STOCHASTIC PROGRAMMING MODELS FOR SCHEDULING AIRLIFT OPERATIONS*

J. L. Midler and R. D. Wollmer

The RAND Corporation
Santa Monica, California 90406

ABSTRACT

This article describes an analytic approach to flight scheduling within an airlift system. The model takes explicit account of the uncertainty present in cargo requirements or demands. For computational feasibility, the approach consists of two related models: (1) a monthly planning model that produces an initial schedule, and (2) a daily model for making periodic changes in the schedule. Both are formulated as two-stage stochastic linear programs. A detailed mathematical description of each model and its physical interpretation is given.

The monthly model determines the number of flights each type of aircraft in the fleet. Excess demands on certain routes are assumed to be met, at least in part, by spot procurement of commerical lift from outside the system. The flight assignment is determined by minimizing the expected total system cost, which consists of operating costs, costs of reallocating aircraft to different routes, spot commercial procurement costs, and other penalty costs of excess demand. The model accounts for limitations on the number of flying hours and the carrying capacities of various aircraft in satisfying demands.

In the daily model the number of aircraft of each type to switch from one route to another and the number of commercial flights on spot contract to add on the current day are the principal decision variables. These are determined by balancing operating, procurement, and redistribution costs against the expected costs of additional cargo delay. The current state of the system—the amount of unmoved cargo on various routes andathe position of aircraft throughout the system—plays a role in determining these decisions.

A description of two variants of an algorithm recently developed for this class of problems is presented. Both versions, which use ideas from convex programming, make extensive use of linear programming codes for the brunt of the calculations. The models may thus be solved by augmenting existing linear programming routines.

I. INTRODUCTION

This paper provides the technical details of a mathematical model for scheduling flights in a large airlift system. The model takes explicit account of the uncertainty surrounding the requirements at the time scheduling decisions must be made, as well as of the various physical constraints on the system. The proposed models are adaptive in the sense that current information about the state of the system (cargo levels at various origins, distribution of vehicles throughout the system) is used periodically in making decisions regarding changes in the flight schedule.

The approach discussed here consists of two components—a monthly and a daily flight planning model. Both are formulated as stochastic two-stage linear programs, and can be solved by algorithms based upon standard linear programming calculations. Consequently, the model can be applied to large-scale systems. This should permit scheduling the system as a whole rather than the two divisions separately, with transfers of aircraft between then on an ad hoc basis, as is currently the practice.

*This research is supported by the United States Air Force under Project RAND. Views or conclusions contained in this study should not be interpreted as representing the official opinion or policy of the United States Air Force.

The paper is organized as follows. The next section contains a brief description of the scheduling problem and discusses the overall strategy employed in treating it analytically. This is followed in Sec. III by a mathematical statement of the monthly and daily models, which are similar in general structure. Section IV contains a discussion of two variants of an algorithm for solving these models, preceded by a short discussion of pertinent concepts for two-stage linear programming under uncertainty. In Sec. V we present a small numerical example for the monthly model.

II. PROBLEM STATEMENT

Description of the Airlift System

Airlift missions within the system are of two basic types: channel missions and special missions. Channel missions fly on a regular basis between fixed origins and destinations: Fixed airports in the continental United States, aerial ports of embarkation (APOEs) and overseas bases, and aerial ports of debarkation (APODs). They carry cargo and passengers.* They are scheduled in advance for each calendar month, although the schedules (number of flights and departure times) are changed during the month, often daily, based upon deviations from expected cargo generation. Many of these flights make intermediate stops for refueling and crew changes. Channel cargo is shipped into APOEs for storage until shipped.

Special missions are flights from points of original cargo generation, such as factories or warehouses, to points at or near the ultimate consignee. Such flights arise on an irregular basis, and are scheduled if sufficient advance notice is provided. The time and place of cargo pickup is usually specified. Often these flights arise after the beginning of the month and require a temporary diverting of aircraft from channel missions, since certain special missions are designated higher priority than channel traffic. The users of a special mission must, in effect, charter an entire aircraft, although the airlift operator decides what type of aircraft will be supplied.

The system is divided administratively into two operating divisions. One division comprises APOEs on the West Coast and serves points in Asia. The other comprises APOEs on the East Coast and serves bases in Europe, the Middle East, and South America, although currently several channels serve destinations in Asia. Headquarters assigns special missions to the two divisions, usually on the basis of whether the cargo originates west or east of the Mississippi River. Often aircraft belonging to one division are temporarily assigned to the other division to meet these demands when the type of aircraft required is in short supply.

Scheduling Problem

In carrying out its operations, the operator is faced with random fluctuations in the amount of cargo to be carried over its various routes.

There is monthly uncertainty because shippers' forecasts of requirements for the month as reported to the airlift operator at various lead times differ significantly from the actual amount of requirements shipped. Secondly, the amount of cargo actually received during a day is partly unpredictable due to random variability in production schedules for manufactured items, delays in transit and depot processing, and other factors.

---

*Due to the large size of the system only cargo flights are considered in the models to be developed. (Also most of the passengers are carried by commercial flights.) Passengers can be handled in the models developed below by converting them to ton (or cube) equivalents if flights can carry mixed loads. Otherwise demand would have to be disaggregated into separate commodity classes significantly increasing the size of the models.

A scheduling problem arises, since demand is not known with certainty and aircraft differ with respect to their payloads, flying times, and operating costs over different routes. Therefore, aircraft of different types must be assigned to various routes over the monthly planning period. In addition, the system can obtain additional lift from commercial sources on call contracts to handle excessive demands on some routes. Furthermore, the assignments of aircraft cannot violate constraints for flying hours.

