ment for forwarder $i \in \mathcal{M}$ is given as follows: $x_1^* = a_1^*(\kappa)$ and $x_i^* = a_i^*(\kappa - \sum_{k=1}^{i-1} x_k^*)$ for each $i \in \{2, 3, \ldots, m\}$.

Proof. Since the carrier has κ units of cargo capacity to allocate to forwarders 1 through m, the optimal objective function is $u_1(\kappa)$. From the definition, $a_i^*(y)$ represents the optimal allotment for forwarder i, if the carrier has y units to allocate to forwarders i through m. Thus, the optimal allotment for forwarder 1 is $x_1^* = a_1^*(\kappa)$. The carrier reserves x_1^* to forwarder 1, leaving $\kappa - x_1^*$ to allocate to forwarders 2 through m. Then, $x_2^* = a_2^*(\kappa - x_1^*)$. The carrier reserves x_2^* to forwarder 2, leaving $\kappa - (x_1^* + x_2^*)$ to allocate to forwarders 3 through m. Then, $x_3^* = a_3^*(\kappa - (x_1^* + x_2^*))$. The carrier reserves x_3^* for forwarder 3, leaving $\kappa - (x_1^* + x_2^* + x_3^*)$ to allocate to forwarders 4 through m, and so on. □

The number of iterations needed to solve (9) is of order $m\kappa^2$. This may create some computational burden, if the carrier is endowed with large cargo capacity. In the next section, we develop some heuristic solutions, implementable for an industry-sized problem.

3 Heuristic Solutions

The fact that the expected actual usage $E[U_i(a)]$ of forwarder i that receives allotment a is not concave in $a \in \mathbb{Z}_+$ results from the all-or-none acceptance rule. To develop some heuristics, we suppose that the shipment can be partially accepted. Upon receiving the k-th booking request with space requirement $W_{i,k}$ from forwarder i, the airline accepts $\min\{W_{i,k}, Z\}$ where Z is the unused portion of the allotment. We will show that the expected total usage is concave under the assumption that the request is partially accepted.

Suppose that forwarder i receives allotment x. Let $Y_{i,k}(x)$ denote its cumulative usage after the carrier's accept/reject decision of the k-th booking request. The actual allotment usage under the partial acceptance assumption is $V_i(x) = Y_{i,N_i}(x)$. The finite sequence of the cumulative usages $\{Y_{i,k}(x) : k = 1, 2, \ldots, N_i\}$ is determined by the following recurrence equation:

$$Y_{i,k}(x) = Y_{i,k-1}(x) + \min\left\{(x - Y_{i,k-1}(x)), W_{i,k}\right\} \quad (10)$$

for each $k = 1, 2, \ldots, N_i$, and $Y_{i,0}(x) = 0$. In (10), the carrier books the shipment up to the unused portion of the allotment $(x - Y_{i,k-1}(x))$. We assume that the forwarder is willing to break the consolidated shipment and deliver $\{(x - Y_{i,k-1}(x)), W_{i,k}\}$.

From our construction, $Y_{i,k}(x) \geq X_{i,k}(x)$ for all k, with probability 1. Consequently, the expected actual allotment usage under the assumption that the request is partially accepted is at least that under the all-or-none assumption; i.e.,

$$E[V_i(x)] \geq E[U_i(x)] \quad (11)$$

Furthermore, if $W_{i,k}$ is a Bernoulli random variable, which takes on values $\{0, 1\}$, then with probability 1, $Y_{i,k}(x) = X_{i,k}(x)$ for all k, so $U_i(x) = V_i(x)$.

For each $i \in \mathcal{M}$, let D_i denote the sum of all space requirements of N_i booking requests from forwarder i; i.e., $D_i = \sum_{k=1}^{N_i} W_{i,k}$ is the total space requirement of forwarder i.

Proposition 3. *Suppose that forwarder i receives allotment $a \in \mathbb{R}_+$. Under the assumption that the request is partially accepted, $Y_{i,k}(a) = \min\{\sum_{j=1}^{k} W_{i,j}, a\}$ for each $k = 1, 2, \ldots, N_i$.*

Proof. The proof is done using mathematical induction. Clearly, $Y_{i,1}(x) = \min\{W_{i,1}, x\}$. Next, assume that $Y_{i,k-1}(x) = \min\{\sum_{j=1}^{k-1} W_{i,j}, x\}$. Substituting this expression in (10), we get

$$Y_{i,k}(x) = \min\{\sum_{j=1}^{k-1} W_{i,j}, x\} + \min\left\{(x - \min\{\sum_{j=1}^{k-1} W_{i,j}, x\}), W_{i,k}\right\}$$

