Multi-vehicle sequential resource allocation for a nonprofit distribution system

BURCU BALCIK1,∗ SEYED IRAVANI2 and KAREN SMILOWITZ2

1Industrial Engineering Department, Ozyegin University, Istanbul 34794, Turkey
E-mail: burcu.balcik@ozyegin.edu.tr
2Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL 60208, USA

Received October 2011 and accepted December 2013

This article introduces a multi-vehicle sequential allocation problem that considers two critical objectives for nonprofit operations: providing equitable service and minimizing unused donations. This problem is motivated by an application in food redistribution from donors such as restaurants and grocery stores to agencies such as soup kitchens and homeless shelters. A set partitioning model is formulated that can be used to design vehicle routes; it primarily focuses on equity maximization and implicitly considers waste. The behavior of the model in clustering agencies and donors on routes is studied, and the impacts of demand variability and supply availability on route composition and solution performance are analyzed. A comprehensive numerical study is performed in order to develop insights on optimal solutions. Based on this study, an efficient decomposition-based heuristic for the problem that can handle an additional constraint on route length is developed and it is shown that the heuristic obtains high-quality solutions in terms of equity and waste.

Keywords: Vehicle routing, resource allocation, clustering, equity, nonprofit operations, food banks

1. Introduction

Nonprofit organizations play an increasingly large role in delivering essential services to vulnerable and underserved members of society. According to the most recent hunger study by Feeding America (FA), the largest hunger relief organization in the United States, 37 000 000 Americans rely on FA for food and demand at member food banks has increased by 46% over the past 4 years (Feeding America, 2010). Food banks receive, store, and distribute food to agencies such as shelters, soup kitchens, and senior centers that directly serve communities. Rising demand and limited resources in economic downturns increase the importance of effective design and management of food distribution operations. Matthew Knott, the Vice President of Strategic Planning at the FA, notes (Matthew Knott, personal communication, 2009):

Increasingly, food banks are getting into the business of distribution. Traditionally, agencies came to food banks to collect donations. However, due to high transportation costs and limited resources (such as vehicles, staff), more agencies are looking to the food banks to deliver the donations.

Lien et al. (2013) address one such food distribution program: the Food Rescue Program (FRP) at the Greater Chicago Food Depository (GCFD), which collects the surplus perishable food from grocery stores, restaurants, and the like, and distributes the food to agencies.

Each day, five FRP vehicles are dispatched from the GCFD and each vehicle visits three to 17 donors and two to 11 agencies. Each donor or agency site is visited by one vehicle. Site visits take an average of 20 to 40 minutes, which limits the number of visits in a day per vehicle. Due to operating schedules of donors and agencies, vehicles visit donors before agencies. Collection activities begin at 6 a.m. and take about 3 to 5 hours. Agencies are run by volunteers and have limited operating hours. Typically, neither donation nor demand amounts are known before the driver visits the sites. Upon arriving at a donor, the driver accepts the full donation as vehicle capacity is generally not limiting. Once the driver arrives at an agency, the volunteers determine needs; demand depends on several factors, including the offered variety, current inventory, supplies provided by other sources, storage capabilities, and budget considerations. Ideally, agencies would determine and communicate demand information in advance. However, due to staffing

*Corresponding author
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/uiie.
and timing constraints at the agencies, volunteers are often not present at the locations to communicate with the driver a priori. Occasionally, drivers must wait at an agency site for the volunteers to arrive and/or determine their demand. After an agency expresses their demand, the driver decides the delivery amount. Although the driver tries to meet the agency’s demand, s/he also strives to reserve food for the remaining agencies along the route.

Lien et al. (2013) model the Sequential Resource Allocation (SRA) problem on a single FRP route and analyze (i) agency sequencing on a route and (ii) food allocation among the agencies. They show that using a waste-minimizing resource allocation objective results in inequitable allocation of food among agencies. Hence, they propose an alternative objective that maximizes the expected minimum fill rate among the agencies on the route, where fill rate is the ratio of the amount of food allocated to the demand. This objective achieves equity and low waste (high distribution) by raising the lowest service level on a route such that all participating agencies benefit. Lien et al. (2013) develop resource allocation and sequencing policies based on this objective and show that the policies also perform well in terms of minimizing waste.

In this article, we extend the single-route analysis of Lien et al. (2013) to a multi-route setting and integrate route composition into the decision making. Routing decisions highly affect the degree to which the supplies and demands can be matched; however, incorporating the SRA decisions into route design while ensuring equity and low waste is challenging. We formulate the equity-maximizing Multi-vehicle SRA problem (MSRA-e) that assigns agencies and donors to vehicles while maximizing the expected minimum fill rate among network agencies and examine the implications of this objective on route composition and service performance. We study the interaction of the equity-maximizing objective with supply availability and demand variability to develop a deeper understanding of the behavior of the MSRA-e. To evaluate the service performance of the MSRA-e in terms of waste, we define a benchmark problem, the waste-minimizing multi-vehicle SRA (MSRA-w), which solely focuses on maximizing resource utilization; comparison with this benchmark shows that MSRA-e achieves near-minimal waste while providing equitable service.

Solving the MSRA-e optimally requires (i) enumeration of all possible donor-agency assignments to vehicles; (ii) enumeration of all possible sequences of agencies for each assignment; and (iii) solution of a complex stochastic problem to evaluate each sequence, which is not practical for realistic size problem instances. Using the insights from our analysis, we develop an efficient decomposition-based heuristic for the MSRA-e that generates high-quality routes in terms of equity and waste. We generalize this heuristic to include travel time constraints.

In Section 2, we review the related literature. In Section 3, we define the MSRA-e and formulate a set-partitioning model. In Section 4, we present a numerical study of the MSRA-e, and in Section 5, we develop a decomposition-based heuristic for the problem. In Section 6, we extend our heuristic to incorporate a travel time constraint. Finally, we conclude the article and discuss future research in Section 7.

2. Literature review

SRA policies are studied in cases where resources or demands are stochastic and observed sequentially. Su and Zenios (2005) consider an equitable SRA problem in which resources are observed sequentially addressing allocation of organs among patient groups. Lien et al. (2013) present a review of sequential resource allocation problems in the literature for fixed delivery routes. Importantly, the papers reviewed focus on cost and profit objectives. Lien et al. (2013) develop equitable sequencing and allocation policies for a single route. Using stochastic dynamic programming, they characterize the structure of the optimal SRA policy along a route. They show that the sequence in which the agencies are visited may significantly affect equity. However, obtaining the optimal sequence and allocation policy for a set of agencies is tractable only for small problem instances. Lien et al. (2013) develop heuristics that result in near-optimal sequencing and allocation policies for the single-route problem.

The Vehicle Routing Problem (VRP) determines delivery routes for a set of capacitated vehicles, each of which originates from a depot, visits a set of customers to satisfy demand, and returns to the depot. Most VRPs focus on efficiency-based objectives such as minimizing total travel costs/times and minimizing number of vehicles. Although efficiency-based objectives serve the needs of commercial distribution problems, routing applications in the public/nonprofit sector often require other objectives that capture their non-financial performance requirements; see the review in Balcik et al. (2010). For instance, Campbell et al. (2008) show that minimizing the total travel time can lead to inequitable response times among demand locations served in a disaster relief environment. The authors propose minmax and minsum objectives to yield more equitable distribution. Huang et al. (2012) extend this work by considering demands at each recipient and additional equity metrics.

