legs connecting with Flight 2. By extending this reasoning it can be shown that, at least theoretically, the passenger load on every leg in a system can depend not only on the seating capacity of that leg but also on the seating capacities of every other leg in the system.

The implication is that decisions such as deter-

mining the optimal aircraft to assign to a given leg cannot be made if the effects on the loads on other legs are ignored. Nevertheless, cross-leg effects have not been included in previous formulations of the aircraft assignment problem such as those described in DANTZIG<sup>[2]</sup> LEGENDRE and MINOUX<sup>[5]</sup> and DASKIN and PANAYOTOPOULOUS.<sup>[3]</sup> Similarly, such cross-leg effects might have significant impact on the optimal setting of capacity limits for various fare classes on a leg. Nonetheless, they have apparently not been incorporated in approaches to the capacity control problem such as those discussed by BELOBABA.<sup>[1]</sup>

In this paper, we address the problem of determining consistent itinerary flows for an airline. We take as given the seating capacity per leg and an underlying demand for travel on each itinerary being offered. We seek to determine the set of on-board loads and the number of passengers travelling on each itinerary who are accommodated. Key issues influencing the results include the rate at which passengers book on different itineraries and whether or not passengers who cannot be accommodated on their itinerary of choice seek to rebook on other itineraries serving the same market. We should emphasize that our approach is descriptive: we are seeking to determine what the loads and itinerary flows would be for an airline in the absence of intervention. However, we do discuss the implications for the underlying problem raised by "yield management" approaches to setting booking limits.

Although we state both the problem and our approach in the context of passenger airline planning, the approach is applicable to any transportation system with a regular schedule in which connecting traffic plays a significant part, such as railroads, buses, and freight carriers. In addition, it is applicable to situations in which a fixed number of units of capacity (hotel rooms, rental cars, etc.) are offered for rent for different periods starting at different times. These applications are discussed in Section 6.

In Section 2, we define the problem and describe an algorithm for determining the consistent itinerary flows for a service network and corresponding set of seating capacities. The algorithm is developed assuming that passengers book at the same rate on all itineraries and unaccommodated passengers do not seek to rebook. The operation of the algorithm is illustrated with a simple example. In Section 3, we extend the algorithm to the case of variable booking curves and rebooking. In Section 4, we discuss the results of an application of the algorithm to a large schedule. Section 5 discusses

the implications of airline "yield management" actions for the calculation of itinerary flows. Section 6 summarizes the major results of this research and discusses possible further applications and extensions.

## 2. PROBLEM FORMULATION AND SOLUTION ALGORITHM

IN THIS section, we formulate (by means of an illustration) the problem of calculating on-board itinerary flows. In particular, we show how the intuitively appealing method of Proportional Flows can result in inconsistent loads and itinerary flows. We then introduce an algorithm for determining consistent loads and itinerary flows for all the legs within a schedule.

### 2.1. Problem Description

The problem of determining consistent itinerary flows is best illustrated by means of an example. Consider the extremely simple system shown in Figure 1. Here, there are two legs serving a network of three cities. Flight 1 flies from city A to city B and Flight 2 from city B to city C. We assume that these two legs serve three markets; the A-B market is served by Flight 1, the B-C market by Flight 2, and the A-C market by Flight 1 connecting to Flight 2. Assume that aircraft with seating capacities of 100 have been assigned to each leg and that the unconstrained market demands are as shown in the figure. The questions we seek to answer include: What will be the actual on-board load on each leg? How many passengers on each itinerary will be accommodated? What would be the value of an additional seat on each leg?

A naive approach to determining on-board loads and itinerary flows would be to take a three-step approach; first determine the unconstrained demands, then the loads, and finally the itinerary flows. *Unconstrained demand* refers to the number of passengers who would book on a given leg in the absence of any capacity constraints. For example, the unconstrained demand for Flight 1 is 120 and



Market	Unconstrained Demand
A - B	60
B - C	100
A - C	60

Fig. 1. Example system.

the unconstrained demand for Flight 2 is 160. Since the unconstrained demands are greater than the seating capacities assigned to these legs, we might assume that each leg would operate at a 100% load factor, that is, each leg would carry a load of 100 people. To calculate the itinerary flows, we might assume that the itinerary flow on a leg is proportional to the unconstrained demand for that itinerary. Then, for example, the itinerary flows on each leg would be:

Flight	Market	Fraction	×	Capacity	=	Itinerary Flow
1	A-B	(60/120)		100		50
1	A-C	(60/120)		100		50
2	B-C	(100/160)		100		62.5
2	A-C	(60/160)		100		37.5

The itinerary flows calculated in this way for the A-C market immediately show a discrepancy. The number of people in this market carried on Flight 1 is 50 while the number carried on Flight 2 is 37.5. Obviously, the number of people carried on each leg in the A-C market must be the same, yet this method of calculation produces different numbers. The problem of calculating consistent itinerary flows is to determine the set of on-board itinerary flows and loads for a network consistent with the underlying itinerary demands and the capacities assigned to each leg.

**2.2. Calculating Consistent On-Board Loads: The Basic Case**

In this section, we describe an algorithm for calculating consistent on-board loads in the basic case; that is, assuming:

1. Booking patterns are the same for all itineraries.
2. Passengers not accommodated on their desired itinerary are lost (i.e., do not rebook on some other itinerary).

Both of these assumptions are relaxed in Section 3.

We begin by establishing some definitions. A *leg* is a single nonstop service between an origin and a destination. A leg may be a nonstop leg itself or may be a single segment of a multistop leg. The *capacity* allocated to the leg is the number of seats available in the aircraft that flies that leg. An *itinerary* is a set of one or more legs that serves an origin-to-destination market. Thus, for example, all of the combinations of legs that a passenger can use to travel from ORD to SFO constitute possible itineraries in the ORD-SFO market.

Markets are specified using two indices—*i* and *j*

denote the market from origin *i* to destination *j*. Itineraries are defined using three subscripts, of which the first two define the market served by that itinerary. Thus  $I_{ijk}$  is the *k*th itinerary serving market *ij*. Legs are indexed by a single subscript *m*. The mapping between itineraries and legs is defined by an index function  $\delta_{ijkm}$ , where  $\delta_{ijkm} = 1$  indicates that leg *m* is part of itinerary  $I_{ijk}$  while  $\delta_{ijkm} = 0$  indicates that leg *m* is not part of itinerary  $I_{ijk}$ . A set of  $\delta_{ijkm}$ 's defines a specific leg network.

We represent the behavior of passengers seeking to make reservations for itineraries by a "booking curve." Let  $f_{ijk}(t)$  be the fraction of passengers wishing to travel on itinerary  $I_{ijk}$  who have sought to make a reservation by time *t*. The total number of prospective passengers who have sought to make reservations by time *t* is given by  $d_{ijk}f_{ijk}(t)$ , where  $d_{ijk}$  is the total demand on itinerary  $I_{ijk}$ . We assume that the time period of interest is normalized so that  $f_{ijk}(0) = 0$  and  $f_{ijk}(1) = 1$  for all itineraries. Further, we assume initially that bookings occur at a constant rate over time. This implies that  $f_{ijk}(t) = t$  for all itineraries. This assumption will be relaxed later.

We are given a set of unconstrained itinerary demands,  $d_{ijk}$ . We are also given an assignment of capacity to legs such that  $c_m$  is the seating capacity assigned to leg *m*. We wish to determine  $l_{ijk}$  the number of passengers accommodated on itinerary  $I_{ijk}$ . Given this, we can easily calculate other quantities of interest, such as  $L_m$ , the constrained on-board load on leg *m*:

$$L_m = \sum_i \sum_j \sum_k \delta_{ijkm} l_{ijk}.$$

It is equally straightforward to calculate total system revenue and on-board yields once the itinerary flows have been calculated.