As time unfolds and demand becomes known, the scheduler can transfer those aircraft from routes where demand is less than anticipated to those where it is higher than expected. Thus, the scheduling problem is a dynamic one of periodically reassigning capacity to routes in order to minimize the expected costs of the operation over time. Assigning insufficient airlift increases both the costs due to delay and the procurement costs of "emergency" commercial lift; assigning excess capacity results in an opportunity cost of foregoing service elsewhere in the system. Since both situations are stochastic, the best that can be done is to minimize the cost entailed in these situations on the average.

Basic Approach

We turn now to an overview of our strategy in treating the scheduling problem. Ideally, the problem should be treated as a multiple period (30 days or longer), sequential decision problem with account being taken of daily demand variation and physical constraints on aircraft in each of the ensuing periods in the current horizon. Such an approach is not feasible from a computational standpoint for a system the size of the one considered or even each division separately. Therefore, broadly following the current procedures, we break the process into two parts—monthly and daily.

A monthly schedule is determined that accounts for uncertainty for the forthcoming month considered as one aggregate period in which decisions planned in advance may be modified. Although this is far from ideal, it is superior to ignoring uncertainty and the subsequent opportunity to remedy its effects. The second-state decisions, as we shall see, are "virtual" or fictitious, but are nonetheless important because they help provide a good first-stage decision, the initial assignments that serve as the starting point for the daily modifications.

The monthly flights are apportioned to the days. For example, if 15 flights of a given aircraft are found from the monthly model, this may be interpreted as one flight per aircraft every two days during the month if the given aircraft type can fly the mission.

Once the month commences, daily amendments to the number of flights scheduled for that particular day are determined by the daily model, which is also a two-stage optimization. Changes in assignments optimal for the forthcoming 2-day period are determined. They minimize the sum of costs incurred on the current day, given the amount of cargo on hand plus the expected value of tomorrow's costs, which depend upon both the amount of new cargo and the changes made on the first day. Consecutive decisions of the daily model are thereby chained together in a recursive fashion with outputs from the model for day t becoming inputs to the model for day $t+1$.

III. MATHEMATICAL FORMULATION

In this section we formulate the monthly and daily models as two-stage linear programs under uncertainty [2], [10] with a random right-hand side in the second stage. Several of the program's variables represent numbers of different kinds of flights and consequently must be integer valued. However, this aspect of the problem is neglected in order to obtain a program that is computationally solvable, thus leaving the user to adopt a rounding scheme. The general form for problems of this type is the following:

$$\min_{x\geq 0}\left\{cx+E_\xi[\min_{y\geq 0} qy|Wy=\xi-Tx]\right\} \qquad \text{subject to } Ax=b. \tag{1}$$

In this problem, c, q, and b are known vectors of dimension n_1, n_2, and m_1, respectively. A, T, and W are known matrices of dimension $m_1\times n_1$, $m_2\times n_1$, and $m_2\times n_2$, respectively. The vectors x and y represent the first- and second-stage decision vectors. ξ is a random vector with a known probability distribution and $E(\cdot)$ denotes the expectation with respect to that distribution.

The problem can be given the following interpretation. Before the random variable ξ is observed, one must choose nonnegative values of x satisfying $Ax=b$. The immediate cost of this choice is cx. After x is chosen and the random variable ξ is observed, nonnegative values must be chosen for y which satisfy $Wy=\xi-Tx$ at a cost of qy. The chosen value of y will, for each observed value of ξ, be the one which minimizes qy. Letting $Q(x)=E_\xi\{\min_{y\geq 0} qy|Wy=\xi-Tx\}$, we wish to minimize the first-stage cost, cx, plus the expected second-stage cost, $Q(x)$. The problem can be viewed as a dynamic programming problem with only two stages, with y a function of the state vector $\xi-Tx$. Note that x does not depend on the observed values of ξ but only on the distribution of ξ, while y does depend on the observed values of ξ and on the chosen values of x.

Monthly Model

In the monthly model the sequence of decisions is the following:

1. The number of flights of each aircraft type in the MAC fleet over each route* is assigned before monthly requirements** on each route are known with certainty.

2. After monthly requirements are observed, some flights assigned to routes with lower than expected demand are switched to routes with higher than expected demand. Commercial airlift is also sometimes added to routes with excess demand.

The monthly model has the following structure. The first-stage constraints state that for each aircraft type the total number of flying hours allocated to all routes cannot exceed the total number of flying hours available of that type. The second-stage constraints are of two types. The first specifies that the number of flying hours of a given aircraft type diverted from a particular route to other routes cannot exceed those initially assigned to it. The second type are demand balance equations, which state that for each route the total carrying capacity (that originally assigned, minus that diverted to other routes plus that diverted from other routes) minus unused carrying capacity plus unsatisfied demand is equal to total demand for that route.

The objective function consists of the cost of the final flying program (the initial plus the amended assignments) plus penalty costs of excess demand or supply. The cost of excess demand is reflected in both the cost of additional commercial lift plus the extra flying time consumed in switching aircraft from one route to another. Specifically, the program is as follows.

Find min Z, x_{ij}, x_{ijk}, y_j^+, $y_j^-\geq 0$ such that

*A route is a channel or special mission from each origin. Routes are assumed to be point-to-point with no intermediate on-load or off-load of cargo; however, e.g., a trip from San Francisco to Japan via Hawaii is considered a different route from a flight which goes via Alaska.