Several equity metrics and objectives are studied in the operations research literature. Marsh and Schilling (1994) review 20 equity metrics used in locating public facilities, including the Gini index, minmax, range, and deviation. Balcik et al. (2010) observe that minmax- and deviation-type metrics and objectives are most frequently used in public/nonprofit sector routing applications. Minmax- (or maxmin-) type objectives that incorporate equity by improving the condition of the least advantaged member, thereby seeking an egalitarian solution, are also widely
Multi-vehicle allocations in nonprofit organizations

1281

studied in resource allocation (Ibaraki and Katoh, 1988; Luss, 1999). Given the number of alternatives in measuring equity, some studies explore the trade-offs between the equity and other objectives such as efficiency (Mandell, 1991; Felder and Brinkmann, 2002).

This article contributes to the literature by introducing a multi-vehicle sequential allocation problem that incorporates two critical objectives for nonprofit distribution systems, namely, equity and waste. We consider an objective function that maximizes the expected minimum service level provided to the agencies, study the implications of this objective for route design and solution performance when multiple vehicles are used, and develop insights about the impact of demand variability and supply availability on route composition. We develop a heuristic approach for the MSRA-e that finds high-quality solutions in terms of both equity and waste and maintains route length feasibility.

3. The MSRA-e

Given a network of donors and agencies and set of vehicles, the MSRA-e is the problem of determining routes and allocation policies for delivery of perishable food collected from donors to agencies to achieve equitable and efficient (low waste) distribution. We describe the problem setting in detail in Section 3.1, present the mathematical model in Section 3.2, explain equitable allocation and sequencing policies in Section 3.3, and discuss solution requirements in Section 3.4.

3.1. Problem setting

Figure 1 presents an illustrative example of FRP operations with two vehicles. The supply at each donor \( q_j \) for all \( j \in M = \{ A, B, C, D \} \) is shown, along with the observed demand \( d_i \), the allocation amount \( x_i \) for all \( i \in N = \{ 1, 2, 3, 4 \} \), and the fill rate (the ratio of the allocated amount to observed demand) at each agency, \( \beta_i = x_i / d_i \forall i \in N \). Donor visits precede agency visits and each site is visited by one vehicle. All donations are accepted, and food allocation at agencies is determined by the driver once demands are observed.

The following three sets of decisions are involved in the MSRA-e.

1. Clustering: Which set of donors and agencies should be assigned to the same vehicle?
2. Sequencing: What should be the sequence of visits to agencies in each cluster?
3. Allocation: How should food be allocated to agencies along each route?

The MSRA-e optimizes these decisions according to an equity-based objective to achieve fair distribution of food among network agencies, while maintaining high distribution of donations. Specifically, we consider an objective that achieves fair food distribution by maximizing the lowest service level in the network; that is, by maximizing the expected minimum fill rate among all agencies (i.e., \( \max \mathbb{E}[\min\{\beta_1, \beta_2, \ldots, \beta_{|N|}\}] \)). This objective also reduces the expected waste as maximizing the expected minimum fill rate among agencies implicitly increases supply distribution in the network.

If \( K = 1 \), the MSRA-e simplifies to the single-route problem studied in Lien et al. (2013), which does not include the clustering decisions. Given the complexity of the integrated clustering, sequencing, and allocation decisions in the multi-route setting, we focus on a version of the problem with a single food category to enhance our understanding about the structure of equitable routes. In reality, there are eight categories, one or more of which are in high demand by the agencies. Here we assume that the routes are designed based on one most critical food category with the highest demand, such as meat. We consider a setting in which donation amounts are represented by their averages and, therefore, assumed to be known when clustering.
decisions are made. To study the impact of variability in supply, in Section 4.2.4, we test the robustness of the MSRA-e to changes in supply.

3.2. Set-partitioning formulation

We develop a set-partitioning formulation of the MSRA-e. Let \( R \) represent the set of feasible candidate routes, where a feasible route is one that meets the precedence criterion of visiting all donors before agencies and respects route length limitations. We define a covering parameter, \( \delta_{ir} \), such that \( \delta_{ir} = 1 \) if node \( i \in N \cup M \) is on route \( r \in R \) and is zero otherwise. Let \( y_r \) represent the choice to use route \( r \in R \): \( y_r = 1 \) if route \( r \in R \) is chosen and is zero otherwise. Let \( f_r \) represent the expected minimum fill rate among all nodes on route \( r \) given specified allocation and sequencing policies that maximize the minimum fill rate of the nodes on route \( r \). Given values of \( f_r \) for all \( r \in R \), \( Z_f \) is the expected minimum fill rate across all routes. The set-partitioning formulation is

\[
\max Z_f = \min_{i \in N} \left\{ \sum_{r \in R} f_r \delta_{ir} y_r \right\}, \quad (1a)
\]

subject to

\[
\sum_{r \in R} \delta_{ir} y_r = 1, \quad \forall i \in N \cup M \quad (1b)
\]

\[
\sum_{r \in R} y_r = K, \quad (1c)
\]

\[
y_r \in \{0, 1\}, \quad \forall r \in R. \quad (1d)
\]

The objective function (1a) maximizes \( Z_f \), the expected minimum fill rate. For the purposes of maximizing the expected minimum fill rate across all routes, objective function (1a) attributes to each node the fill rate of the route. One can linearize function (1a) with the following constraint:

\[
Z_f \leq \sum_{r \in R} f_r \delta_{ir} y_r, \quad \forall i \in N. \quad (2)
\]

Constraints (1b) ensure that all nodes (agencies and donors) are assigned to a route. Constraint (1c) selects \( K \) routes, ensuring utilization of all vehicles. Constraints (1d) define the binary route selection variables.

3.3. Equity-based resource allocation and sequencing policies

We define a cluster of nodes \( c \) in the set of all feasible clusters \( C \) as the set of donors \( M_c \) and agencies \( N_c \) visited by a vehicle. Only one optimal route (sequence and allocation policy) for each cluster is necessary in the set \( R \). Finding the optimal route for a cluster requires the analysis of the optimal sequencing and allocation for that cluster. Let \( D_i \) denote the random variable for demand at agency \( i \in N \). Let \( s_i \) denote the supply upon arrival at agency \( i \in N \); the initial supply \( s_0 = s_1 = \sum_{j \in M} q_j \). Although \( s_i \) is cluster-specific, we omit cluster-specific superscripts for ease of notation. The following steps are applied to obtain the optimal sequencing and resource allocation solutions for a given cluster of agencies and donors.

1. Enumeration of sequences. Sequences are feasible if their corresponding length satisfies the route length constraints. All feasible agency sequences are enumerated. Since all donors are visited before agencies and all donations are accepted, donor sequencing does not impact the MSRA-e objective function, and a single donor sequence is sufficient.