Our method for calculating consistent on-board itinerary flows is to gradually increase the parameter *t* from 0 toward 1, calculating the total demand for seats on each leg as we go. Assume that the unconstrained demand for at least one leg is greater than its seating capacity. Then, as *t* increases, the demand for seats on some leg will come to equal its capacity. When this occurs, we say the leg has become *full*. For the first leg that becomes full, we can calculate the on-board flows for each itinerary including that leg by using the proportional rule described earlier. We say that these itineraries are now *blocked*. As we continue to increase *t*, more legs will become full. As each leg becomes full, we can calculate the on-board flows for each newly

blocked itinerary. Ultimately either all legs will be full, in which case we are done, or  $t$  will reach 1, in which case the flows for all unblocked itineraries will be equal to their unconstrained demand.

To specify the algorithm, we need to define two additional variables. At each iteration,  $k_m$  represents the seating capacity on leg  $m$  not taken up by currently blocked itinerary flows. Similarly,  $u_{ijk}$  is a binary variable that, on each iteration, keeps track of whether or not itinerary  $I_{ijk}$  is blocked. By convention,  $u_{ijk} = 0$  indicates that itinerary  $I_{ijk}$  is blocked while  $u_{ijk} = 1$  indicates that itinerary  $I_{ijk}$  is not blocked.

Given this notation, we can describe the algorithm as follows:

**Step 0.** Initialize:

$$\begin{aligned} u_{ijk} &= 1 && \text{for all itineraries} \\ k_m &= c_m && \text{for all legs} \\ l_{ijk} &= 0 && \text{for all itineraries} \\ t &= 0 \end{aligned}$$

**Step 1.** For each leg  $m$ , calculate:

$$\begin{aligned} r_m &= \sum_i \sum_j \sum_k \delta_{ijkm} u_{ijk} d_{ijk} \\ t_m &= k_m / r_m \end{aligned}$$

( $r_m$  is the remaining demand in unblocked itineraries.)

**Step 2.** For the system, determine:

$$m^* = \arg \min(t_m)$$

( $m^*$  is the index of the next leg that will become full.)

If  $t_{m^*} \geq 1$ , go to Step 5. (No more legs will become full.)

If  $t_{m^*} < 1$ , go to Step 3.

**Step 3.** For all  $i, j, k$  such that  $\delta_{ijkm^*} = 1$  and  $u_{ijk} = 1$ , set:

$$\begin{aligned} l_{ijk} &= t_{m^*} d_{ijk} \\ u_{ijk} &= 0 \end{aligned}$$

(For all newly blocked itineraries, determine their on-board flows and indicate that they are now blocked.)

If all itineraries have been blocked (ie.,  $u_{ijk} = 0$  for all  $i, j, k$ ) go to Step 5.

Otherwise go to Step 4.

**Step 4.** For all legs  $m$ , set:

$$k_m = c_m - \sum_i \sum_j \sum_k \delta_{ijkm} l_{ijk}$$

(Recalculate remaining capacity.)

Go to Step 1.

**Step 5.** Set all  $l_{ijk} = l_{ijk} + u_{ijk} d_{ijk}$

(Set the onboard flow equal to unconstrained demand for all unblocked itineraries.)

Stop.

This is a relatively broad-brush portrayal of the underlying algorithm. It is easy to see how the implementation could be made much more efficient. For example, we can identify *ab initio* the set of legs for which the unconstrained demand is less than the capacity, that is, those for which:

$$\sum_i \sum_j \sum_k \delta_{ijkm} d_{ijk} \leq c_m.$$

Since these legs will never become full, they do not need to be included in the calculations in Steps 1 and 4. We term them *slack legs*. In addition, all itineraries incorporating *only* slack legs will never be blocked and their on-board flows can be set to their unconstrained demands and do not need to be included in any of the calculations.

### 2.3. Example Application

Consider the eight-station, seven-leg example shown in Figure 2. We assume that all of the legs connect, but we will consider only itineraries consisting of one or two legs. There are thus 13 markets, each served by a single itinerary. These markets with their associated demands are shown in Table I. Table II shows the seating capacities associated with each leg.

We begin by calculating  $r_m$  for all legs as described in Step 1. In so doing, we find that legs 1, 3, and 6 are slack legs and can be excluded from further calculations. The situation at step 1 of Iteration 1 is shown below:

*Iteration 1*

Leg	$r_m$	$k_m$	$t_m$	Status
1	75	100	—	Slack
2*	195	150	0.769	Open
3	40	100	—	Slack
4	335	315	0.940	Open
5	105	100	0.952	Open
6	140	150	—	Slack
7	110	100	0.909	Open

Leg 2 is the first leg to become full, so we set  $m^* = 2$  and  $t_{m^*} = 0.769$ . We can now calculate the onboard flows for those itineraries that incorporate leg 2:

$$l_2 = 0.769 \times 110 = 84.6$$

$$l_5 = 0.769 \times 85 = 65.4.$$

We update  $k_m$  and begin Iteration 2. For example,  $r_4$  equals the unconstrained demand for leg 4

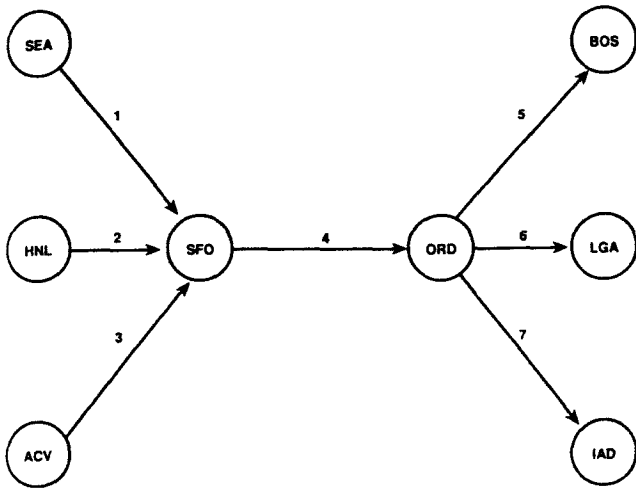


Fig. 2. Seven-leg, eight-station example system.

TABLE I

Example Problem: Itineraries, Net Fares, and Loads

<i>l</i>	Market	Itinerary	Net Fare	Unconstrained Demand
1	SEA-SFO	(1)	129	45
2	HNL-SFO	(2)	410	110
3	ACV-SFO	(3)	69	25
4	SEA-ORD	(1, 4)	250	30
5	HNL-ORD	(2, 4)	520	85
6	ACV-ORD	(3, 4)	360	15
7	SFO-BOS	(4, 5)	430	40
8	SFO-LGA	(4, 6)	410	35
9	SFO-IAD	(4, 7)	325	30
10	SFO-ORD	(4)	265	100
11	ORD-BOS	(5)	155	65
12	ORD-LGA	(6)	135	105
13	ORD-IAD	(7)	144	80

TABLE II

Example Problem: Seating Capacities

Leg	Seating Capacity
1	100
2	150
3	100
4	315
5	100
6	150
7	100

from itineraries that do not involve the full leg 2:

$$r_4 = 30 + 15 + 40 + 35 + 30 + 100 = 250.$$

Similarly,  $k_4$  equals the seats remaining on leg 4 after subtracting those assigned to people traveling on itineraries involving the full leg 2:

$$k_4 = 315 - 65.4 = 249.6.$$

The full results for Iteration 2 are given below:

Iteration 2

Leg	$r_m$	$k_m$	$t_m$	Status
1	—	—	—	Slack
2	—	—	—	Full
3	—	—	—	Slack
4	250	249.6	0.998	Open
5	105	100	0.952	Open
6	—	—	—	Slack
7*	110	100	0.909	Open

This time,  $m^* = 7$  and  $t_{m^*} = 0.909$ . We can now calculate the on-board flows for two more itineraries:

$$l_9 = 0.909 \times 30 = 27.3$$

$$l_{13} = 0.909 \times 80 = 72.7.$$

Proceeding to Iteration 3 we obtain:

Iteration 3

Leg	$r_m$	$k_m$	$t_m$	Status
1	—	—	—	Slack
2	—	—	—	Full
3	—	—	—	Slack
4	220	222.3	—	Slack
5*	105	100	0.952	Open
6	—	—	—	Slack
7	—	—	—	Full

and  $m^* = 5$ ;  $t_{m^*} = .95$ ;

$$l_7 = 0.95 \times 40 = 38.00$$

$$l_{11} = 0.95 \times 65 = 61.75.$$

Note that on this iteration, Leg 4 has “gone slack,” that is, the remaining demand is less than the available capacity.