**Requirements as used here will mean net requirements from which capacity of commercial airlift on an annual contract (for which the frequency of trips and routes is fixed) and any other fixed capability that is available is subtracted. This has the effect of merely lowering the mean of the distribution of requirements.

$$\sum_{i,j} c_{ij}x_{ij}+E\left\{\min\left[\sum_{i,j,k\neq j}\left(c_{ijk}-c_{ij}\frac{a_{ijk}}{a_{ij}}\right)x_{ijk}+\sum_j c_j^+y_j^+ +\sum_j c_j^-y_j^-\right]\right\}=Z. \tag{2}$$

First stage:

$$\sum_j a_{ij}x_{ij}\leqslant F_i, \qquad \text{all } i. \tag{2.1}$$

Second stage:

$$x_{ij}-\sum_{k\neq j}\frac{a_{ijk}}{a_{ij}}x_{ijk}\geqslant 0, \qquad \text{all } i, j \tag{2.2}$$

$$\sum b_{ij}x_{ij}-\sum_{i,k\neq j}\left[b_{ij}\left(\frac{a_{ijk}}{a_{ij}}\right)\right]x_{ijk}+\sum_{i,k\neq j} b_{ij}x_{ijk}-y_j^+-y_j^-=d_j, \qquad \text{all } j, \tag{2.3}$$

where x_{ij}= number of flights of aircraft type i initially assigned to route j during the month;

x_{ijk}= number of flights of aircraft type i assigned to route k using hours made available by canceling route j flights;

y_j^+ = demand on route j which is satisfied by commercial lift, if permitted, or unsatisfied demand if commercial lift not permitted;

y_j^- = unused capacity on route j;

and

a_{ij}= number of hours required by aircraft type i for a flight initially assigned to and flown on route j;

a_{ijk}= number of hours required by aircraft type i for a flight on route k that uses hours made available by canceling route j flights ($a_{ijk}\geqslant a_{ik}$);

b_{ij}= the carrying capacity (tons or any other appropriate measure) of a flight of an aircraft of type i on route j;

F_i= the maximum number of flying hours for aircraft of type i available during the month;

d_j= total demand (tons or any other measure) for route j (a random variable);

c_{ij}= cost of flight of aircraft type i initially assigned to and flown on route j ($c_{ij}\geqslant 0$);

c_{ijk}= cost per flight of aircraft type i assigned to route k from hours made available by canceling route j flights ($c_{ijk}\geqslant c_{ik}$);

c_j^+ = cost per ton of commercial augmentation on route j. If commercial augmentation is not available to carry excess demand, this may instead represent a shortage cost;

c_j^- = cost of a unit of unused carrying capacity on route j.*

Note that since a flight assigned to and flown on route j takes a_{ij} hours while a flight flown on route k from hours diverted from route j takes a_{ijk} hours, such a flight on route k results in the cancellation of (a_{ijk}/a_{ij}) flights on route j. Thus, of the flights initially assigned to route j, those that are actually flown number

$$x_{ij}-\sum_{k\neq j}\frac{a_{ijk}}{a_{ij}}x_{ijk}.$$

This is reflected in expressions (2), (2.2), and (2.3).

*This model and the daily model described below represent generalizations of the pioneering Ferguson-Dantzig model [4] applying linear programming under uncertainty to aircraft allocation. In contrast to their model, the monthly model considers reassignment of aircraft once demand is known and hence its second stage consists of a bonafide linear program. Also the demand distributions in our models are continuous rather than discrete.

The first-stage constraints correspond to $Ax=b$ and the second-stage constraints to $Wy=\xi-Tx$ in expression (1). Notice that the constraints (2.2) are placed in the second stage despite the fact that the right-hand side is not a random vector. This is done because the switching variables x_{ijk} are assumed to be determined *after* demand becomes known and thus they depend upon information about the d_j. Also, they are to be chosen simultaneously with the y_j^+ and y_j^- to minimize the overall second-stage cost function.

Daily Scheduling Model

The daily scheduling model is also a two-stage linear program under uncertainty. Its principal first-stage decision variables are the number of flights of each type to switch from one route to another in the forthcoming day—variables that played the role of second-stage recourse variables in the monthly model. In addition, the daily model is couched in terms of unmoved cargo stocked at the various origins (aerial ports of embarkation) for various lengths of time; the cargo has not yet been shipped, and subsequently incurs certain delay costs. Consequently, the amount of cargo in storage for various lengths of time waiting to be loaded on different aircraft types over the different routes constitutes another set of first-stage decision variables.

In the first stage the constraints consist of balance limitations on the number of flying hours, bounds on switches, capacity restrictions on shipments, and inventory balance equations for cargo which has been in the port for over 1 day—quantities known with certainty. The second stage consists of demand equations for newly arriving cargo which is treated as a random variable. These equations include terms to account for a possible under- or over-supply of capacity. It will be assumed that for a given route new cargo has a lower priority (i.e., lower penalty or delay cost) than old cargo. The reason for this will become clear later.

The objective function is again one of cost minimization. In this case the costs include (1) incremental operating costs resulting from flight changes, (2) daily commercial procurement, (3) delay costs for cargo that remains unmoved for various lengths of time, and (4) expected penalty costs resulting from misallocation of capacity.

To write the daily model, we introduce the following notation (variables refer to the t-th day of the month):

x_{ij} = number of flights of aircraft type i previously scheduled to fly over route j, a known constant;

x_{ijk} = number of flights of aircraft type i switched to route k using hours made available by canceling route j flights;

y_{ij}^l = capacity reserved for cargo l days old shipped on aircraft type i over route j $(l=0, 1, \ldots, L)$;

w_{hj}^l = capacity reserved for cargo l days old shipped on commercial lift of type h over route j;

z_{hj} = number of flights of commercial lift of type h procured for route j;

s_j^l = amount of cargo l days old remaining to be shipped over route j at the end of the day;

H_j^l = inventory of cargo l days old for route j for shipment at the beginning of the day (H_j^0 represents new cargo and is a random variable);

v_j = unused capacity on route j at the end of the day;

a_{ij} = number of hours required by aircraft type i scheduled and flown on route j;

a_{ijk} = number of hours required by aircraft type i for a flight on route k using hours made available by canceling route j flights $(a_{ijk} \geqslant a_{ik})$;

b_{ij} = the carrying capacity of aircraft type i over route j;

c_{ij} = cost per flight of aircraft type i scheduled and flown on route j $(c_{ij} \geqslant 0)$;

c_{ijk} = cost per flight of aircraft type i assigned to route k using hours made available by canceling route j flights $(c_{ijk} \geq c_{ij})$;

β_{hj} = cost per flight of commercial lift of type h over route j;

γ_j^l = unit cost per ton of one additional day's delay for cargo l days old which is to be shipped over route j. It is assumed that $\gamma_j^{\circ} = \min_l \gamma_j^l$.