If $\sum_{j=1}^{k-1} W_{i,j} < x$, then $Y_{i,k}(x) = \sum_{j=1}^{k-1} W_{i,j} + \min\{x - \sum_{j=1}^{k-1} W_{i,j}, W_{i,k}\} = \min\{x, \sum_{j=1}^{k} W_{i,j}\}$;

otherwise, $Y_{i,k}(x) = x$. Hence, $Y_{i,k}(x) = \min\{x, \sum_{j=1}^{k} W_{i,j}\}$. □

For short-hand notation, denote $\rho_i(x) = p_i E[V_i(x)]$. It follows from (11) that the optimal solution to the following mathematical program

$$\xi = \max\left\{\sum_{i=1}^{m} \rho_i(a_i) : \sum_{i=1}^{m} a_i \leq \kappa, \quad a_i \geq 0 \text{ for each } i \in \mathcal{M}\right\} \tag{12}$$

yields an upper bound on the carrier's problem (8); i.e., $\xi \geq \zeta$. Unlike $\pi_i(x)$ in the carrier's problem, the function $\rho_i(x)$ is concave on $\mathbb{Z}_+$. Similar to $\pi_i(a)$, the function $\rho_i(a)$ is not differentiable on $\mathbb{R}_+$, since $W_{i,k}$ and N_i are assumed to be nonnegative integer-valued random variables. To solve (12), we present two heuristic approaches.

1. Continuous approximation, in which the total space requirement is modeled as a nonnegative real-valued random variable, denoted by $\hat{D}_i$. Since the cumulative distribution function of $\hat{D}_i$ is continuous, the expected actual allotment usage of forwarder i that receives allotment x, $E[\min(\hat{D}_i, x)]$, is differentiable on $\mathbb{R}_+$. This allows us to apply a standard nonlinear programming technique (e.g., KKT conditions).

2. Lagrangian relaxation, in which we dualize the capacity constraint, $\sum_{i=1}^{m} x_i \leq \kappa$. The objective function of the relaxed problem is concave, so the relaxed problem for a fixed value of the Lagrange multiplier is easy to solve. We then use subgradient optimization to update the Lagrange multiplier in such a way that the capacity constraint is likely to be tighter on the subsequent iteration.

We use boldface type to denote an m-dimensional vector; e.g., $\mathbf{x} = (x_1, x_2, \ldots, x_m)$.

3.1 Continuous Approximation

Recall that the total space requirement of forwarder i, $D_i = \sum_{k=1}^{N_i} W_{i,k}$, is a $\mathbb{Z}_+$-valued random variable. Using continuous approximation, we model it as an $\mathbb{R}_+$-valued random variable $\hat{D}_i$, whose mean μ_i and variance σ_i^2 are chosen such that

$$\mu_i = E[D_i] = E[N_i]E[W_{i,1}], \tag{13}$$
$$\sigma_i^2 = \mathrm{var}(D_i) = \mathrm{var}(W_{i,1})E[N_i] + (E[W_{i,1}])^2 \,\mathrm{var}(N_i),$$

Problem (12) becomes

$$\hat{\xi} = \max\Big\{ \sum_{i=1}^{m} p_i E[\min(\hat{D}_i, x_i)] : \sum_{i=1}^{m} x_i \le \kappa, \quad x_i \ge 0 \text{ for each } i \in \mathcal{M} \Big\} \tag{14}$$

Let F_i denote the cumulative distribution function of $\hat{D}_i$. Index the forwarders such that $p_i \ge p_{i+1}$ for all $i \in \mathcal{M}$, where $p_{m+1} = 0$. Since $\hat{D}_i$ is an $\mathbb{R}_+$-valued random variable, F_i is continuous and strictly increasing, and the quantile F_i^{-1} is a well-defined function.

Proposition 4. *The necessary and sufficient conditions for $\hat{\mathbf{x}}^*$ to be an optimal solution to (14) are as follows: There exists $\ell^* \in \mathcal{M}$ and a positive $\lambda^* \in [p_{\ell^*+1}, p_{\ell^*})$ such that*

$$\hat{x}_i^* = \begin{cases} F_i^{-1}(1 - \lambda^*/p_i) & \text{for each } i = 1, 2, \ldots, \ell^* \\ 0 & \text{for each } i = \ell^* + 1, \ell^* + 2, \ldots, m \end{cases}$$

and $\sum_{i=1}^{\ell^*} \hat{x}_i^* = \kappa$.