We have now completed the process. All legs are either labeled full or slack. All itinerary flows that have not been explicitly calculated can be set to their unconstrained demands. The final results are shown in Table III.

TABLE III

Resulting On-Board Loads and Marginal Values

Leg	Capacity	Unconstrained Demand	Actual Load	Capacity Marginal Value (\$/Seat)
1	100	75	75.0	0
2	150	195	150.0	458
3	100	40	40.0	0
4	315	335	310.5	0
5	100	105	100	260
6	150	140	140	0
7	100	110	100	194

**2.4. Calculating Marginal Values of Capacity**

Once the itinerary flows have been calculated, we can determine the marginal value of capacity for each leg. The marginal value of capacity for a leg is defined as the additional revenue that would result from an additional seat on that leg. Denote the marginal revenue on leg  $m$  by  $\lambda_m$ . If leg  $m$  is not full, then  $\lambda_m = 0$ , since addition of a seat to that leg will not result in any added revenue. If leg  $m$  is full, then its marginal value depends on the unconstrained demands and net fares associated with the itineraries that were blocked as a result of the leg becoming full. If the leg does not share any of these itineraries with any other full legs, it is straightforward to calculate  $\lambda_m$ : it is just the weighted average of the net fares for the itineraries that were blocked when leg  $m$  became full. To illustrate, consider leg 2 from the example in Section 2.3. Itineraries 2 and 5 were blocked by leg 2. From the solution, we can see that neither of those itineraries included any other full legs. Therefore, the marginal revenue associated with leg 2 is the weighted average of the net fares for those two itineraries, weighted by their corresponding unconstrained demands:

$$\lambda_2 = [110/(110 + 85)] \times \$410 + [85/(110 + 85)] \times \$520 = \$458.$$

When a full leg shares one or more itineraries with other full legs, the computation of the marginal revenue associated with that leg is more complex. Specifically, there is the possibility that increasing capacity on one leg will allow passengers to book on itineraries that include another full leg, thereby displacing other passengers on that leg. To develop a technique for calculating leg marginal values, we need to establish some new notation.

For the remainder of this section only, we assume that the legs have been numbered in the order in which they become full. Slack legs are numbered in arbitrary order after full ones. Define  $Z(m)$  as the set of itineraries that are blocked by leg  $m$ . (In the example above,  $Z(2) = \{2, 5\}$ .) By convention, define  $Z(m) = \emptyset$  for slack legs. For each leg, define  $D_m$  as the total unconstrained demand on all itineraries blocked by  $m$ . That is:

$$D_m = \sum_{I_{ijk} \in Z(m)} d_{ijk}.$$

Note that  $D_m = 0$  for a slack leg. Now, define  $v_m$  by:

$$v_m = \sum_{I_{ijk} \in Z(m)} d_{ijk} p_{ijk} / D_m$$

where  $p_{ijk}$  is the net fare associated with itinerary  $I_{ijk}$ . Finally, define  $w_{mn}$  by:

$$w_{mn} = \begin{cases} 0 & \text{for } n \leq m \\ \sum_{I_{ijk} \in Z(m)} \delta_{ijkn} d_{ijk} / D_m & \text{for } n > m \end{cases}$$

$w_{mn}$  defines the interaction of each full leg with the legs that are not yet full. For  $n > m$ ,  $w_{mn} > 0$  means that at least one itinerary blocked by leg  $m$  also includes leg  $n$ . The magnitude of  $w_{mn}$  is the proportion of the total unconstrained demand for all itineraries that were blocked by leg  $m$  that is represented by itineraries that include leg  $n$ . By convention, set  $v_m = 0$  and  $w_{mn} = 0$  whenever  $D_m = 0$ . The marginal values must obey the relationship:

$$\lambda_m = v_m - \sum_n w_{mn} \lambda_n \quad \text{for all } m.$$

This relationship can be written in matrix form as:

$$(I + W)\underline{\lambda} = \underline{v}.$$

Since  $I + W$  is upper triangular, the corresponding set of equations can be solved for  $\underline{\lambda}$  by sequential substitution.

It is important to realize that, under the CIF model, total revenue is not guaranteed to be a concave function of leg capacity. For example, consider the simple system shown in Figure 1. Assume that the unconstrained itinerary demands are as shown and that the capacity on leg 1 is 100. Then  $\lambda_2$ , the marginal value associated with leg 2 as a function of  $c_2$ , is shown below:

$$\begin{aligned} 0 \leq c_2 \leq 133.33 & \quad \lambda_2 = (5p_{BC} + 3p_{AC} - 3p_{AB})/8 \\ 133.33 \leq c_2 \leq 160 & \quad \lambda_2 = p_{BC} \\ c_2 \geq 160 & \quad \lambda_2 = 0. \end{aligned}$$

If  $p_{AB} + p_{BC} > p_{AC}$  (a not unreasonable occurrence), then the marginal value of capacity on leg 2 will decrease as  $c_2$  is increased from less than 133.33 to a value between 133.33 and 160. Clearly, total revenue would, in this case, not be a concave function of leg capacity. This would imply that the deterministic aircraft problem, such as studied by Dantzig<sup>[2]</sup> and Legendre and Minoux<sup>[5]</sup> is not, in general, guaranteed to have a concave objective function when full passenger flows are taken into account. Hence, the deterministic aircraft assignment problem accounting for full passenger flows may not be solvable by techniques that assume concavity of the objective function.

Finally, note that nothing in the calculation of the marginal values guarantees that they will be non-negative. It is in fact, straightforward to con-

struct examples in which  $\lambda_m < 0$  for one or more legs in a system. In this case, increasing seating capacity on a leg might result in a decrease in system revenue!

### 3. EXTENSIONS TO THE BASIC ALGORITHM

WE HAVE now developed the basic algorithm and demonstrated its application in the simple case in which all itineraries book at the same uniform rate and rejected passengers do not rebook. In this section, we discuss how the algorithm might be extended to relax both of these requirements in turn.

#### 3.1. Variable Booking Curves

In the previous section we showed how to calculate consistent on-board loads assuming that passengers book in at the same uniform rate for all itineraries. In other words, we assumed that  $f_{ijk}(t) = t$  for all itineraries. Some reflection will show that exactly the same algorithm will work if all itineraries have the same monotonically non-decreasing booking curve—that is, if  $f_{ijk}(t) = f(t)$  for all  $i, j, k$ . The basic algorithm may not work, however, if different itineraries have different booking curves. Fortunately, it can be easily modified to accommodate this possibility.