2. Evaluation of the resource allocation policy. For each agency sequence, we solve for the optimal equitable resource allocation policy. Given a sequence of agencies, \( 1 \rightarrow 2 \rightarrow \cdots \rightarrow N \), where \( N = |N| \), Lien et al. (2013) characterize the structure of the optimal allocation policy using the following stochastic dynamic programming model: suppose that upon service completion at agency \( i \rightarrow 1 \), the fill rates for agencies 1 to \( i - 1 \) are \( \beta_1, \beta_2, \ldots, \beta_{i-1} \). Let \( \beta_{i-1}^{\min} = \min(\beta_1, \beta_2, \ldots, \beta_{i-1}) \). The optimal allocation amount, \( x_i \), at agency \( i \), given supply \( s_i \) and observed demand \( d_i \), is

\[
z^{(i)}(s_i, \beta_{i-1}^{\min}, d_i) = \max_{x_i} \mathbb{E}_{d_{i+1}} 
\]

\[
1 \wedge \beta_{i-1}^{\min} \wedge \frac{x_i}{d_i} \wedge z^{(i+1)}(s_i - x_i, \beta_{i+1}^{\min}, d_{i+1}),
\]

where \( a \wedge b = \min\{a, b\} \) and \( z^{(i)}(s_i, \beta_{i-1}^{\min}, d_i) \) is the optimal expected minimum fill rate for the sequence. At the last agency in the sequence, the optimal allocation is \( x_{\max}^c = s_{\max} \wedge d_{\max} \), and

\[
z^{(N)}(s_{\max}, \beta_{\max}^{\min}, d_{\max}) = 1 \wedge \beta_{\max}^{\min} \wedge \frac{s_{\max} \wedge d_{\max}}{d_{\max}}.
\]

To find the optimal inventory allocation for the sequence, solve for \( z^{(1)}(s_1, \beta_0^{\min}, d_1) \), where \( \beta_0^{\min} = 1 \). The optimal expected minimum fill rate of the sequence, \( f_r \), is

\[
f_r = \mathbb{E}_{D_1} \left[ z^{(1)}(s_1, 1, d_1) \right]. \quad (5)
\]

3. Selection of the best sequence. The sequence with the highest expected minimum fill rate is used in the set-partitioning formulation for the MSRA-e.

3.4. Computational effort

For a given cluster, solving the MSRA-e requires the enumeration of all feasible agency sequences and evaluation of a complex stochastic dynamic program for each such sequence. We show how this impacts the computational effort required to solve the problem.

Enumerating the clusters and sequences: For a problem instance with \( N = |N| \) agencies and \( K \) vehicles,
Multi-vehicle allocations in nonprofit organizations

$N \geq K$, the maximum number of clusterings for agencies is $C_N = C(N, 1) + \cdots + C(N, N - K + 1)$, where $C(a, b)$ represents the total number of $b$-element subsets of a set with $a$ elements. Similarly, given $M = |M| \geq K$ donors, the number of donor combinations is $C_M = C(M, 1) + \cdots + C(M, M - K + 1)$. The number of clusters to enumerate is $C_N \times C_M$. For each cluster with $N_c > 1$, we enumerate all $N_c!$ feasible agency sequences.

Evaluating resource allocation policy: The number of computations to solve the resource allocation problem is dependent on the dimension of the discrete state space. We use a discrete distribution to characterize the demand and discretize the incoming supply and minimum fill rate values to evaluate $z_i^0(s_i, \beta_{\text{min}}^{-1}, d_i)$ at each agency $i$. Let $l_s$, $l_d$, and $l_f$ represent the number of discrete intervals for supply, fill rate and demands, respectively. The number of discrete intervals for resource allocation, $l_{s_0}$, is dependent on supply, demand, and fill rate. If $s_i \geq d_i$, then $x_i$ is only constrained by the minimum fill rate; $l_{s_0} = l_d \beta_{\text{min}}^{-1}$. If $s_i < d_i$, then $x_i$ is also limited by the supply amount, and hence $l_{s_0} = l_d (\beta_{\text{min}}^{-1} \times \frac{1}{\beta})$. For a sequence with $N_c$ agencies, there are $l_s l_d l_f$ states for nodes 2 through $N_c$. Although one computation is necessary to find the optimal allocation at the last node since $s_{N_c} = s_{N_c} \wedge d_{N_c}$, one needs to make $l_{s_0}$ computations to determine the optimal allocation at nodes 2 to $N_c - 1$. There are $l_d$ states for the first node in the sequence as $s_1 = s_0$ and $\beta_{\text{min}}^0 = 1$, and each state requires $l_{s_0}$ computations to determine the optimal allocation. The number of evaluations to find the expected minimum fill rate, $f_r$, with $N_c$ agencies is $l_{s_0} l_{d_0} + \sum_{i=1}^{N_c-1} (l_s l_d l_f d_{i-1}) + l_s l_d l_f d_0$.

Solving the set partitioning model: ILOG CPLEX 11 is used to solve the set partitioning model with a 2.4 GHz 64-bit (4 MB L2 cache) CPU machine with 8 GB of RAM.

Combined effort: Clearly the major computational effort consists in evaluating all feasible sequences. Consider an instance with six agencies, six donors, and two vehicles. The resulting 3844 clusters yield 76,632 sequences, requiring 81 hours to evaluate the equity-based resource allocation policy and obtain the optimal expected minimum fill rates. The set-partitioning formulation is solved within a second by CPLEX. The computational effort required to evaluate the equity-based resource allocation policy for each enumerated route limits the exact MSRA-e solutions to small instances.

4. Numerical study

This section examines the MSRA-e to develop a deeper understanding of the problem behavior and provide insights based on optimal solutions to a set of test instances. While Model (1) allows for general constraints on route length by node service time and travel length, our analysis in Sections 4 and 5 focuses on the unconstrained case to isolate the effects of the objective function in route composition and allocation decisions. In Section 6, we consider the time-constrained case.

4.1. Test instances

Table 1 summarizes the 3537 test instances; detailed parameters for instances are available from the authors. The number of nodes ranges from six to 15; the smallest instance has four agencies and two donors, whereas the largest instance has seven agencies and eight donors. These represent the largest instances that can be solved optimally with a reasonable computational effort without a travel time constraint. In Section 6, we can consider larger instances with the travel time constraint imposed.

Demands for agencies are assumed to be independent and are observed from a discrete distribution, created based on a gamma distribution with 20 discretization intervals. Similar to Lien et al. (2013), we use a gamma distribution to ensure positive demand values in settings with low and high demand variability. Three levels of agency demand means are considered: low (25), medium (50), and high (75). Each agency has either a low (0.5) or high (1.5) Coefficient of Variation (CV). Demand distributions among agencies are identical, partially identical, or different. Identically distributed demands limit the number of sequences to evaluate and allow the consideration of larger instances in terms of the number of agencies.

Three total supply levels are considered: scarce, equal, and ample, relative to total demand. In instances with scarce supplies, total supply is 75% of the total demand means, whereas in the settings with ample supplies, total supply is 25% larger than the total demand means. Distribution of supply among donors is identical, wide range, or narrow range. In the narrow range, half of the donors have about 50% more supply than remaining donors, whereas in the

| Number of agencies | Number of donors | Number of vehicles | Demand distributions (Mean, CV) | Similarity of agency demands | Total supply level | Donor supplies |
|--------------------|------------------|-------------------|-------------------------------|-----------------------------|-----------------|--------------|
| 4–7                | 2–8              | 2–7               | (25, 0.5) (25, 1.5)           | Identical                   | Scarce          | Identical    |
|                    |                  |                   | (50, 0.5) (50, 1.5)           | Partially identical         | Equal           | Narrow range |
|                    |                  |                   | (75, 0.5) (75, 1.5)           | Different                   | Ample           | Wide range   |
The formulation of the MSRA-w and the waste-based re-allocation explicitly reduces waste while achieving equity. We use the MSRA-w as a benchmark, which optimizes clustering, sequencing, and allocation decisions to minimize total expected waste in the system. Solving the MSRA-w requires much lower computational effort than the MSRA-e, since expected waste can be easily computed based on total supply and expected demand in a cluster. The MSRA-e implicitly considers waste while achieving equity by raising the fill rate of all agencies and thereby increasing expected demand (which reduces waste). The MSRA-w greedily allocates resources to the first agencies visited to minimize waste, which can lead to wide ranges in fill rates along the route. Furthermore, MSRA-w routes become more equitable as supply level is increased. For instance, the average optimality gaps of MSRA-w in terms of the expected minimum fill rate are between 12.9 and 21.1% when supply is scarce, whereas the average optimality gaps range from 4.4 to 8.2% when supply is ample. This is intuitive as the differences between MSRA-e and MSRA-w allocations may decrease in situations in which supplies are abundant. Effects of supply availability in MSRA-e are further analyzed in Section 4.2.3.