Then the problem for day t can be stated. Find

$$\text{Min } Z,\ x_{ijk},\ s_j^l,\ z_{hj},\ y_{ij}^l,\ w_{hj}^l,\ v_j \geq 0,$$

such that

$$\sum_{i,j,k\neq j}\left(c_{ijk} - c_{ij}\left(\frac{a_{ijk}}{a_{ij}}\right)\right)x_{ijk} + \sum_{j,l\neq 0}\gamma_j^l s_j^l + \sum_{h,j}\beta_{hj}z_{hj} + E\left\{\min_{s_j^{\circ}}\sum_j \gamma_j^{\circ}s_j^{\circ}\right\} = Z. \tag{3}$$

First Stage:

$$\sum_{k\neq j}\frac{a_{ijk}}{a_{ij}}x_{ijk} \leq x_{ij},\ \text{all } i \text{ and } j; \tag{3.1}$$

$$\sum_{l>0} y_{ij}^l + y_{ij}^{\circ} + \sum_{k\neq j}\left(b_{ij}\frac{a_{ijk}}{a_{ij}}\right)x_{ijk} - \sum_{k\neq j} b_{ij}x_{ikj} \leq b_{ij}x_{ij},\ \text{all } i \text{ and } j; \tag{3.2}$$

$$\sum_{l>0} w_{hj}^l + w_{hj}^{\circ} - b_{hj}z_{hj} \leq 0,\ \text{all } h \text{ and } j; \tag{3.3}$$

$$\sum_i y_{ij}^l + \sum_h w_{hj}^l + s_j^l = H_j^l,\ \text{all } j \text{ and } l > 0. \tag{3.4}$$

Second Stage:

$$s_j^{\circ} - v_j = H_j^{\circ} - \sum_i y_{ij}^{\circ} - \sum_h w_{hj}^{\circ},\ \text{all } j. \tag{3.5}$$

Equation (3.1) states, as in the monthly model, that the number of flying hours switched off a given route and aircraft type cannot exceed those originally assigned.

The meaning of Eqs. (3.2) and (3.3) is that the total tonnage loaded of cargo already received plus the amount reserved for new cargo on a given aircraft type over a given route must be less than the carrying capacity of all aircraft of that type assigned to the route and for the commercial vehicles, respectively.* Equation (3.4) states for each duration of cargo destined for each route, the total tonnage allocated for the aircraft and commercial airlift plus the amount not moved must equal the inventory of cargo at the beginning of the day (a known constant for all but newly arriving cargo). We note that the assumption that newly arrived cargo has lowest penalty cost (mathematically stated, $\gamma_j^{\circ} = \min_l (\gamma_j^l)$ for all j) assures that no new cargo will be carried unless all old cargo is carried. This assures that for all but newly arrived cargo, the amount shipped uses all the allotted capacity. If this assumption were not made, it would be more reasonable to determine the y_{ij}° and w_{ij}° simultaneously with the y_{ij}^l and $w_{ij}^l (l > 0)$ as a function of the actual amount of new cargo. This would necessitate shifting equations (3.2) through (3.4) into the second stage, making for a large linear program for the second stage. The present formulation permits the second stage to have a very simple structure which reduces the computational burden in solving it. This matter will be discussed in more detail below.

*For routes on which commercial augmentation is not permitted, the corresponding z_{hj} are omitted.

Finally, Eq. (3.5) states that for each route the inflow of new cargo minus the total capacity reserved for shipment of new cargo on all types of operator-owned and commercial vehicles is equal to any cargo not shipped minus any unused capacity.

In addition, we may wish to include constraints that take into consideration the capacity limitations of enroute bases. Let $J(p)$ equal the set of routes which use base p as an enroute stop for refueling or crew change. Let α_p be the maximum number of flights that can be accommodated at base p in a single day. Then these constraints can be expressed as

$$\sum_{i,\, j\epsilon J(p)} \left(x_{ij} - \sum_{k \neq j} \frac{a_{ijk}}{a_{ij}} x_{ijk} + \sum_{k \neq j} x_{ikj} \right) \leq \alpha_p.$$

Letting $H_j^l(t)$ be the inventory of cargo for route j that is l days old on day t, and letting quantities in parentheses show time dependence of other quantities in a similar way, note that $H_j^l(t+1) = s_j^{l-1}(t)$, using the convention that $s^{l-1}(0)$ refers to cargo not shipped on the last day of the preceding month.

In addition, the quantities x_{ij} are updated daily by the relations

$$x_{ij}(t+1) = x_{ij}(t) - \sum_{k \neq j} \frac{a_{ijk}}{a_{ij}} x_{ijk} + \sum_{k \neq j} x_{ikj},$$

where the x_{ijk} and x_{ikj} are the optimal values of the variables as determined by the model for day t.

In the monthly model there are m (number of individual aircraft types) constraints in the first stage and $(mn+n)$ in the second stage (n being the total number of routes). The assumption that there are five aircraft types and 100 routes presents a linear program for the first stage with five constraints and 600 for the second stage; however, in the second stage the first mn are of the generalized upper-bound type, to be discussed later, to which special methods may be applied.

In the daily model there are at most $n(2m+H+L)$ constraints in the first stage and n in the second stage, where H is the number of individual commercial vehicles and L is the number of duration classes of old cargo. For $H=3$, $L=3$, $m=5$, and $n=100$ this yields one first-stage program with 1600 constraints. The second stage contains 100 equations. While this is within the range of linear programming codes for "third generation" computers, the first-stage problem can be simplified by treating it by decomposition methods where Eq. (3.2) is the coupling equation and (3.1), (3.3) and (3.4) are the two disjoint subproblems. Again, for the former subproblem, the majority of the constraints (3.1) are of the generalized upper-bound variety.