We assume, as before, that the booking curve  $f_{ijk}(t)$  is monotonically non-decreasing and that  $f_{ijk}(0) = 0$  and that  $f_{ijk}(1) = 1$  for all  $i, j, k$ . The endpoint requirements are merely normalization conventions. Given these representations of the booking curves, we can determine the on-board flows using the same algorithm as in Section 2 with Steps 1 and 2 modified as shown below:

*Step 1.* For each leg  $m$ , determine  $\tau_m$  such that:

$$\sum_i \sum_j \sum_k \delta_{ijk} u_{ijk} f_{ijk}(\tau_m) d_{ijk} = k_m$$

*Step 2.* For the system, determine:

$$m^* = \operatorname{argmin}(\tau_m)$$

$$t_{m^*} = \tau_m$$

If  $t_{m^*} \geq 1$ , go to Step 5 (no more legs will become full)

If  $t_{m^*} \leq 1$ , go to Step 3.

The major complication in solving the problem with variable booking curves is the calculation of  $\tau_m$  in Step 1. Note that if  $f_{ijk}(t) = t$  for all  $t$ , then we have the basic algorithm. Otherwise,  $\tau_m$  will need to be explicitly calculated for all legs for which  $r_m > k_m$ . The form of the calculation depends upon the form of the function  $f_{ijk}(t)$ . Assume for example that  $f_{ijk}(t)$  is piecewise linear with:

$$f_{ijk}(t) = f_{ijk}(t_h) + m_{ijk}^h (t - t_h) \quad \text{for } t_h \leq t < t_{h+1}.$$

To calculate  $\tau_m$ , assume that the current value of  $t$  is in the interval  $[t_h, t_{h+1})$ . Then, let

$$\tau = c_m - \sum_i \sum_j \sum_k \delta_{ijk} m_{ijk}^h \left[ \frac{l_{ijk}(f_{ijk}(t_h) - m_{ijk}^h t_h) u_{ijk} d_{ijk}}{\sum_i \sum_j \sum_k \delta_{ijk} m_{ijk}^h u_{ijk} d_{ijk}} \right].$$

If  $\tau \leq t_{h+1}$ , then  $\tau_m = \tau$ . Otherwise, if  $t_{h+1} = 1$ , set  $\tau_m = 1$  and stop, else set  $h \leftarrow h + 1$  and redo the above calculation.

If the booking curves have more complex forms, then robust line search techniques such as binary search may be necessary to calculate  $\tau_m$  (see LUENBERGER<sup>[6]</sup> for an overview of such techniques). The necessity to use such techniques can substantially increase the time required to find the solution.

#### 3.2. Passenger Reflow

So far we have assumed that passengers not accommodated on their desired itinerary are lost to the system. In actuality, this is often untrue. A passenger who cannot book on a desired itinerary may choose to book on other itineraries serving the same market (assuming, of course, that other such itineraries exist). We can capture this effect through a passenger reflow model.

Consider a market served by  $N > 1$  itineraries. Suppose a particular itinerary  $I_{ijk}$  has just become blocked on leg  $m$  at time  $t_{m^*}$ . There remain  $(1 - f_{ijk}(t_{m^*}))d_{ijk}$  passengers in this market who wish to travel via the  $I_{ijk}$  itinerary but cannot be accommodated. Some fraction of these passengers, say  $\alpha_{ijk}$ , will not seek to rebook. The remaining fraction  $1 - \alpha_{ijk}$ , will reflow over the remaining unblocked itineraries serving the  $i$  to  $j$  market. Assume that  $b_{ijkq}$  is the fraction of passengers from itinerary  $I_{ijk}$  who seek to rebook, that rebook on itinerary  $I_{ijq}$ . Then, we can capture this effect by modifying  $d_{ijq}$ , and  $f_{ijq}(t)$  for  $t \geq t_{m^*}$  as shown:

$$d'_{ijq} = d_{ijq} + b_{ijkq} d_{ijk} (1 - f_{ijk}(t_{m^*})) (1 - \alpha_{ijk}) \quad (2)$$

$$f_{ijq}(t) = [f_{ijq}(t) d_{ijq} + (f_{ijk}(t) - f_{ijk}(t_{m^*})) \cdot (1 - \alpha_{ijk}) b_{ijkq} d_{ijk}] / d'_{ijq} \quad (3)$$

To demonstrate this type of reflow, consider the two booking curves shown in Figure 3. We assume that these two booking curves are for two different itineraries serving the same market and that the demands on these itineraries are given by  $d_{i,j1} = 10$  and  $d_{i,j2} = 20$ . The cumulative booking curves are represented as piecewise linear with intervals defined at  $(t_0, t_1, t_2, t_3) = (0, 1/3, 2/3, 1)$ . As shown,

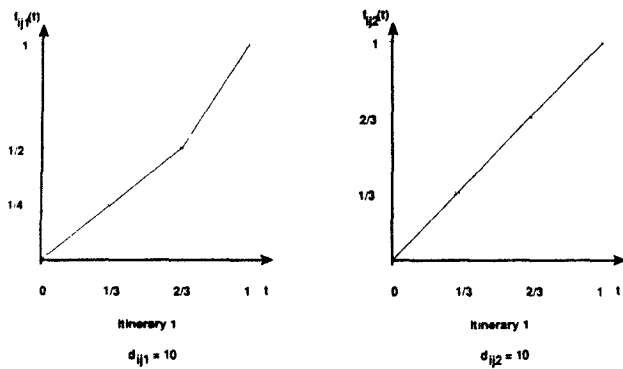


Fig. 3. Cumulative booking curves for reflow example.

we assume that  $f_{i,j1}(t_s) = (0, 1/4, 1/2, 1)$  and  $f_{i,j2}(t_s) = (0, 1/3, 2/3, 1)$  for  $s = 0, 1, 2, 3$ . We further assume that  $\alpha_{i,j1} = \alpha_{i,j2} = 0$ . That is, for either itinerary all passengers that cannot be accommodated will seek to rebook on the other itinerary. Finally, assume that itinerary 2 is the first of these two itineraries to become blocked and it becomes blocked at  $t_{m^*} = 1/2$ . How then should  $d_{i,j1}$  and  $f_{i,j1}(t)$  be modified to account for reflow?

Since there are only two itineraries serving this market,  $b_{i,j21} = 1$ . Resetting  $t_1 = 1/2$ , we can calculate the updated values of  $d'_{i,j1}$  and  $f'_{i,j1}(t)$  for  $t \geq 1/2$  as shown below:

$$d'_{i,j1} = 10 + (1/2) \times 20 = 20$$

$$f'_{i,j1}(1/2) = [(3/8) \times 10] / 20 = 0.1875$$

$$f'_{i,j1}(2/3) = [(1/2) \times 10 + (2/3 - 1/2) \times 20] / 20 = 0.42$$

$$f'_{i,j1}(1) = 1$$

with  $f'_{i,j1}(t)$  linear between these critical points. After these modifications, the algorithm continues as before.

An issue so far unaddressed is the determination of the reflow parameters  $b_{i,jkq}$ . Of course, they might be entered exogenously. However, for a large network, estimation of these parameters and direct input would be a job of almost unmanageable complexity. In addition, it seems that the reflow coefficients should depend on the set of itineraries that are currently blocked. A simple, but realistic approach is to estimate the reflow coefficients based on the fractions of passengers in each market who choose each itinerary. Let  $e_{i,jk}$  be the fraction of people in the  $i$  to  $j$  market who choose itinerary  $k$ . Then we can calculate  $e_{i,jk}$  by:

$$e_{i,jk} = d_{i,jk} / \sum_n d_{i,jn}$$

Assume that itinerary  $k$  has just become blocked. Then, we might estimate the reflow parameter  $b_{i,jkq}$  as shown in Equation 4.

$$b_{i,jkq} = e_{i,jq} u_{i,jq} / \sum_n e_{i,jn} u_{i,jn} \quad (4)$$

where  $u_{i,jn} = 1$  means that itinerary  $I_{i,jn}$  is not blocked while  $u_{i,jn} = 0$  means that itinerary  $I_{i,jn}$  is blocked.