4.2.1. Equity and waste levels

The MSRA-e implicitly reduces waste while maximizing equity. We use the MSRA-w as a benchmark, which optimizes clustering, sequencing, and allocation decisions to minimize total expected waste in the system. Solving the MSRA-w requires much lower computational effort than the MSRA-e, since expected waste can be easily computed based on total supply and expected demand in a cluster. The formulation of the MSRA-e and the waste-based resource allocation and sequencing policies are explained in Appendix A.

Let \( Z^*_F \) and \( \hat{w} \) represent the expected minimum fill rate and expected waste from the optimal MSRA-e solution, respectively, and let \( \hat{f} \) and \( \hat{Z}^*_w \) represent the expected minimum fill rate and expected waste from the optimal MSRA-w solution, respectively. Table 2 reports the average and maximum optimality gaps in expected minimum fill rate (i.e., \( Z^*_F - \hat{f} \)) and expected waste (i.e., \( \hat{w} - \hat{Z}^*_w \)) by supply level.

**Observation 1:** Optimal MSRA-e routes achieve near-optimal levels of expected waste, whereas optimal MSRA-w routes result in larger optimality gaps in expected minimum fill rate.

As shown in Table 2, MSRA-e results in, on average, 1.3 to 2.8% more expected waste than the optimal expected waste found by the MSRA-w. The average optimality gaps of MSRA-w in terms of the expected minimum fill rate are between 4.4 and 21.1%. The optimality gaps of the MSRA-w in expected minimum fill rates increase at a higher rate with the number of agencies compared with optimality gaps in the expected waste found by the MSRA-e. Indeed, for the instances with seven agencies, the maximum optimality gap in expected minimum fill rate becomes 49.0% for MSRA-w, whereas all MSRA-e solutions perform within 9.7% of the optimal expected wastes. The results are intuitive as the MSRA-e implicitly considers waste while achieving equity by raising the fill rate of all agencies and thereby increasing distribution (which reduces waste). The MSRA-w greedily allocates resources to the first agencies visited to minimize waste, which can lead to wide ranges in fill rates along the route. Furthermore, MSRA-w routes become more equitable as supply level is increased. For instance, the average optimality gaps of MSRA-w in terms of the expected minimum fill rate are between 12.9 and 21.1% when supply is scarce, whereas the average optimality gaps range from 4.4 to 8.2% when supply is ample. This is intuitive as the differences between MSRA-e and MSRA-w allocations may decrease in situations in which supplies are abundant. Effects of supply availability in MSRA-e are further analyzed in Section 4.2.3.

**Table 2. Comparison of MSRA-e and MSRA-w in terms of equity and waste by agency count**

| Supply Level | Ave. (%) | Max. (%) |
|--------------|----------|----------|
|               |          |          |
| Scarce supply | \( \hat{w} - \hat{Z}^*_w \) | 1.6 | 5.7 |
|               | \( Z^*_F - \hat{f} \) | 12.9 | 32.2 |
| Equal supply  | \( \hat{w} - \hat{Z}^*_w \) | 1.7 | 7.0 |
|               | \( Z^*_F - \hat{f} \) | 7.4 | 17.8 |
| Ample supply  | \( \hat{w} - \hat{Z}^*_w \) | 1.3 | 6.4 |
|               | \( Z^*_F - \hat{f} \) | 4.4 | 11.6 |

4.2.2. Cluster size and composition

Lien et al. (2013) find that equity-maximizing solutions sequence agencies with highly variable demand early on a route since (i) demand at the first agency visited is known before making allocation decisions and (ii) demand variability is reduced for later decision epochs. Recall that supply is allocated at each agency considering the demand expectations of the unvisited agencies along a route. For a given supply, as the number of agencies along a route increases, the expected minimum fill rate tends to deteriorate due to the increase in the number of decision epochs and sources of variability. We show how the choice of objective function impacts cluster structures.

The MSRA-e and MSRA-w often lead to different clusters: 70% of test instances report different clustering solutions. In general, the clusters are not different in instances with identical agencies or instances with many vehicles and few donors. In such cases, few alternative clustering solutions exist. Clusters for the other instances differ in terms of size and/or composition.

**Observation 2:** In general, for a given number of vehicles and agencies, the largest cluster in the MSRA-e solution is smaller than that in the MSRA-w.
Table 3. Comparison of cluster size in MSRA-e and MSRA-w

| Number of agencies | N = 4 | N = 5 | N = 6 | N = 7 |
|--------------------|------|------|------|------|
| Average $\max c$(MSRA-e) | 4.3  | 5.2  | 5.0  | 6.0  |
| Average $\max c$(MSRA-w) | 5.1  | 5.8  | 6.1  | 6.7  |
| Percentage of instances with $\max c$(MSRA-e) $>$ $\max c$(MSRA-w) | 4.6% | 6.8% | 6.9% | 11.1% |
| Percentage of instances with $\max c$(MSRA-w) $>$ $\max c$(MSRA-e) | 40.4% | 37.6% | 47.0% | 32.2% |

Table 3 presents an analysis of cluster size for the MSRA-e and the MSRA-w. Let $\max c$(MSRA-e) and $\max c$(MSRA-w) represent the maximum cluster size in terms of number of donors and agencies in a solution of MSRA-e and MSRA-w, respectively. The table compares the average values of the $\max c$(MSRA-e) and $\max c$(MSRA-w) over all test instances and the percentage of instances in which the MSRA-e or MSRA-w resulted in the largest cluster.

The average size of the largest clusters in the MSRA-w is larger than those obtained for the MSRA-e. Although there are instances in which the MSRA-e results in larger clusters, the maximum cluster size is larger in solutions of the MSRA-w for a higher proportion of instances. These results are consistent with the findings in Lien et al. (2013) that smaller clusters with fewer decision epochs tend to perform better in maximizing the expected minimum fill rates compared with larger clusters. Furthermore, demand at the first agency is observed before any decisions are made; therefore, the variability of this agency does not impact allocation decisions. Alternatively, the MSRA-w tends to create larger clusters to aggregate demand and supply. Since the waste-based allocation policy does not reserve supply for the remaining agencies in the cluster, increasing the number of demands and supplies in a cluster does not necessarily affect the expected waste negatively. Indeed, the MSRA-w tends to benefit from risk pooling from agencies and supply pooling from donors by assigning more nodes to the same cluster to decrease the likelihood of total demand realized falling short of supplies. Under the assumption of uncapacitated vehicles, without the constraints forcing the utilization of all vehicles, the optimal MSRA-w solution would always use a single cluster.

Observation 3: In general, the MSRA-e tends to spread demand variability among clusters, whereas the MSRA-w assigns the agencies with high demand variability to the same cluster.