Proof. Since F_i is continuously differentiable for each $i \in \mathcal{M}$, the objective function $\sum_{i=1}^{m} p_i E[\min(\hat{D}_i, x_i)]$ is continuously differentiable on $\mathbb{R}_+^m$. Moreover, it possesses continuous second partial derivatives, so the KKT conditions are necessary for a point to be an optimal solution. The KKT conditions are also sufficient, since the objective function is concave on $\mathbb{R}_+^m$. (All constraints are linear, so the constraint qualifications/regularity conditions are satisfied.)

We associate a vector of multipliers $\boldsymbol{\mu} \ge \mathbf{0}$ with the nonnegativity constraints and $\lambda \ge$

0 with the capacity constraint to form the Lagragian function:

$$\mathcal{L}(\mathrm{x}|\boldsymbol{\mu}, \lambda) = \sum_{i=1}^{m} p_i E[\min(\hat{D}_i, x_i)] + \lambda(\kappa - \sum_{i=1}^{m} x_i) + \sum_{i=1}^{m} \mu_i x_i$$

For a feasible solution $\mathbf{x}$ to be an optimal solution to (14), the KKT conditions

$$p_i \bar{F}_i(x_i) - \lambda + \mu_i = 0 \qquad \text{for } i \in \mathcal{M}$$

$$\lambda(\kappa - \sum_{i=1}^{m} x_i) = 0$$

$$\mu_i x_i = 0 \qquad \text{for } i \in \mathcal{M}$$

are both necessary and sufficient. Since $\boldsymbol{\mu} \geq \mathbf{0}$, we can eliminate them; the above conditions can be written as

$$p_i \bar{F}_i(x_i) \leq \lambda \qquad \text{for } i \in \mathcal{M} \tag{15}$$

$$\lambda(\kappa - \sum_{i=1}^{m} x_i) = 0 \tag{16}$$

$$[p_i \bar{F}_i(x_i) - \lambda] x_i = 0 \qquad \text{for } i \in \mathcal{M} \tag{17}$$

Note that $\bar{F}_i(0) = 1$ and $\bar{F}_i(x)$ is strictly decreasing on $[0, \kappa]$. From (17), we conclude that $x_i > 0$ if and only if $p_i > \lambda$. From this result and the way the forwarders are indexed, there exists $\ell \in \mathcal{M}$ such that $x_i > 0$ for $i \leq \ell$ and $x_i = 0$ for $i > \ell$. For each $i \leq \ell$, $p_i \bar{F}_i(x_i) = \lambda$; the multiplier must be $\lambda < p_\ell$. For $i > \ell$, $x_i = 0$ and $p_{\ell+1} \leq \lambda$. Finally, suppose that $\lambda = 0$. Then, it follows from (15) that x_i is the largest possible value of $\hat{D}_i$, and the capacity constraint $\sum_{i=1}^{m} x_i \leq \kappa$ is violated. Hence, $\lambda > 0$, and (16) becomes $\sum_{i=1}^{m} x_i = \kappa$. □

Proposition 4 asserts that the first ℓ^* forwarders receive positive allotments, whereas the last $(m - \ell^*)$ forwarders receive zero allotments. For forwarders 1 through ℓ^*, its allotment is chosen such that the marginal revenue is equal to the Lagrange multiplier [i.e., $p_i\bar{F}_i(\hat{x}_i^*) = \lambda^*$], and that the sum of their allotments equals the capacity. That is,

$$\sum_{i=1}^{\ell^*} F_i^{-1}(1-\lambda^*/p_i) = \kappa \text{ where } p_{\ell^*+1} \le \lambda^* < p_{\ell^*} \tag{18}$$

The two-phase algorithm in Table 1 finds an optimal solution to (14). We search for the optimal number of forwarders that would receive positive allotments (denoted as ℓ^*) in Phase I and for the optimal Lagrange multiplier (denoted as λ^*) in Phase II. The values of ℓ^* and λ^* must satisfy (18).