This particular representation of reflow is only one of several ways in which the underlying phenomenon might be represented. It has the advantage of being simple and requiring no new data, with the exception of the  $\alpha_{i,jk}$ 's. It is easy to conceive of more complex models explicitly representing the sequential behavior of potential passengers seeking to make a reservation. Whether such models improve the accuracy of the overall approach remains to be seen.

#### 4. APPLICATION

IN THE previous sections we have shown that cross-leg effects can have a significant impact on by-leg loads and yields and hence on system revenue. In this section we will investigate how widespread these effects might be in a realistic airline schedule given realistic levels of demand. If they are widespread, then ignoring them in making aircraft assignment and yield management decisions might lead to suboptimal solutions. On the other hand, if they are not widespread, then cross-leg effects might well be ignored in making such decisions without substantial loss.

To determine the importance of consistent itinerary flows in a real-world setting, we investigate the results of the algorithm for a large, realistic (though hypothetical) airline schedule for a domestic U.S. carrier. This schedule contains 2456 legs serving 10,342 markets via 25,059 itineraries. Markets with demand less than one passenger per day are ignored. The schedule was devised to contain substantial connection possibilities at fifteen hubs. Legs into and out of these hubs were timed to provide arrival and departure banks. Once the schedule had been devised, a standard market model was used to determine the corresponding demands in each of the 10,342 markets. Demand in each market was distributed among possible itineraries using relative preference weights for the itineraries. Aircraft were chosen to serve the various legs from a fleet consisting of 10 different aircraft types. The assignment of the fleet to the schedule was optimized by solving a utilization-based assignment problem (see Legendre and

Minoux.<sup>[5]</sup> Since the size of the fleet was limited, a given leg was not necessarily assigned the most profitable aircraft type.

For this schedule and set of underlying demands, we calculated on-board loads and revenues in two different ways. The Consistent Itinerary Flows (CIF) method followed the algorithm described in Section 2. The Proportional Flow (PF) method calculated the on-board load on each leg as the minimum of the unconstrained load on that leg and the seating capacity being offered on that leg. In the PF method, the average yield for a leg was assumed to be the demand-weighted average of the yields for itineraries incorporating that leg. For itineraries incorporating more than one leg, the itinerary yield was allocated among legs in the itinerary based on relative mileage. For example, consider an itinerary with a total yield of \$120 consisting of leg 1 with length 1000 miles and leg 2 with length 500 miles. For each passenger on this itinerary, \$80 would be allocated to leg 1 and \$40 to leg 2. As discussed in Section 1, the loads and average yields generated by the PF methods do not necessarily correspond to any internally consistent set of itinerary flows.

To understand the differences in loads and yields produced by the CIF and the PF methods, it is instructive to divide the legs in a schedule into five

categories as shown in Table IV. For legs in Category 1, the unconstrained demand is less than the seating capacity and none of the itineraries incorporating that leg are blocked on other legs. In this case, CIF and PF will generate the same load and revenue estimates. For legs in Category 2, the unconstrained demand is less than the capacity but at least one of the itineraries incorporating that leg is blocked on some other leg. In this case, PF will overestimate both the on-board load and revenue relative to CIF. For legs in Categories 3 through 5, unconstrained demand is greater or equal to the seating capacity. For the legs in Category 3, at least one itinerary has been blocked on another leg, and, as a result, the CIF load is less than the seating capacity. Since the PF load will be equal to the seating capacity, the PF load will be greater than the CIF load for legs in this category. In addition, PF will misestimate revenue, through it may be either an over- or an under-estimation. For legs in Category 4, the CIF load is equal to the seating capacity and at least one itinerary was previously blocked. Obviously, in this case, both CIF and PF will produce loads equal to the seating capacity. However, PF will either over- or underestimate revenue. Finally, for legs in Category 5, the unconstrained demand is greater than the seating capacity and no itinerary is previously blocked. In

TABLE IV  
Flight Leg Categories

Category	Definition	Constrained On-Board Load	Total Leg Revenue	Comments
I	UOD <sup>a</sup> < capacity. No blocked itineraries	CIF = PF	CIF = PF	The revenue and loads on these legs are independent of approach
II	UOD < capacity. At least one blocked itinerary	CIF = PF	CIF < PF	For these legs, PF <i>overestimates</i> both load and revenue
III	UOD ≥ capacity. At least one itinerary previously blocked. CIF load < capacity	CIF < PF	CIF ≅ PF	PF <i>overestimates</i> load and misestimates revenue
IV	UOD ≥ capacity. At least one itinerary previously blocked. CIF load = capacity	CIF = PF	CIF ≅ PF	Load = capacity for both CIF and PF. PF misestimates on-board revenue
V	UOD ≥ capacity. No itinerary previously blocked. CIF load = capacity	CIF = PF	CIF = PF	The revenue and loads calculated for these legs are independent of approach

<sup>a</sup>UOD = unconstrained demand.

TABLE V  
System-Level Comparison: CIF vs. PF

	Method of Calculation		Change	% Change
	PF	CIF		
System revenue	\$16,967,000/day	\$16,918,000/day	\$49,000/day	0.29
System load factor	70.7	70.4	0.3	0.43

this case, both CIF and PF will calculate the correct on-board loads and revenues.

The overall accuracy of the PF method in estimating loads and revenues for a schedule will depend upon the proportion of legs in the schedule that fit into each of the categories described above. If a sizeable proportion of the legs fall into Categories 2 through 4, then we would expect the load and revenue estimates generated by PF to differ substantially from those generated by CIF. On the other hand, if the vast majority of legs fall into Categories 1 and 5, then the estimates generated by PF are likely to be close to those generated by CIF, at least on a system basis.

Table V shows the system load factors and revenues generated by PF and CIF. For the schedule, PF overestimated both system loads and revenue, but by less than one-half of one-percent. In other words, for this schedule, the system load factors and revenues estimated by CIF and PF are essentially the same.

Table VI shows the breakdown of legs into the five Categories of Table IV. About 54% of the legs fall in Categories 1 and 5, in which case CIF and PF produce the same loads and revenues. Of the remaining three Categories, Category 2 is the largest, containing almost 45% of the legs. Recall that, for legs in this category, the unconstrained demand is less than the seating capacity and at least one of the itineraries served by that leg is blocked by some other leg. Although the number of legs in this category is substantial, on the average PF overestimates loads and yields by less than 1% for legs in this category. Legs in Category 3 show the greatest difference between PF- and CIF-calculated loads and revenues. For legs in this category, PF overestimates loads by more than 2% and revenue by more than 1% relative to CIF. However, considerably less than 1% of all legs in the schedule fall into this category.

For this application, using the Proportional Flows method to calculate on-board loads and revenues

does not result in *system* load and revenue estimates significantly different from those calculated using CIF. What about the marginal revenue from added capacity? For the PF method, the marginal revenue from adding another seat on a leg is equal to the average revenue if the unconstrained load is greater than or equal to the seating capacity. Otherwise it is zero. Table VII summarizes the difference between PF and CIF in calculating marginal revenues for legs in each category. Note that PF will estimate a marginal revenue different from CIF only for legs in Categories 3, 4, and 5—which account for less than 3% of the legs in the system. Although the number of legs is small, the difference in marginal revenue can be very significant for some legs. Figure 4 shows the distribution of the percentage difference in marginal revenue as computed by the two different methods for the 71 legs in Categories 3, 4, and 5. The percentage shown is calculated as  $[(\text{PF value} - \text{CIF value}) / \text{CIF value}] \times 100$ . On the average, for these legs, PF estimates marginal revenues about 24% lower than CIF (24% is the mean of the distribution shown in Figure 4).