The MSRA-w solutions exploit risk pooling further by clustering agencies with highly variable demand together. Conversely, the MSRA-e solutions can benefit from distributing agencies with high demand variability among different clusters and placing these agencies first in the visit sequence. Therefore, MSRA-e generally avoids clustering the high CV agencies together and spreads the variability among different routes, which is against the practice of risk pooling.

Given Observations 2 and 3, one may prefer many short routes over fewer long routes to maximize equity due to benefits of risk spreading; however, increasing the number of routes also divides donor supplies among more vehicles and might cause inefficiencies due to potential effects of partitioning. In the following subsection, we examine the impact of supply availability and demand variability on inventory pooling/spreading in MSRA-e.

4.2.3. Effects of supply availability and demand variability

To isolate the impacts of supply availability and demand variability in MSRA-e, we analyze a set of test cases designed to highlight these impacts. The 864 scenarios are summarized in Table 4.

To investigate the benefits of inventory pooling/spreading, we compare one- and two-vehicle MSRA-e solutions for the scenarios. Let $\Delta_{2-1}$ denote the percentage difference in the expected minimum fill rates for
Table 5. Change (in percentage) in $\Delta_{2-1}$ values with respect to increased supplies

| Number of agencies | Number of donors | Average change in $\Delta_{2-1}$ | Maximum change in $\Delta_{2-1}$ | % of instances with a negative change in $\Delta_{2-1}$ |
|--------------------|-----------------|---------------------------------|---------------------------------|----------------------------------|
|                    |                 | Scarce $\rightarrow$ equal       | Scarce $\rightarrow$ equal      | Scarce $\rightarrow$ Equal $\rightarrow$ ample |
| 2                  | 2               | $-3.1$                          | $-7.1$                          | 92.6                              |
| 2                  | 4               | $-4.0$                          | $-7.0$                          | 100                               |
| 2                  | 6               | $-4.0$                          | $-6.0$                          | 100                               |
| 4                  | 2               | $-2.6$                          | $-3.8$                          | 96.3                              |
| 4                  | 4               | $-3.2$                          | $-3.8$                          | 100                               |
| 4                  | 6               | $-2.8$                          | $-3.8$                          | 96.3                              |
| 6                  | 2               | $-2.3$                          | $-5.0$                          | 100                               |
| 6                  | 4               | $-2.7$                          | $-5.0$                          | 100                               |
| 6                  | 6               | $-0.5$                          | $-8.1$                          | 69.0                              |

one- and two-vehicle solutions. If the difference is positive, the two-vehicle solution is better than the one-vehicle solution in terms of the expected minimum fill rate and hence inventory spreading is more beneficial. A negative difference corresponds to the case in which using one vehicle performs better and hence inventory pooling is more beneficial. We find that two-vehicle solutions lead to more equitable solutions for 76% of the instances. We examine the changes in $\Delta_{2-1}$ values with respect to supply availability and demand variability in more detail. Additional experiments to compare one- and three-vehicle solutions reveal that the trends described in Observations 4 and 5 for $\Delta_{2-1}$ hold for $\Delta_{3-1}$ as well.

Supply availability.

We examine the one-vehicle and two-vehicle MSRA-e solutions as supply availability increases: (i) from scarce to equal and (ii) from equal to ample. Table 5 presents the average and maximum changes in $\Delta_{2-1}$ values and the percentage of instances with a negative change in $\Delta_{2-1}$ values as supply level is increased gradually, aggregated over all demand distributions for a fixed number of agencies and donors. A negative change in Table 5 implies a decrease in $\Delta_{2-1}$ when supply increases from scarce to equal or from equal to ample. We find that 91% of instances with scarce supplies have a positive $\Delta_{2-1}$ value, whereas the percentages of positive $\Delta_{2-1}$ values decrease to 85 and 52% as the supply availability increases from scarce to equal and from equal to ample, respectively. Figure 2 further shows the absolute $\Delta_{2-1}$ values and the trends with increased supply levels for a subset of instances with identical donor supplies.

Observation 4: In general, the benefits of inventory pooling in MSRA-e increase as supply increases.

As an example, consider the instances with six agencies and four donors. For these instances, $\Delta_{2-1}$ values decrease as the supply level is gradually increased. Specifically, as the total supply increases from scarce to equal and from equal to ample, the average changes in $\Delta_{2-1}$ values are $-2.7\%$ and $-2.4\%$ and the maximum changes are $-5.0\%$ and $-3.7\%$, respectively. Serving the nodes along a single route becomes more attractive with increased supply levels.

As the number of agencies along a route increases, the number of allocation decisions and the sources of variability increase along the route and thus shorter routes may be beneficial in the MSRA-e. However, in an extreme case when supplies are large (total supply is 1000 units for a four-agency instance with average demand of 10 at each agency),

![Fig. 2. The trends in $\Delta_{2-1}$ values with respect to supply availability (four identical donors).](image)
having fewer agencies on multiple short routes may not improve the objective function. Since supplies are abundant, the number of agencies along a route does not impact allocation decisions. As supply becomes scarce, partitioning the nodes among multiple routes may become useful if the benefit of fewer agencies (fewer decision epochs) outweighs the benefits of pooling supply along a single route.

### Demand variability.

As Fig. 2 shows, the benefits of inventory pooling in MSRA-e is larger for instances involving low CV agencies. We further examine the one-vehicle and two-vehicle MSRA-e solutions relative to demand variability. Table 6 presents the average change in $\Delta_{2-1}$ values for instances with identical demand CVs, when demand variabilities are increased from low CV to high CV. Focusing instances with identical donor supplies, Fig. 3 illustrates the absolute $\Delta_{2-1}$ values and the trends with increased demand CVs in more detail. From the results, we make the following observation.

**Observation 5:** In general, the benefits of inventory pooling in MSRA-e decrease as the demand variability increases.

As shown in Table 6, as agency demand CVs are increased from low to high, the change in $\Delta_{2-1}$ values is positive in almost all cases. The results show that the benefit of using two vehicles compared to one vehicle increases as the demand variability increases. As discussed in Observation 3, distributing agencies with high demand variabilities among different clusters tends to improve equity; as agency demands are less predictable, it becomes more difficult to make reliable resource allocation decisions along a single route and using multiple routes may become more desirable. However, if all agency demands are known beforehand, serving the agencies along multiple routes does not have much effect on the equity objective. Figure 3 also shows that the benefits of inventory pooling is generally larger for the instances with ample supplies, as one would expect given the Observation 4.

These observations suggest that the reasons for using multiple routes in a nonprofit setting can be different than those in commercial settings. In the traditional cost-based Vehicle Routing Problem, under the triangular inequality, a single route is more efficient if the vehicle capacity is not limiting. In the MSRA-w, a single route minimizes the expected waste. Even without travel time restrictions, a single route in the MSRA-e may decrease equity.

#### 4.2.4. Robustness to supply variation

To understand the impact of supply variation on MSRA-e, we conduct additional experiments, summarized in Table 7. The data set consists of 42 base instances for a network with six agencies and four donors visited by two vehicles and 693 scenarios generated by increasing and/or decreasing the base supply amounts at a subset of donors by 10%. In each scenario, we raise or lower the supply level(s) of one, two, or three donor(s) by 10% from the base supply level(s).