IV. SOLUTION TECHNIQUES

This section describes two alternative computational methods for solving two-stage linear programs under uncertainty. These methods may be applied to the monthly and daily model. They assume the random variables have a continuous distribution. Each is a variant of an algorithm for convex programming, which is required for this type of problem. The first is a cutting-plane or constraint-generating procedure, while the second is a gradient method. We outline the steps required for the general problem exemplified by the monthly model and indicate certain simplifications that can be made in the daily model whose second-stage has a special structure. Before proceeding to the algorithms, we review some salient facts concerning two-stage linear programming under uncertainty.

Basic Facts from Linear Programming Under Uncertainty

As mentioned in Sec. III, both the monthly and daily problems are special cases of the problem:

(4)
$$\text{Find } x \geq 0, \text{ min } Z \text{ such that}$$
$$Z = cx + Q(x)$$
$$Ax = b$$
$$\text{where } Q(x) = E\left[\min_{y \geq 0} qy | Wy = \xi - Tx\right].$$

Wets [9] has shown that $Q(x)$ is convex in x. Thus the above linear program under uncertainty is equivalent to a deterministic convex program in x. The algorithms to be developed focus and are based on certain properties of the function $Q(x)$.

One important property (to be proved later in the Constraint Generation section) is that the linear function $\rho(x°) - (\pi(x°)T)x$ is a support function* (or tangent plane) to $Q(x)$ at $x°\epsilon K$, where K is $\{x|Ax = b\}$ and $\pi(x°)$ is the expectation of the optimal dual variables and $\rho(x°)$ is the expectation of the demand terms in the dual objective function. Specifically, let $\pi(x°, \xi)$ be an optimal solution to

(5)
$$\max_{\pi} \left\{\pi(\xi - Tx°) | \pi W \leq q\right\}.$$

Then

$$\pi(x°) = E_\xi \pi(x°, \xi) \text{ and } \rho(x°) = E_\xi\{\pi(x°, \xi)\xi\}.$$

This property forms the basis for the cutting-plane method.

Another important property [9] is that if the distribution of ξ is absolutely continuous, $Q(x)$ is differentiable at $x°\epsilon K$ and has the gradient $-\pi(x°)T$, where again $\pi(x°)$ is the expected value of the solution to (5). This property is utilized in the gradient approach developed by Wets [9].

In general, the $\pi(x°)$ and $\rho(x°)$ cannot be explicitly computed; consequently, the methods to be described resort to sampling procedures to estimate these quantities.

Our monthly and daily problems both possess the property of *complete recourse*. In other words, each x satisfying $Ax = b$ yields a feasible program for all possible ξ in the second-stage constraints $Wy = \xi - Tx$. This permits us to eliminate one of the steps in the algorithms to be described for finding feasibility cuts [8, pp. 31–34]. Moreover, the daily model exhibits simple recourse, i.e., $W = [I, -I]$. This permits simplification of one major step in the following algorithm.

Cutting-Plane Method

A cutting-plane algorithm has been developed by Wets and Van Slyke [8] for solving the general problem. It can be viewed as an application of Kelley's cutting-plane method [6] to the equivalent convex program.

In this approach the convex function $Q(x)$ appearing in the objective function is treated as a constraint. The objective function is replaced by $z(x) = cx + \theta$ and the constraint,

(6)
$$Q(x) \leq \theta,$$

considered in the first-stage program where θ is treated as an additional decision variable. Since the second-stage costs are nonnegative in our problem, we may consider θ as a nonnegative variable.

*A linear function $H(x)$ is a support for convex function $z(x)$ defined on $X \subseteq E^n$ at $x°$ if $z(x) \geq H(x)$ for all $x\epsilon X$ and $z(x°) = H(x°)$.

The constraint (6) is represented by a set of support planes $\rho^k-(\pi^k T)x$; i.e., a convex function is represented as the upper envelope of its supporting planes. Four constraints are generated by solving the second-stage problem for a fixed x which solves an augmented first-stage problem. Thus the technique consists of a decomposition of the overall problem into two linear-programming subproblems. The method can be viewed, in fact, as a particular case of Bender's partitioning method [1] applied to the original problem.

The particular problem that is solved at any iteration, say the k^{th}, is the following:

1. Find $x,\ \theta \geqslant 0$, min $\tilde{z}$ such that

$$
\begin{aligned}
&cx+\theta=\tilde{z}\\
&\text{a) } Ax=b\\
&\text{b) } (\pi^k T)x+\theta \geqslant \rho^k, \qquad k=1,\ldots,S,
\end{aligned}
\tag{7}
$$

where S is an index of a subset of previously generated constraints to be described more fully below.

The constraints (7b) are generated sequentially and have the property that for all x satisfying (7a), $(x,\ Q(x))=(x,\ \theta)$ satisfies (7b).

2. Let $(\bar{x},\ \bar{\theta})$ be the solution obtained from step 1 and solve for $Q(\bar{x})$. If $Q(\bar{x})=\bar{\theta}$, $\bar{x}$ solves (4). Otherwise, a constraint of the type (7b) is generated that is violated by $(\bar{x},\ \bar{\theta})$.