It is not clear what the overall effect on the differences between system revenues, loads, and marginal revenues would be of increasing system demand. In general, we would expect the differences to increase as legs move from Category 1—the largest class of legs in the example—into Categories 2, 3, and 4. However, on a system basis, some of the differences are likely to cancel out, and, as demand is increased further, legs will begin to move out of Categories 2, 3, and 4 into Category 5. The effect on system revenue of uniformly increasing system demand in the example while keeping the capacity assigned to each leg constant is shown in Figure 5. For demands 50% higher than the base, the difference in system revenue between PF and CIF has risen to almost 2%. However, this difference decreases slightly if the demand multiplier is increased further to 1.7.

TABLE VI  
Distribution of Legs by Category

Leg Category	No Legs	% Legs	PF Revenue (\$1000)	CIF Revenue (\$1000)	% Difference	PF Loads	CIF Loads	% Difference
I	1,287	52.4	7,950	7,950	0.0	103,378	103,378	0.0
II	1,098	44.7	7,896	7,848	0.6	102,851	102,042	0.9
III	7	0.3	120	118	1.7	1,330	1,301	2.2
IV	27	1.1	461.9	462.4	-0.1	5,122	5,122	0.0
V	37	1.5	539	539	0.0	6,210	6,210	0.0
Total:	2,456	100.0	16,967	16,918	0.3	218,891	218,053	0.4

**5. IMPLICATIONS FOR REVENUE MANAGEMENT**

THE DISCUSSION in the preceding sections has focused on the case in which an airline will accept any request for a booking as long as the passenger requesting the booking can be physically accommodated on every leg in the requested itinerary. However, for several years airlines have sought to increase their revenue by systematically refusing low-fare booking requests in the expectation of re-

ceiving a higher fare for the same seat from a later booking. This process, known variously as "yield management" or "revenue management" has attracted considerable attention as a mechanism by which airlines can increase revenue without increasing capacity. A good introduction to the basic issues involved can be found in Belobaba.<sup>[1]</sup> Active intervention on the part of an airline can clearly have an important effect on the composition of the traffic on each leg. In this section, we will describe

TABLE VII  
*Marginal Revenues by Category*

Category	No Legs	% Legs	Marginal Revenue		CIF vs PIF	Comments
			CIF	PF		
I	1287	52.4	0	0	CIF = PF	For these legs, UOD < capacity, thus marginal revenue = 0 independent of technique
II	1098	44.7	0	0	CIF = PF	
III	7	0.3	0	> 0	PF > CIF	PF will <i>overestimate</i> marginal revenue
IV	27	1.1	> 0	0	PF ≈ CIF	For these legs, PF will either over- or underestimate marginal revenue
V	37	1.5	> 0	> 0	PF ≈ CIF	

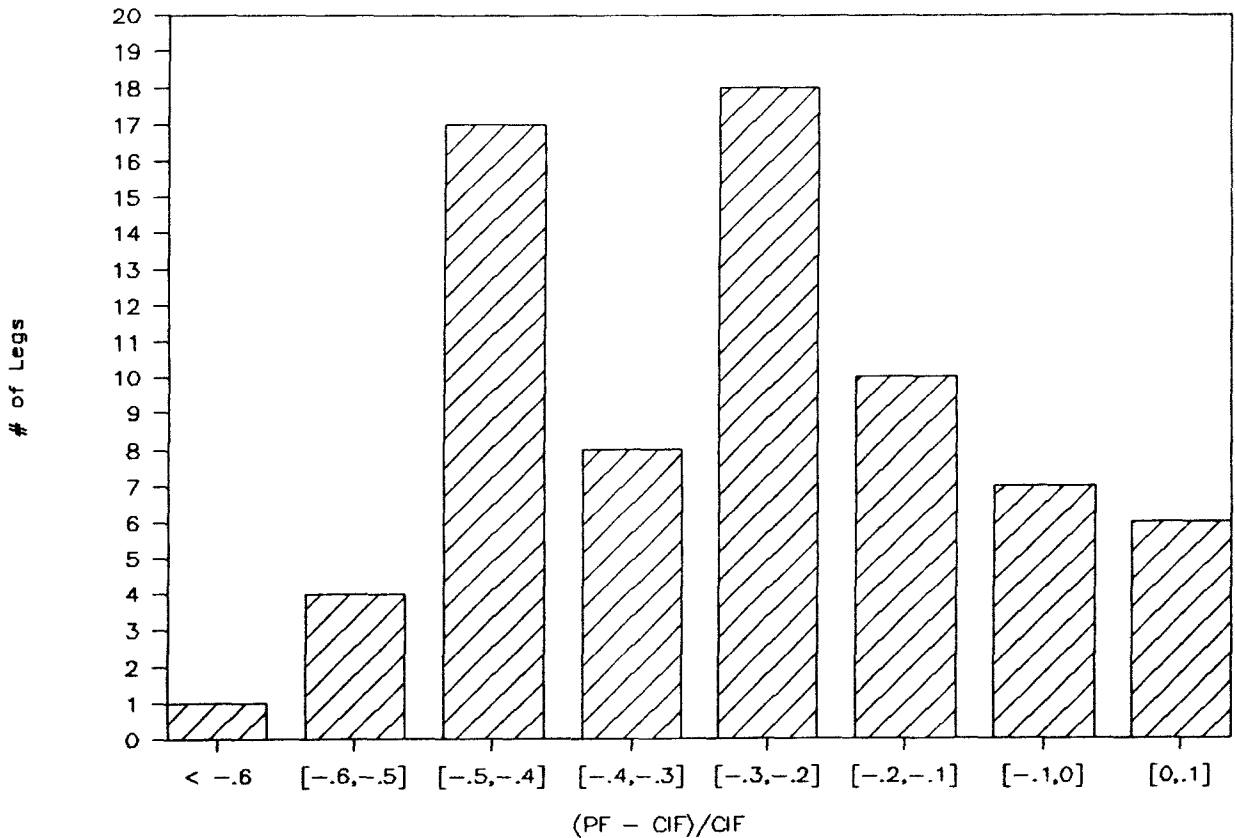


Fig. 4. Distribution of legs by difference in marginal revenue (Categories 3, 4, and 5).