![Fig. 3. The trends in $\Delta_{2-1}$ values with respect to demand variability (four identical donors).](image)

---

**Table 6. Change (in percentage) in $\Delta_{2-1}$ values with respect to increased demand variability**

| Number of agencies | Demand means | Number of donors = 2 | | Number of donors = 4 | | Number of donors = 6 |
|-------------------|--------------|----------------------|--------------|----------------------|--------------|----------------------|
|                   |              | Scarc | Equal | Ample |              | Scarc | Equal | Ample |              | Scarc | Equal | Ample |
| 2                 | 150          | 4.7   | 5.9   | 5.4   | 0.4           | 1.0   | 2.7   | 2.7   | 1.4           | 1.7   | 3.0   |
|                   | 100, 200     | 2.7   | 3.5   | 3.9   | 2.8           | 3.5   | 3.9   | 2.8   | 0.5           | 0.6   | 2.5   |
| 4                 | 75           | 5.1   | 4.9   | 4.1   | 3.2           | 2.2   | 2.8   | 2.8   | 2.2           | 2.2   | 5.5   |
|                   | 50, 100      | 4.0   | 3.2   | 2.5   | 3.4           | 2.2   | 2.6   | 2.8   | 2.2           | 2.1   | 2.1   |
| 6                 | 50           | 4.5   | 2.9   | 2.6   | 5.3           | 2.5   | 2.4   | 4.9   | 6.6           | 2.0   | 2.0   |
|                   | 25, 75       | 3.6   | 2.3   | 2.3   | 3.4           | 2.4   | 2.4   | 3.9   | 0.8           | 2.5   | 2.5   |
|                   | 25, 50, 75   | 4.4   | 2.6   | 2.4   | 4.5           | 2.5   | 2.3   | 4.7   | 2.7           | 2.5   | 2.5   |
Since all donations are collected before any food is allocated, given a fixed cluster, one can make sequencing and allocation decisions based on actual supplies collected. Therefore, variations in donation levels do not affect sequencing and allocation decisions; however, supply variation can impact clustering. To study the impact of variability in supply on the performance of the clusters, we compare the optimal solution values under actual supply levels, which deviate in the 10% range, and the best solution values that can be achieved with the clusters designed based on average supply levels. Let \( z'' \) denote the optimal expected minimum fill rate found based on actual supplies, and \( w'' \) is the corresponding percentage of expected waste with respect to total supply. Let \( z' \) represent the maximum expected minimum fill rate that can be achieved by using actual supplies in clusters designed according to average supply levels and \( w' \) denote the corresponding expected waste percentage. Then \( z'' - z' \) and \( w' - w'' \) values represent the losses in expected minimum fill rates and expected waste percentages due to the deterministic supply assumption. Average and maximum \( z'' - z' \) and \( w' - w'' \) values are presented in Table 8.

The results in Table 8 indicate that although supply variation may lead to different clusterings, the average impact of supply variation on equity and waste levels is small and the MSRA-e solutions are robust to supply fluctuations of 10%. Nevertheless, improving communication with donors to access to perfect supply information in advance or incorporating supply uncertainty explicitly in route design could improve the FRP service performance.

To achieve equitable solutions with low waste, the MSRA-e creates many short routes and distributes demand variability among the routes. We now use these insights to develop an efficient heuristic for the MSRA-e next.

### 5. A decomposition-based heuristic

The Decomposition-Based heuristic (DBH) decomposes the problem into three phases: clustering, sequencing, and allocation, and builds on existing heuristics for sequencing and allocation from Lien et al. (2013). The clustering phase partitions donors and agencies to satisfy a proxy objective for equity. The sequencing phase generates a feasible sequence in each cluster. The allocation phase allocates resources along the cluster routes. The initial solution is improved via a set of moves to change the clusters. Given that each move requires a re-evaluation of allocation policies among impacted clusters, a solution algorithm based on an extensive search strategy is not reasonable. Our approach focuses on obtaining a high-quality initial solution and applying a limited number of moves to achieve a near-optimal solution.

#### 5.1. Constructing an initial solution

The three phases are described in detail next.

**Phase 1. Clustering.**

The clustering phase assigns donors and agencies to vehicles based on agencies’ demand distributions without explicitly considering sequencing or allocation. Since demand variability affects cluster sizes/compositions and resulting fill rates, we introduce a multi-objective clustering model. The first objective, \( g(F) \), maximizes \( F \), the expected minimum fill rate across all agencies. This is a proxy for the true expected minimum fill rate since we do not solve the sequencing and allocation problems in this phase. The second objective, \( g(V) \), minimizes \( V \), the maximum total demand variance assigned to a cluster across all clusters; that is, if \( \sigma_i^2 \) denotes the demand variance at agency \( i \), \( V = \max_{i \in C} (\sum_{j \in N} \sigma_j^2) \). This objective is motivated by the tendency of the MSRA-e to spread the high demand variability among different clusters, as discussed in Observation 3. We introduce weights for each objective, \( \lambda_V \) and \( \lambda_F \) (\( \lambda_V + \lambda_F = 1 \)), which represent the weights for variance, \( g(V) \), and fill rate, \( g(F) \), objectives, respectively. We define the following parameters and variables:

- \( \mu_i \): expected demand at agency \( i \in N \) (i.e., \( \mu_i = \mathbb{E}(D_i) \forall i \in N \));
- \( \delta_i \): if node \( i \in N \cup M \) is assigned to cluster \( c \in C \); 0, otherwise;
- \( p_{ic} \): proportion of demand of agency \( i \in N \) satisfied by cluster \( c \in C \).

The clustering formulation is as follows:

\[
\min \lambda_V g(V) - \lambda_F g(F), \quad (6a)
\]
subject to

\[ F \leq \sum_{c \in C} p_{ic}, \quad \forall i \in \mathcal{N}, \]  

(6b)

\[ V \geq \sum_{i \in \mathcal{N}} \sigma_i^2 \delta_i, \quad \forall c \in \mathcal{C}, \]  

(6c)

\[ p_{ic} \leq \delta_i, \quad \forall i \in \mathcal{N}, \forall c \in \mathcal{C}, \]  

(6d)

\[ \sum_{i \in \mathcal{N}} \mu_i p_{ic} \leq \sum_{j \in \mathcal{M}} q_j \delta_{jc}, \quad \forall c \in \mathcal{C}, \]  

(6e)

\[ \sum_{i \in \mathcal{C}} \delta_i = 1, \quad \forall i \in \mathcal{N} \cup \mathcal{M}, \]  

(6f)

\[ \sum_{i \in \mathcal{N} \cup \mathcal{M}} \delta_{ic} \geq 1, \quad \forall c \in \mathcal{C}, \]  

(6g)

\[ \delta_{ic} \in \{0, 1\}, \quad \forall i \in \mathcal{N} \cup \mathcal{M}, \forall c \in \mathcal{C}, \]  

(6h)

\[ 0 \leq p_{ic} \leq 1, \quad \forall i \in \mathcal{N}, \forall c \in \mathcal{C}. \]  

(6i)

The weighted objective function (6a) minimizes the maximum total demand variance in a cluster among all clusters and maximizes the minimum fill rate across network agencies. Constraints (6b) define the minimum fill rate. Constraints (6c) determine the maximum demand variance accumulated in a cluster among all clusters. Constraints (6d) guarantee that the demand of an agency is met by its assigned cluster. Constraints (6e) limit the total demand satisfied in a cluster by the total cluster supply. Constraints (6f) ensure that each agency and donor is assigned to one cluster. Constraints (6g) assign at least one agency and one donor to each cluster, ensuring the utilization of all vehicles. Finally, Constraints (6h) define the binary assignment variables and (6i) define the continuous demand satisfaction variables.