The idea behind Phase I is as follows. We want to find ℓ^* such that $\sum_{i=1}^{\ell^*} F_i^{-1}(1-\lambda/p_i) = \kappa$ for some $\lambda \in [p_{\ell^*+1}, p_{\ell^*})$. For a fixed value of $\ell \in \mathcal{M}$, the function $\sum_{i=1}^{\ell} F_i^{-1}(1-\lambda/p_i)$ is decreasing in $\lambda \in [p_{\ell+1}, p_\ell)$. Then,

$$\sum_{i=1}^{\ell} F_i^{-1}(1-p_\ell/p_i) < \sum_{i=1}^{\ell} F_i^{-1}(1-\lambda/p_i) \le \sum_{i=1}^{\ell} F_i^{-1}(1-p_{\ell+1}/p_i)$$

Let LB(ℓ) and UB(ℓ) denote the quantities on the left- and right-hand sides, respectively. Note that $\sum_{i=1}^{\ell} F_i^{-1}(1-\lambda/p_i)$ is continuous in λ. The intermediate value theorem asserts that if LB(ℓ) $\le \kappa \le$ UB(ℓ), then there exists $\lambda(\ell) \in [p_{\ell+1}, p_\ell)$ such that $\sum_{i=1}^{\ell} F_i^{-1}(1-\lambda(\ell)/p_i) = \kappa$. Phase I determines the smallest integer ℓ^* such that LB(ℓ^*) $\le \kappa \le$ UB(ℓ^*).

Phase II searches for $\lambda^* \in [p_{\ell^*+1}, p_{\ell^*})$ that solves $\sum_{i=1}^{\ell^*} F_i^{-1}(1-\lambda^*/p_i) = \kappa$, where ℓ^* is found in Phase I. This can be done using a one-dimensional search procedure. In Table 1 Phase II, we present a bisection method. In iteration $t \ge 1$, we employ the midpoint rule (traditionally called the Bolzano search plan) for selecting the trial solution $\lambda_t = (\lambda_t' + \lambda_t'')/2$. If $\sum_{i=1}^{\ell^*} F_i^{-1}(1-\lambda_t/p_i) > \kappa$, we need to increase the lower bound $\lambda_{t+1}' = \lambda_t$. If $\sum_{i=1}^{\ell^*} F_i^{-1}(1-\lambda_t/p_i) < \kappa$, we need to decrease the upper bound $\lambda_{t+1}'' = \lambda_t$. [Again, note that $\sum_{i=1}^{\ell^*} F_i^{-1}(1-\lambda/p_i)$ is decreasing in λ.] By construction, the length of the interval of uncertainty in iteration $(t+1)$ is halved that in iteration t. We stop when $\sum_{i=1}^{\ell^*} F_i^{-1}(1-\lambda_t/p_i)$ is close to κ.

- Phase I: Search for $\ell^* \in \mathcal{M}$.
  1. Initialization: Set $\ell = 1$.
  2. Iteration $\ell \geq 1$:
     (a) Set $\text{UB}(\ell) = \sum_{i=1}^{\ell} F_i^{-1}(1 - p_{\ell+1}/p_i)$.
     (b) Set $\text{LB}(\ell) = \sum_{i=1}^{\ell} F_i^{-1}(1 - p_\ell/p_i)$.
  3. Stopping:
     (a) If $\ell = m$, set $\ell^* = m$ and go to Phase II.
     (b) If $\text{LB}(\ell) \leq \kappa \leq \text{UB}(\ell)$, set $\ell^* = \ell$ and go to Phase II. Otherwise, set $\ell = \ell + 1$ and perform the next iteration.
- Phase II: Search for $\lambda^* \in [p_{\ell^*+1}, p_{\ell^*})$.
  1. Initialization: Select a small tolerance $\epsilon > 0$.
     (a) Set initial lower bound $\lambda_1' = p_{\ell^*+1}$.
     (b) Set initial upper bound $\lambda_1'' = p_{\ell^*}$.
  2. Iteration $t \geq 1$:
     (a) Compute new multiplier $\lambda_t = (\lambda_t' + \lambda_t'')/2$.
     (b) Computer allotment $x_{it} = F_i^{-1}(1 - \lambda_t/p_i)$ for $i = 1, 2, \ldots, \ell^*$.
  3. Stopping: If $|\sum_{i=1}^{\ell^*} x_{it} - \kappa| < \epsilon$, set $\lambda^* = \lambda_t$ and $\hat{x}_i^* = x_{it}$, and stop. Otherwise,
     (a) If $\sum_{i=1}^{\ell^*} x_{it} > \kappa$, set $\lambda_{t+1}' = \lambda_t$, and $\lambda_{t+1}'' = \lambda_t''$.
     (b) If $\sum_{i=1}^{\ell^*} x_{it} < \kappa$, set $\lambda_{t+1}'' = \lambda_t$, and $\lambda_{t+1}' = \lambda_t'$.
     (c) Set $t = t + 1$ and perform the next iteration.