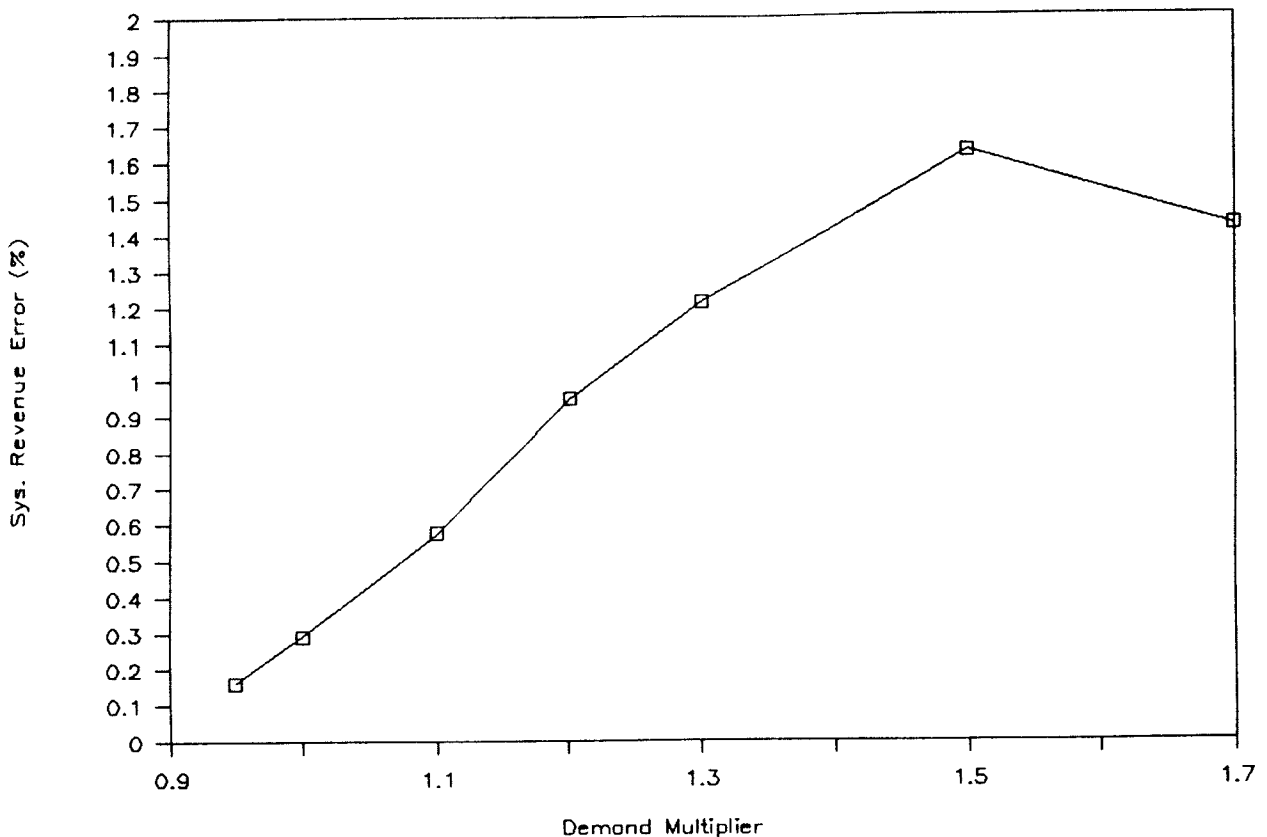


Fig. 5. Sensitivity of PF system revenue estimation error to aggregate demand level.

briefly how two different approaches to revenue management can affect traffic flows and marginal revenues.

A complete approach to airline revenue management must consider at least two characteristics of the underlying problem. The first characteristic is *uncertainty*. At the time a booking request is received, an airline is uncertain about the number and composition of future booking requests. This is the aspect of revenue management that is addressed in most treatments of the subject, including ROTHSTEIN,<sup>[7]</sup> Belobaba<sup>[1]</sup> and others. The second important characteristic of the problem is the interconnectedness of the legs in the system. The airline must make the decision to accept or reject a booking request for a complex itinerary accounting for the value of the passengers that might be displaced on each leg of that itinerary. This is the aspect of revenue management on which the CIF approach sheds the most light.

To understand the effect of revenue management on itinerary flows, we need to introduce the concept of a *booking class*. For each itinerary, an airline may sell tickets to passengers in different market segments at different fares. Each of these different

fares corresponds to a separate booking class. Let  $s$  be an index over booking classes. Then, associated with an itinerary  $I_{ijk}$  will be one or more booking classes, each with a different associated fare. When booking requests are received for a leg, they have associated with them the corresponding booking class and possibly other information. The general revenue management problem is to determine which booking requests to accept and which to reject based on the information available.

The ability of an airline to perform effective revenue management is often constrained by the capabilities of the computerized reservation system (CRS) that communicates the booking requests to the airline. All CRSs transmit the associated booking class with a booking request. If the CRS communicates full information about the desired itinerary as well, then the airline has all the important information necessary to determine the revenue associated with that request. However, currently most CRSs do not transmit full itinerary information with each booking request. Rather, they transmit a separate request for each leg in the itinerary. This means that the airline must make a decision whether to accept a booking request on the

basis of a leg/booking-class combination, with no indication whether or not the leg is part of a more complex itinerary.

When bookings can only be controlled by leg and booking class, the airline must choose a set of booking limits, one for each booking-class and leg combination. Passengers then begin to book. A booking request on a particular itinerary in a particular booking class is accepted as long as the booking limit for that class has not been exceeded on any of the legs in the itinerary. Otherwise it is rejected. The *deterministic booking-class revenue management problem* (DBRM) is to determine the set of booking limits by leg that will maximize total revenue. To formulate this problem, we need to know how many passengers will fly on each itinerary/booking-class combination given a set of underlying demands and a set of booking limits by leg. Answering this question requires application of the CIF algorithm (or an equivalent approach).

Let  $N$  be a leg network with associated unconstrained demands  $d_{ijk}$ , booking limits  $b_{ms}$ , fares  $p_{ijk}$ , and capacity by leg  $c_m$ . For any set of booking limits, the CIF algorithm can be used to find the corresponding set of loads treating itinerary/booking-class combination as a separate itinerary. For purposes of applying CIF, the leg capacities are given by the corresponding booking limits. Let  $\underline{b}$  be the vector of booking limits and  $\underline{d}$  the vector of unconstrained demands. Then we can denote the resulting mapping by  $l_{ijk}(\underline{d}, \underline{b}, N)$ , where  $l_{ijk}(\underline{d}, \underline{b}, N)$  is the number of passengers on itinerary  $i_{jk}$  in booking class  $s$  who are accommodated when the set of booking limits is  $\underline{b}$ . Then, the deterministic booking-class revenue management problem can be written as:

$$\text{Maximize}_{\underline{b}} \sum_i \sum_j \sum_k \sum_s p_{ijk} l_{ijk}(\underline{d}, \underline{b}, N) \quad (5.0)$$

Subject to

$$\sum_s b_{ms} \leq c_m \quad \text{for all } m \quad (5.1)$$

$$b_{ms} \geq 0 \quad \text{for all } m, s. \quad (5.2)$$

It is not clear that this problem can be efficiently solved. As discussed in Section 2.4, total revenue is not concave in  $\underline{b}$ . Therefore, techniques that rely upon local information to determine a direction of improvement may not find a global optimum. Indeed, the problem is potentially made even more complex by allowing different booking rates among itinerary/booking-class combinations and the possibility of reflow among booking classes when booking limits are binding.

Ideally, of course, an airline would like to be able

to control bookings in terms of itinerary and booking class. We call this problem the *deterministic itinerary revenue management problem* (DIRM). For each itinerary/booking-class combination, let  $b_{ijk}$  be the associated booking limit. Then DIRM can be written as:

$$\text{Maximize}_{b_{ijk}} \sum_i \sum_j \sum_k \sum_s p_{ijk} b_{ijk} \quad (6.0)$$

Subject to

$$\sum_i \sum_j \sum_k \sum_s \delta_{ijkm} b_{ijk} \leq c_m \quad \text{for all } m \quad (6.1)$$

$$b_{ijk} \leq d_{ijk} \quad \text{for all } i, j, k, s \quad (6.2)$$

$$b_{ijk} \geq 0 \quad \text{for all } i, j, k, s. \quad (6.3)$$

A network approach to solving DIRM can be found in Glover et al.<sup>[4]</sup>

Clearly, for the same set of underlying demands, the maximum revenue obtainable from setting booking limits by itinerary/booking-class will be at least as great as that obtainable from setting booking limits by leg/booking-class. For systems with considerable connecting traffic, it may be considerably greater. In other words, the revenue from solving DIRM will be greater than the revenue from solving DBRM. This superiority of itinerary/booking-class revenue management to leg/booking-class revenue management holds equally true when uncertainty is taken into account. This is a strong motivation for airlines to try to manage bookings by itinerary/booking-class rather than by leg/booking-class. The abilities of airlines to develop such systems will be limited by whether or not the majority of their booking requests are communicated on an itinerary of a leg basis.