Formulation of the objective function: Since \( F \) and \( V \) differ in magnitude, we scale them to obtain the dimensionless objectives, \( g(F) \) and \( g(V) \), through normalization; see Marler and Arora (2004) for a discussion. To normalize, lower bounds and upper bounds are developed for the values of \( F \) and \( V \). Let \( F_{LB} \) and \( F_{UB} \) represent lower and upper bounds for the variable \( F \) and \( V_{LB} \) and \( V_{UB} \) represent lower and upper bounds for the variable \( V \). Let \( \mathcal{N}_c \) denote the subset of \( \mathcal{N} \) consisting of the agencies with \( N - K + 1 \) largest demand variances, and let \( \mathcal{N}_M \) denote the subset of \( \mathcal{N} \) consisting of the agencies with \( N - K + 1 \) largest demand means. The bounds for \( F \) and \( V \) are

\[ F_{LB} = \min_{j \in \mathcal{M}} \left( \frac{q_j}{\sum_{i \in \mathcal{N}_c} \mu_i} \right), \]  

(7a)

\[ F_{UB} = \min \left( 1, \sum_{j \in \mathcal{M}} q_j / \sum_{i \in \mathcal{N}_c} \mu_i \right), \]  

(7b)

\[ V_{LB} = \sum_{i \in \mathcal{N}_c} \sigma_i^2, \quad V_{UB} = \sum_{i \in \mathcal{N}_c^2}. \]  

(7b)

Given \( K \) clusters, \( F_{LB} \) is calculated by dividing the minimum donation quantity by the largest total demand that can be assigned to a cluster. The upper bound for the expected minimum fill rate \( F_{UB} \) is set to the minimum of one and the ratio of total supply to total demand. The lower bound for the maximum total demand variance assigned to a cluster \( V_{LB} \) is simply equal to the largest demand variance across the agencies, while \( V_{UB} \) is the largest sum of variances that can be assigned to a cluster. Using these bounds, we obtain the scaled objectives:

\[ g(F) = \frac{F - F_{LB}}{F_{UB} - F_{LB}}, \quad g(V) = \frac{V - V_{LB}}{V_{UB} - V_{LB}}. \]  

(8)
If the lower and upper bounds are equal, we set \( g(F) = F/F_{UB} \) and \( g(V) = V/V_{UB} \).

**Phase 2. Sequencing.**

We use the sequencing heuristic of Lien et al. (2013) that orders agencies to maximize the expected minimum fill rate along the route. Agencies are ordered in decreasing CV values; if the CV values of agencies are identical, those agencies are ordered according to decreasing standard deviation values (i.e., the agency with the larger mean is visited first).

**Phase 3. Allocation.**

We incorporate the Two-Node Decomposition (TND) heuristic developed by Lien et al. (2013) to solve the equitable resource allocation problem along a route; see Appendix B. Given route \( r \) with agencies sequenced \( 1 \rightarrow 2 \rightarrow \cdots \rightarrow N_r \) in cluster \( c \), the TND heuristic decomposes the problem into sequential two-node subproblems. Upon visiting an agency, supply is divided: one portion for serving the current two-node subsequence and another reserved for serving the remaining nodes. Supplies are allocated among the two nodes based on the demand observed at the first node and the fill rates at the previously visited nodes.

The TND heuristic requires two computations at each node: (i) reserving supply and (ii) allocating supply. Although the TND heuristic determines the allocation policy efficiently, one still needs to simulate over all possible demand realizations at each agency to calculate \( f_r \). The computational effort increases with the number of agencies and the number of discretization intervals for demands.

**The procedure.**

Our heuristic performs as follows. The clustering phase is performed \( \theta \) times, with each repetition consisting of a unique combination of the weights of \( \lambda_Y \) and \( \lambda_F \). Each solution to the clustering model is evaluated with the sequencing and allocation heuristics. The solution with the largest expected minimum fill rate is selected for continues improvement.

**5.2. Improving the initial solution**

The improvement step involves a series of node reassignments. Each reassignment is followed by the sequencing and allocation heuristics from the construction stage. Since each move requires re-evaluating the allocation policy to assess the expected minimum fill rate of the modified routes, only a limited number of moves is considered. In each iteration, two types of moves are applied sequentially: (i) agency moves (agency removal/insertion and exchanges) and (ii) donor moves (donor removal/insertion and exchanges). Any move that improves the MSRA-e objective is accepted.

Each iteration begins by selecting two clusters, one with the lowest expected minimum fill rate (the worst-performing cluster) and one with the highest expected minimum fill rate (the best-performing cluster). Each move attempts an inter-cluster node reassignment to improve the expected minimum fill rate for the worst-performing cluster. If an improving solution cannot be found for the two clusters, the cluster with the next highest expected minimum fill rate is evaluated for reassignment. We limit the number of clusters considered for inter-cluster moves with the worst-performing cluster by a parameter \( P \), which we set to \( |C| - 1 \) in our numerical studies. Moves are performed in the order listed, until an improving solution is found. The algorithm stops if no improving solution is found or the maximum iteration limit is reached.

**Agency moves.**

Agency moves make changes to the demand composition of the worst-performing cluster to improve its expected minimum fill rate.

**Agency removal/insertion:** An agency is removed from the worst-performing cluster and inserted into the best-performing cluster. This move is attempted only if the worst-performing cluster includes at least two agencies. Two node selection strategies are applied: (i) select the agency with the minimum demand mean and (ii) select the agency with the minimum standard deviation. Only the agencies with the smallest means and/or standard deviations are considered for removal and insertion, since it is more likely that insertion of large demands decrease the objective of the pairing cluster below the current objective.

The selected agency is inserted into the pairing cluster to a position determined according to the sequencing heuristic. The resource allocation problem is evaluated in the two clusters. If the move improves the current MSRA-e objective value, it is accepted and next iteration proceeds. We use the second node selection strategy only if the first does not result in an improving solution and the agency with the minimum demand mean and the minimum standard deviation is not the agency with the minimum demand mean. At most two removal/insertion attempts are made for a pair of clusters.

**Agency exchange:** Two agencies are selected from the corresponding clusters: (i) select the pair of agencies with the smallest mean difference across all node pairs and (ii) select the pair of agencies with the smallest standard deviation difference across all node pairs. These strategies focus on decreasing the demand in the worst-performing cluster. The second selection strategy is applied if the move attempt with the first strategy does not lead to an improving solution and a different pair of agencies exists under the second node selection strategy. The insertion positions of the agencies are determined based on the sequencing heuristic and the resource allocation problem is solved in the modified clusters. The new solution is accepted if improving; otherwise, the nodes remain in their original clusters.

**Donor moves.**

Donor moves attempt to improve the objective by reassigning donors among the worst- and best-performing clusters.

**Donor removal/insertion:** A donor is removed from the current best-performing cluster and inserted in the
worst-performing cluster. The donor with the smallest supply is chosen for removal. The resource allocation problem is resolved in the clusters affected by the move.

**Donor exchange:** We consider exchanging two donors between the worst-performing cluster and the current best-performing cluster. Among the candidate donor pairs for exchange, we select the two donors that lead to the smallest change in supply amounts in the clusters. Only one exchange attempt is considered for a pair of clusters.

### 5.3. Computational analysis of the DBH

This section evaluates the performance of the DBH to solve the MSRA-e.