Constraint Generation – For fixed $x=\bar{x}$ and a given $\xi=\tilde{\xi}$, the second-stage linear program becomes:

$$
\begin{gathered}
\text{Find } y\geqslant 0,\ Q(\bar{x},\ \tilde{\xi}) \text{ where}\\
Q(\bar{x},\ \tilde{\xi}) = \min qy \text{ subject to } Wy=\tilde{\xi}-T\bar{x}.
\end{gathered}
\tag{8}
$$

Let $\pi(\bar{x},\tilde{\xi})$ be the simplex multipliers corresponding to an optimal basis in (8). It follows that $\pi(\bar{x},\ \tilde{\xi})(\tilde{\xi}-T\bar{x})=Q(\bar{x},\ \tilde{\xi})$ and $\pi(\bar{x},\ \tilde{\xi})W\leqslant q$. Thus, $\pi(\bar{x},\ \tilde{\xi})$ is feasible in the following program for all x:

$$
\begin{gathered}
\text{Find } \pi, \max w \text{ where}\\
\pi(\tilde{\xi}-Tx)=w\\
\pi W\leqslant q.
\end{gathered}
\tag{9}
$$

Since (9) is dual to (8) for general x, it follows from duality theory that

$$\pi(\bar{x},\ \tilde{\xi})(\tilde{\xi}-Tx)\leqslant Q(x,\ \tilde{\xi})$$

for all x. Letting $\pi(\bar{x})=E_\xi\pi(\bar{x},\ \xi)$ and $\rho(\bar{x})=E_\xi[\pi(\bar{x},\ \xi)\xi]$, it follows that for all x, $(\pi(\bar{x})T)x+Q(x)\geqslant\rho(\bar{x})$. Hence, the following constraint must be satisfied by all $(x,\ Q(x))$:

$$(\pi(\bar{x})T)x+\theta\geqslant\rho(\bar{x}). \tag{10}$$

Algorithm – All constraints of the type (7b) are of type (10). That is, each $(\pi^k,\ \rho^k)$ is equal to some $(\pi(\bar{x}),\ \rho(\bar{x}))$. Since for every x satisfying $Ax=b$, $(x,\ Q(x))$ is feasible for (7), it follows that in (7), min $\tilde{z}\leqslant$ min z, where min z solves (4). Thus, if $(\bar{x},\ \bar{\theta})$ solves (7) and $Q(\bar{x})=\bar{\theta}$, $c\bar{x}+Q(\bar{x})\leqslant$ min z and $\bar{x}$ solves (4).

On the other hand, if $Q(\bar{x})\neq\bar{\theta}$, it follows (since $c\bar{x}+Q(\bar{x})\geqslant$ min z) that $Q(\bar{x})>\bar{\theta}$. Also $(\pi(\bar{x})T)\bar{x}+Q(\bar{x})=\rho(\bar{x})$ by taking expectations in the expression $\pi(\bar{x},\ \tilde{\xi})(\tilde{\xi}-Tx)=Q(\bar{x},\ \tilde{\xi})$. Thus $(\bar{x},\ \bar{\theta})$ violates the expression (10).

One problem which arises is that it is possible to generate a large number of constratints of type (7b). However, in reference [7] it is shown that all constraints in (7b) that are not tight (i.e., where equality does not hold) in the optimal solution to (7) may be dropped from succeeding iterations. Thus, the number of such constraints never exceeds the number of rows in T. Formally, the algorithm is as follows with S initially zero in step 1.

1. Solve for $x \geqslant 0$, min $\tilde{z}$ such that

$$cx+\theta=\tilde{z}$$
$$Ax=b$$
$$(\pi^k T)x+\theta-s^k=\rho^k \qquad k=1,\ldots,S.$$

Let $\bar{x}$ be the optimal x obtained.

2. If $s^k>0$, delete the constraint in which s^k appears, $k=1,\ldots,S$ and renumber those remaining.
3. Sample from the distribution of ξ. For each point $\xi^i(i=1,\ldots,n)$ in the sample, solve for

$$y \geqslant 0,\ Q(\bar{x},\xi) \text{ where}$$
$$Q(\bar{x},\xi^i)=\min qy \text{ subject to}$$
$$Wy=\xi^i-T\bar{x}.$$

Increase S by 1 and let $\pi^s=n^{-1}\Sigma\pi(\bar{x},\xi^i)$ and $\rho^s=n^{-1}\Sigma\xi^i\pi(\bar{x},\xi^i)$, where n is the sample size and $\pi(\bar{x},\xi^i)$ are the optimal simplex multipliers for the sample point ξ^i.

4. If $Q(\bar{x})-\theta\leqslant\delta$, where δ is a predetermined constant or variable, terminate. Otherwise, return to 1.

The quantity δ in step 4 may be chosen by any sensible criteria. For example, it may be a constant or it could be a small fraction times the last step 1 solution (i.e., $z(\bar{x})$).

When the second stage exhibits simple recourse (i.e., the only second-stage variables are slack variables) as is the case in the daily model, step 3 can be considerably simplified. In this case the sampling procedure and the need to solve a linear program can be dispensed and the expected value of the shadow prices calculated explicitly. Specifically, $Wy=\xi^i-T\bar{x}$ may be expressed as $q_i^+-q_i^-=\xi^i-Tx$, $i=1,\ldots,m_2$, and from [8]:

$$E\pi_i(x^k)=q_i^+\int_{\xi_i\geqslant T_ix^k}dF(\xi_i)-q_i^-\int_{\xi_i<T_ix^k}dF(\xi_i),\qquad i=1,\ldots,m_2,$$

where $F(\xi_i)$ is the cumulative of the distribution of ξ_i. That is, they are equal to the cost of the i-th excess demand times the probability of its occurrence minus the cost of the excess supply in the i-th equation times the probability of its occurrence.

The Second-Stage Problem — The sampling involved in the calculation of $Q(\bar{x})$ in step 3 of the algorithm involves the solution of many linear programs. The solution time for this step can greatly be reduced by taking advantage of the fact that if a basis is optimal for a program and the right-hand side is changed, that basis will either be optimal or infeasible for the new right-hand side. Thus, the sequence of programs for the second stage may be solved as follows. Solve the program for one sample point. Check each of the other sample points for feasibility with respect to the optimal basis found.