From the above discussion, it is clear that the type and sophistication of revenue management performed by a carrier will have a significant effect on the size and composition of its loads. With no revenue management or revenue management by booking class and leg only, a CIF-type approach is still required to account for the complex interactions of passenger flows. However, a carrier with itinerary/booking class revenue management capabilities is able to directly control the mix of passengers (at least in the deterministic case.) No CIF-type approach is required to determine loads or itinerary flows. This difference extends to the concavity of total revenue with respect to capacity by leg. With no revenue management or revenue management by booking class and leg only, total passenger revenue is **not** guaranteed to be a concave function of seating capacity by leg. With revenue management by itinerary and booking class, total passenger

revenue will be concave with respect to the seating capacity on each leg. (This is a direct consequence of the fact that DIRM is a linear program.) The concavity of total passenger revenue with respect to seating capacity on each leg may have implications for the ability of classical optimization techniques to determine the optimal allocation of aircraft with different seating capacities to a schedule.

## 6. SUMMARY AND DISCUSSION

THE IMPORTANCE of connecting traffic to the profitability and operation of modern airlines is universally recognized. However, the cross-leg revenue and load effects resulting from such connecting traffic are difficult to incorporate into formal optimization approaches to airline planning problems such as aircraft assignment and yield management without making the problems unmanageably complex. For this reason, most approaches to these types of problems make the assumption (often unstated) that loads, yields, and revenues on a particular leg are independent of the seating capacities offered on other legs. As the examples we presented show, this assumption does not correspond to the situation in the real world. The next question is then: How important are such cross leg effects? If they are not widespread, then they can probably be ignored without significant distortion. If, on the other hand, they are widespread, then they need to be incorporated for optimization approaches to provide good solutions.

In this paper we presented a straightforward algorithm for determining loads, yields, and revenues given an assignment of seating capacities to legs and an underlying set of itinerary demands. Although the basic algorithm assumed uniform booking rates across all legs and no reflow, it can be easily modified to accommodate the case of variable booking rates and reflow. In application to a realistic schedule with realistic capacity assignments and loads, we found that calculating loads and revenues under the standard assumption of no cross-leg overestimates system revenue by about 0.3% and system load factor by about 0.4%. The results suggest that while cross-leg effects may not be important for many legs within a schedule, they can be of critical importance for certain legs. These legs are those that have a high level of unconstrained demand and a high fraction of connecting passengers. For a large airline, legs with these characteristics would tend to be those linking their major hubs. For these legs, taking into account the cross-leg effects is crucial in making good capacity allocation decisions.

The application described in this paper was chosen to be realistic in the sense that the underlying schedule was based on the "hub-and-spoke" philosophy prevalent in the industry. Although the results for this application are suggestive, they cannot, of course, be considered definitive. The sensitivity to underlying demand indicates that cross-leg effects can be significant at a system level (> 1% of revenue) at high levels of overall demand. We would expect that they have the most importance in planning for airlines with high load factors and/or high dependence on connecting traffic. Since the trend of recent years has been toward higher load factors and greater reliance on connecting traffic, we expect that the importance of such cross-leg effects will increase over time.

Although the CIF method presented here is a better estimator of on-board loads, yields, and revenues than the PF method, several areas remain for further investigation. Perhaps the most salient area is that of uncertainty. Although the PF and CIF methods may seem reasonable for application to a single day when the demand is known with certainty, they are less applicable to longer time periods in which the underlying demands are both uncertain and fluctuate from day to day. Such demand fluctuation has been incorporated into the PF approach by use of the so-called *rejected demand* model (see SWAN<sup>(8)</sup>). The rejected demand model treats the unconstrained demand for a leg as the mean of a distribution that reflects both underlying uncertainty on demand and its natural day-to-day variation. Under these assumptions, given a fixed seating capacity, the mean load will always be less than the unconstrained demand (since, with some probability, unconstrained demand will be greater than the capacity implying that the load will be equal to the capacity). In addition, the mean load will generally be less than the capacity, even if the mean unconstrained demand is greater than the capacity. This probabilistic model has been widely applied in airline planning applications. The equivalent probabilistic formulation of the CIF approach is not immediately clear. This is particularly true in light of the strong probabilistic dependencies among market demands. Formulation of a probabilistic version is a clear area for further research.

Although the development of the Consistent Itinerary Flows Algorithm in the previous sections was couched in terms of the passenger airlines, it is clearly directly applicable to any scheduled transportation service in which connecting traffic plays a significant part. Some examples include air freight, passenger busses, and passenger railways. In these industries the straightforward CIF algo-

rithm might even be more effective than in the airlines, since revenue management systems are currently more primitive or nonexistent.

An interesting further application of the Consistent Itinerary Flows approach is to rentals. Typically, in the short run a renter owns a fixed stock of capacity (cars, hotels rooms, etc.) that can be rented for different lengths of time starting on different days. The revenue per day charged for the rental may depend both on the length of the rental and/or the starting date. For example, a hotel may offer a standard daily rate, a weekly rate, a monthly rate and a weekend rate. In this case, seeking to determine occupancy, marginal revenue, and total revenue given a forecast of demands for bookings in each of the rates is analogous to determining loads, marginal revenue, and total revenue for an airline. In the case of the hotel, each combination of arrival day and length of stay corresponds to an "itinerary" and each calendar day corresponds to a "leg." The capacity on each "leg" is the number of available rooms. With these correspondences, the problem of determining occupancy and total revenue for the hotel is essentially identical to the Consistent Itinerary Flow problem faced by the passenger airline and is amenable to the same algorithm. The only difference is that hotel bookings might need to be considered for an indefinite period into the future, corresponding to an infinite number of legs and itineraries. In practice, this difficulty can be avoided by finding some future period when the forecast unconstrained occupancy is less than capacity for a number of days in a row. The CIF algorithm can then be applied for all days up to that period assuming that no potential bookings will

be blocked during or after that period. A similar approach could be used for cars or other rental equipment.

#### ACKNOWLEDGMENTS

THE COMMENTS and suggestions of the Editor and an Associate Editor resulted in a number of improvements in clarity and readability of this paper.

#### REFERENCES

1. P. P. BELOBABA, "Airline Yield Management: An Overview of Seat Inventory Control," *Trans. Sci.* **21**, 63-67 (1987).
2. G. B. DANTZIG, *Linear Programming and Extensions*, Princeton University Press, Princeton, N.J., 1963.
3. M. S. DASKIN AND N. D. PANAYOTOPOULOUS, "A Lagrangian Relaxation Approach to Assigning Aircraft to Routes in Hub and Spoke Networks," Research Report, the Transportation Center, Northwestern University, 1987.
4. F. GLOVER, R. GLOVER, J. LORENZO AND C. McMILLAN, "The Passenger-Mix Problem in the Scheduled Airlines," *Interfaces* **12**, 73-80 (1982).
5. J. P. LEGENDRE AND M. MINOUX, "Une Application de la Notion de Dualité en Nombres Entiers: Sélection et Affectation Optimales d'une Flotte d'Avions," *RAIRO Rech. Oper.* **11**, 201-222 (1977).
6. D. C. LUENBERGER, *Introduction to Linear and Non-linear Programming*, Addison-Wesley, Reading, Mass., 1973.
7. M. ROTHSTEIN, "OR and the Airline Overbooking Problem," *Opns. Res.* **33**, 237-284 (1985).
8. W. SWAN, "Load Factor and Denied Boardings," FTL Memorandum M78-1, MIT Flight Transportation Laboratory, Cambridge, Mass., 1978.

(Received, June 1988; revisions received September 1989, May 1990; accepted May 1990)

Copyright 1991, by INFORMS, all rights reserved. Copyright of Transportation Science is the property of INFORMS: Institute for Operations Research and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.