Table 1: Continuous approximation heuristic procedure

3.2 Lagrangian relaxation

In Problem (12), we relax the capacity constraint by multiplying it by Lagrange multiplier $\nu \geq 0$ and bringing it into the objective function, which now becomes

$$\sum_{i=1}^{m} \rho_i(a_i) + \nu(\kappa - \sum_{i=1}^{m} a_i) = \sum_{i=1}^{m} [\rho_i(a_i) - \nu a_i] + \nu\kappa$$

For short-hand notation, denote $c_i(a|\nu) = \rho_i(a) - \nu a$. We want to solve

$$\min\{\xi(\nu) : \nu \geq 0\} \text{ where } \xi(\nu) = \max\{\sum_{i=1}^{m} c_i(a_i|\nu) : a_i \in \mathbb{Z}_+ \text{ for each } i \in \mathcal{M}\} \qquad (19)$$

Recall that in Problem (12), $\rho_i(a) = p_i E[\min(D_i, a)]$, and D_i is a $\mathbb{Z}_+$-valued random variable, so we can restrict our attention to a nonnegative integer allotment. For a fixed value of the Lagrange multiplier, the maximization $\xi(\nu)$ is easy to solve, since the objective function is separable, and $c_i(a|\nu)$ is concave on $\mathbb{Z}_+$ for each $i \in \mathcal{M}$.

Let G_i denote the cumulative distribution function of D_i for each $i \in \mathcal{M}$.

Proposition 5. *For a fixed value of the Lagrange multiplier ν, an optimal solution of $\xi(\nu)$ is as follows. If $\nu > p_i$, then $a_i^*(\nu) = 0$; otherwise,*

$$a_i^*(\nu) = \text{argmax}\left\{a \in \mathbb{N} : \bar{G}_i(a-1) \geq \nu/p_i\right\} \qquad (20)$$

Proof. Since the objective function is separable, we can individually maximize each term $c_i(a|\nu)$. Since D_i is a $\mathbb{Z}_+$-valued random variable, we have that for each $a \in \mathbb{Z}_+$

$$E[\min(D_i, a)] = \sum_{t=0}^{\infty} P(\min(D_i, a) > t) = \sum_{t=0}^{a-1} P(D_i > t) = \sum_{t=0}^{a-1} \bar{G}_i(t) \qquad (21)$$

where the first equation follows from the result $E[Z] = \sum_{t=0}^{\infty} P(Z > t)$ for a $\mathbb{Z}_+$-valued random variable Z. Using (21), we obtain the expression

$$c_i(a|\nu) = p_i E[\min(D_i, a)] - \nu a = p_i \sum_{t=0}^{a-1} \bar{G}_i(t) - \nu a$$

and the first difference is

$$c_i(a|\nu) - c_i(a-1|\nu) = p_i \bar{G}_i(a-1) - \nu$$

If $\nu > p_i$, then the first difference is negative for all $a \geq 0$, so $c_i(a|\nu)$ is nonincreasing, and $a_i^*(\nu) = 0$. Otherwise, we obtain the expression (20). □

The solution $\mathbf{a}^*(\nu)$ found in Proposition 5 can be used to obtain an upper bound:

$$\sum_{i=1}^{m} c_i(a_i^*(\nu)|\nu) \geq \xi \geq \zeta$$

The first inequality follows from the fact that $\xi(\nu) \geq \xi$ for all $\nu \geq 0$, and the second inequality follows from (11). Also, it can be used to obtain a lower bound. Let

$$x_i^*(\nu) = \lfloor \tilde{x}_i^*(\nu) \rfloor \text{ where } \tilde{x}_i^*(\nu) = \left[a_i^*(\nu) - \frac{1}{m^*(\nu)} (\sum_{i=1}^{m} a_i^*(\nu) - \kappa)^+ \right]^+ \tag{22}$$

where $m^*(\nu)$ is the number of non-zero allotments, i.e., the size of the set $\{i \in \mathcal{M} : a_i^*(\nu) > 0\}$. Note that $\mathbf{x}^*(\nu)$ is a feasible solution and can be used to obtain a lower bound $\sum_{i=1}^{m} \rho_i(x_i^*(\nu))$ to the approximated problem (12) or $\sum_{i=1}^{m} \pi_i(x_i^*(\nu))$ to the original problem (8).

We want to solve (19). The Lagragian relaxation algorithm is presented in Table 2. In iteration k, for a fixed value of the Lagrange multiplier ν_k, we solve $\xi(\nu_k)$: Its solution $\mathbf{a}^*(\nu_k)$ is given in Proposition 5, and it is used to construct a feasible solution $\mathbf{x}^*(\nu_k)$ as in (22). The first and latter are used to compute the upper and lower bounds of ξ in (12), respectively. The best upper and lower bounds that we have found during itera-