(The π vector associated with this basis is an optimal one for all sample points that are feasible.) Then solve the program for some sample point that was infeasible and repeat. This procedure reduces the number of linear programs from the number of elements in the sample to the number of optimal bases for the sample points. After one optimal basis is found, others may of course be found from this one by a sequence of dual simplex pivot steps. Hence, each pivot will yield a basis whose cost row is nonnegative. It is suggested, then, that after each iteration, feasibility be checked for the other sample points which have not been solved for, rather than wait until the current one being considered is feasible. It is also suggested that the optimal basis be saved from iteration to iteration.

Gradient Method

In this approach the equivalent convex program is solved by a nonlinear programming algorithm which utilizes the gradient of the objective function of (4) at x°,

$$\nabla z(x^\circ) = (c - \bar{\pi}(x^\circ)T).$$

The method we describe is based upon a variant of the Frank-Wolfe algorithm [5] for convex programming with linear constraints developed by Wets [9] for stochastic programs. Other gradient methods for convex programming, such as the Gradient Projection method, Zangwill's convex simplex method [11], or Zoutendijk's method of feasible directions [12] could be employed as well.

In brief, this method obtains solutions to a sequence of linear approximations to the given problem of the form

(11)
$$\min_x \nabla z(x^{l-1})x$$

subject to

$$Ax = b$$
$$x \geqslant 0,$$

where x^{l-1} is the solution for x to the given problem at the preceding or $(l-1)$st iteration. An improved feasible solution x^l is obtained by adding to x^{l-1} an improving feasible direction $\lambda(\bar{x}^l - x^{l-1})$, where $\bar{x}^l$ is the (basic) optimal solution to (11) and λ(a scalar) is determined by solving a one-dimensional optimization problem to be described below.

Several features of this method should be noted. First, the linear program which is actually solved in the first stage is not augmented but contains only constraints of the first stage; only the cost function varies from iteration to iteration. Second, the computational labor is again split between first and second stages with the expected values of the shadow prices, which are used in the objective function of the first-stage problem, generated in a sampling procedure in solving a set of second-stage programs for a fixed x. Third, since the method is a gradient approach, it is self-correcting in the sense that if in an earlier iteration a poor approximation to the gradient is generated by an unusually large sampling error which leads to a small improvement in the solution, it is likely to be corrected at later iterations. By contrast, in the cutting-plane approach, if an unusually large sampling error generates a support function which cuts inside the feasible region, there is no assurance that future iterations with smaller sampling errors will eliminate this constraint.

Algorithm—The exact steps of the algorithm are as follows:

1. (Initialization.) Set iteration count $l=0$, $x^l=0$ and $\bar{\pi}^{l-0} \equiv \bar{\pi}(x^l)=0$ [the sample estimate of the expected value of the optimal multipliers for the second-stage program].

2. (Generate optimal solution to the linearized equivalent convex program.) Set $l = l+1$. For $l = 1, 2, \ldots$, solve the linear program

(i)
$$\min_x \, (c - \bar{\pi}^{l-1}T)x,$$

subject to

$$Ax = b$$
$$x \geq 0.$$

Denote the optimal solution to (i) by $\bar{x}^l$. For $l = 1$, go to step 5. Otherwise go to step 3.

3. (Seek improved direction.) Solve the scalar optimization problem:

(ii)
$$\min_{0 \leq \lambda \leq 1} \psi(\lambda) = c[(1-\lambda)x^{l-1} + \lambda \bar{x}^l]$$
$$+ \bar{Q}[(1-\lambda)x^{l-1} + \lambda \bar{x}^l],$$

where $\bar{Q}$ is a sample estimate of $Q(x)$ obtained by solving the second-stage program for a random sample of ξ and a fixed x. In (ii) this function is viewed as a function of the scalar parameter λ, which weights the previous solution with the current optimal solution to the linearized problem. Denote the optimal solution to (ii) by λ^*.

Set $x^l = (1-\lambda^*)x^{l-1} + \lambda^* \bar{x}^l$.

4. (Test for convergence.) If $\lambda^* = 0$, terminate with x^{l-1} as the optimal solution. Otherwise, set $x^l = (1-\lambda^*)x^{l-1} + \lambda^* \bar{x}^l$ and go to step 5.

5. (Determine $\bar{\pi}^l$.) This is accomplished as in step 3 of the cutting-plane algorithm by sampling from the distribution of ξ and for each solving the resulting linear programs parametrically in the general case and taking the average of the resulting simplex multipliers, or explicitly calculating $\bar{\pi}^l$ for the case of simple recourse. Go to step 2.

Some additional remarks about the algorithm are in order. The auxiliary minimization problem that is required in step 3 requires a considerable amount of computation.

An alternative convergence criterion is available at step 4. From results in [9] an upper bound on the difference between the minimum value of the objective function $\hat{z}(x^\circ)$ and the current value $\hat{z}(x^l)$ is given by

(12)
$$\hat{z}(x^l) - L_h,$$

where

$$L_h = \max_{h=1,\ldots,l} \hat{z}(x^h) + [c - \pi(x^h)T](\bar{x}^h - x^h).$$

Thus the problem may be terminated when (12) is less than some prespecified constant.

In addition, we note that a majority of the constraints in the second-stage problem of the monthly model, i.e., those of the form

$$\sum_{k \neq j} a_{ijk} x_{ijk} \leq a_{ij} x_{ij},$$

---

*The constraints may be reduced to this form without loss of generality by dividing each coefficient by the right-hand side and rescaling the variables.

and some of the first-stage constraints in the daily model are of a type known as generalized upper bounds [3]. These are constraints which appear schematically as

$$\begin{array}{llllll} 1\ 1\ \dots\ 1 & & & & \leqslant b \\ & 1\ 1\ \dots\ 1 & & & \leqslant b_2 \\ & & \ddots & & \vdots \\ & & & 1\ 1\ \dots\ 1 & \leqslant b_2. \end{array}$$

That is, they are disjoint sums where the right-hand side is a nonnegative constant.* A special algorithm is available for solving problems of this form, which entails carrying a basis where size is reduced by L, the number of generalized upper-bound constraints.

Finally, as mentioned earlier, some rounding scheme is necessary to convert the numbers of assigned flights to integers. One scheme might be to first round up solutions whose fractional parts are 0.5 or greater and round down those whose fractional parts are less than 0.5. Then, if some flying hour constraint is exceeded, subtract one from a sufficient number of variables to satisfy that constraint. The variables to be reduced should be chosen from among those previously rounded up and may be those whose fractional part is closest to 0.5.

V. NUMERICAL EXAMPLE

We conclude with a small-scale numerical example of the solution of the monthly model. The problem consists of two aircraft types and two routes. The data were chosen for convenience. Solutions based on actual data will be reported in the future.

The data for the problem were as follows:

Aircraft type	Flying hours (round trip)		Carrying capacity (in tons)		Cost per flight (dollars)	
	Route 1	Route 2	Route 1	Route 2	Route 1	Route 2
1	24	14	50	75	7200	6000
2	49	29	20	20	7200	4000

Penalty cost ($ per ton)		
Route	Excess demand	Excess supply
1	500	0
2	250	0

In addition, the time to switch between either route was 5 hours for aircraft type 1 and 7 hours for aircraft type 2. The incremental costs for these moves were $1000 and $1500 for these aircraft types, respectively. It was assumed that 720 hours of flying time were available for each aircraft type. This can correspond to 3 planes of each type with an average utilization of 8 hours per day.

Finally, demand was assumed to be independently lognormally distributed on each route with the following parameters:*

Route	μ	σ
1	1000	50
2	1500	300

The following solution was found using an experimental program for the cutting-plane version of the algorithm.**

Aircraft type	Route	
	1	2
1	16.5	23.2
2	6.7	0.0

The total cost of the entire program converged after four iterations (based upon a termination when the relative decrease in costs fell below 1 percent). Each second-stage program was based upon drawing a sample of 25 observations so that, in effect, about 100 linear programming problems were solved during the course of the iterations. Execution time was 10 seconds.

For comparison purposes, a deteministic linear program was run with demand on routes assumed to be equal to the mean used in the lognormal case. This yielded the following results:

Aircraft type	Route	
	1	2
1	18.3	20.0
2	4.2	0.0

It should be noted that this solution differs from the case where the demands are stochastic. It might be supposed that the stochastic version would always assign more aircraft as a hedge against uncertainty; however, this is not the case even in this simple example—more flights of type 1 aircraft to route 1 are made in the deterministic version than in the stochastic.

*The current programs allow only for the lognormal distribution. It is a simple matter to add subroutines to generate random samples from arbitrary distributions whose cumulative distribution function can be tabulated. The lognormal distribution was chosen for experimental purposes because it produces nonnegative random variables which are skewed to the right—characteristics actual demand patterns are likely to exhibit.

**The program was written in FORTRAN and employs a small, all-in-core linear programming subroutine. Currently, the program is being modified to handle much larger problems and well be described in future work. A version of the algorithm for the monthly model using the gradient algorithm was also developed. Preliminary runs indicate that the gradient method takes considerably longer to run than the cutting-plane version because of the number of additional programs that must be solved during the auxiliary minimization problem described in Sec. IV.

ACKNOWLEDGMENTS

We wish to thank Dr. Roger Wets of the Boeing Scientific Research Laboratories for helpful discussions. We also wish to thank Lois Littleton of RAND, who wrote the program described in Sec. V and contributed a number of useful suggestions.

REFERENCES

[1] Benders, J. F., "Partition Procedures for Solving Mixed-Variables Programming Problems," Num. Math., *4*, 238–252 (1962).

[2] Dantzig, G. B. and A. Madansky, "On the Solution of Two-Stage Linear Programs Under Uncertainty," *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* (University of California Press, Berkeley and Los Angeles, 1961) Vol. 1, 156–176.

[3] Dantzig, G. B. and R. M. Van Slyke, "Generalized Upper Bounding Techniques for Linear Programming," J. Computer and System Science, *1*, 213–226 (1967).

[4] Ferguson, A. R. and G. B. Dantzig, "The Allocation of Aircraft to Routes—An Example of Linear Programming Under Uncertain Demand," Mgmt. Sci., *3*, 45–73 (1956).

[5] Frank, M. and P. Wolfe, "An Algorithm for Quadratic Programming," Nav. Res. Log. Quart., *3*, 95–110 (1956).

[6] Kelley, J. L., Jr., "The Cutting-Plane Method for Solving Convex Programs," SIAM J. Appl. Math., *8*, 703–712 (1960).

[7] Murty, K. G., "Two-Stage Linear Program Under Uncertainty: A Basic Property of the Optimal Solution," ORC 66–4, Operations Research Center, University of California, Berkeley (Feb. 1966).

[8] Van Slyke, R. M. and R. J. Wets, "L-Shaped Linear Programs with Applications to Optimal Control and Stochastic Programming," ORC 66–17, Operations Research Center, University of California, Berkeley (June 1967) (Revised); (Also to appear in SIAM J. Appl. Math.)

[9] Wets, R. J., "Programming Under Uncertainty: The Complete Problem," *Zeit. Wahr. und Verw. Geb.*, *4*, 316–339 (1966).

[10] Wets, R. J., "Programming Under Uncertainty: The Equivalent Convex Program," SIAM J. Appl. Math., *14*, 89–105 (1966).

[11] Zangwill, W. J., "The Convex Simplex Method," Mgmt. Sci., *14*, 221–230 (1967).

[12] Zoutendijk, G., *Methods of Feasible Directions* (Elsevier Publishing Company, Amsterdam and New York, 1960).

* * *