#### 5.3.1. Solution quality and speed

We evaluate the performance of the DBH in terms of solution accuracy and time efficiency relative to the exact approach in Section 3. We present results for $\theta = 11$, with $\lambda_F$ ranging from zero to one in increments of 0.1 and $\lambda_V = 1 - \lambda_F$. We find that the performance of the DBH improves as $\theta$ increases, although the marginal benefit decreases quickly. The solution quality of the heuristics is evaluated based on the average and maximum absolute optimality gaps, shown in Table 9.

The average optimality gaps for the DBH solutions range from 0.8 to 1.6% and increase slowly with the number of agencies; only five of the 3537 instances result in gaps over 10%. The average solution time of the DBH across all instances is 12 (CPU) seconds, whereas the maximum solution time is 14 (CPU) minutes. Solving these small size instances optimally through enumeration can take days, as discussed in Section 3.4.

#### 5.3.2. Performance of the DBH initial solutions

To assess the value of the improvement phase, we evaluate the accuracy of the initial solutions obtained by the constructive step of the DBH. Table 10 presents the optimality gaps and solution times of the DBH initial solutions.

When the initial DBH solutions in Table 10 are compared to the final DBH solutions in Table 9, final solutions are about 0.5% better than the initial solutions in terms of the average optimality gaps. The real benefit of the improvement phase comes from addressing the maximum optimality gaps. The initial solutions can perform 10.4% worse than the final solutions. The percentage of instances that can be solved optimally or near-optimally (i.e., with gap \leq 2%) increases by about 9% with the improvement step. The comparison of solution times shows that the constructive step is more time-consuming than the improvement step. Hence, although the constructive step performs generally well in solving the MSRA-e, the improvement step is effective in finding better solutions.

#### 5.3.3. Performance of the DBH clustering solutions

The DBH incorporates the sequencing and allocation heuristics of Lien et al. (2013) for the single-route problem. In the MSRA-e, clustering decisions are particularly

---

**Table 9. Solution accuracy and times of the DBH**

| Number of agencies | $N = 4$ | $N = 5$ | $N = 6$ | $N = 7$ |
|--------------------|---------|---------|---------|---------|
| Average optimality gap (%) | 0.8 | 1.1 | 1.1 | 1.6 |
| Maximum optimality gap (%) | 8.0 | 9.1 | 8.8 | 11.4 |
| % of instances with optimality gap = 0 | 46.9 | 32.5 | 32.7 | 25.5 |
| % of instances with optimality gap \leq 0.02 | 86.3 | 82.5 | 79.6 | 73.1 |
| % of instances with optimality gap \geq 0.10 | 0.0 | 0.0 | 0.0 | 0.5 |
| Average solution time (CPU seconds) | 1.5 | 3.2 | 6.7 | 34.8 |
| Maximum solution time (CPU seconds) | 5.0 | 12.0 | 73.0 | 840.0 |

---

**Table 10. Solution accuracy and times of the initial DBH solutions**

| Number of agencies | $N = 4$ | $N = 5$ | $N = 6$ | $N = 7$ |
|--------------------|---------|---------|---------|---------|
| Average optimality gap (%) | 1.2 | 1.5 | 1.6 | 2.2 |
| Maximum optimality gap (%) | 11.3 | 11.4 | 11.6 | 18.7 |
| % of instances with optimality gap = 0 | 39.4 | 28.5 | 29.3 | 22.0 |
| % of instances with optimality gap \leq 0.02 | 78.1 | 73.6 | 71.3 | 63.6 |
| % of instances with optimality gap \geq 0.10 | 0.7 | 0.1 | 1.0 | 1.7 |
| Average solution time (CPU seconds) | 1.5 | 3.2 | 6.3 | 28.1 |
| Maximum solution time (CPU seconds) | 5.0 | 12.0 | 45.0 | 631.0 |
important in route design; hence, we are interested in the ability of the DBH to select near-optimal clusters. To evaluate the accuracy of the clustering solutions obtained by the DBH, instead of the heuristic of Lien et al. (2013), we use the optimal sequencing and allocation policies within clusters found by the DBH. We then find the optimality gap by comparing the resulting objective function value with that of the optimal. Table 11 reports the optimality gaps due to the clustering solutions obtained by the DBH relative to the optimal MSRA-e value.

When compared to the average DBH optimality gaps reported in Table 9 (ranging from 0.8 to 1.6%), the gaps due to clustering decisions are comparable with the gaps due to sequencing and allocation decisions. As shown in Table 11, the percentage of instances that the DBH results in an optimal clustering solution (with optimal sequencing and allocation policies) ranges from 67.4 to 77.4%. However, in some instances, the gap due to clustering can be relatively high; indeed, the maximum clustering optimality gaps are close to the maximum optimality gaps of the overall heuristic. These results indicate that clustering can be a critical determinant of the accuracy of the DBH and near-optimal clusters can be obtained by using the DBH algorithm.

5.3.4. Expected waste

The optimal MSRA-e solutions in Section 3 maximize equity while maintaining low levels of waste. We evaluate the expected waste in solutions obtained with the DBH relative to the values for the optimal MSRA-w and MSRA-e solutions. Let $H_w$ denote the expected waste of the DBH solution, while $Z^*_w$ and $\tilde{w}$ are the expected waste from MSRA-w and MSRA-e solutions, respectively. The results are presented in Table 12.

As observed from Table 12 (bottom), the gaps between the DBH and the optimal MSRA-e solutions in expected waste are smaller than the optimality gaps. For some cases, the average gap is negative; that is, using the DBH leads to smaller waste than solving the MSRA-e optimally. The small/negative gaps in expected waste indicate that although the focus is on the equity objective in the DBH, the algorithm does not considerably compromise on the waste objective in favor of equity objective.

6. MSRA-e with travel time constraints

In earlier sections, we isolate the impact of the objective function on route design by removing constraints on travel time/distance and obtain insights about clustering of donors and agencies in the FRP system. In some settings, network travel time may limit cluster options. We consider the constrained version of MSRA-e with travel time restrictions for each vehicle. Using the insights from our previous analysis, we adopt the DBH to incorporate travel time into route design. We describe the heuristic in Section 6.1 and present computational results in Section 6.2.

6.1. The decomposition-based heuristic with travel time

The set-partitioning formulation (1a) to (1d) for the MSRA-e incorporates travel time constraints by limiting the feasible route set. Travel time restrictions reduce the computational effort to solve the MSRA-e optimally, since the objective functions need only be evaluated for time-feasible routes. However, even with this limited set, the computational effort precludes enumeration of the feasible route set and evaluation of the expected minimum fill rates for all feasible routes. Therefore, we develop the Decomposition Based Heuristic with Travel Time (DBH$_T$) that generalizes the DBH to construct time-feasible routes while limiting the number of resource allocation evaluations. The DBH$_T$ incorporates travel time in solving the clustering and sequencing problems and implementing improvement moves.

### Table 11. Optimality gaps due to the DBH clustering solutions

| Number of agencies | $N = 4$ | $N = 5$ | $N = 6$ | $N = 7$ |
|--------------------|---------|---------|---------|---------|
| DBH clusters       |         |         |         |         |
| Average optimality gap due to clustering (%) | 0.3 | 0.5 | 0.4 | 0.8 |
| Maximum optimality gap due to clustering (%) | 6.0 | 8.8 | 7.5 | 11.0 |
| % of instances with optimality gap = 0 | 77.4 | 72.3 | 71.9 | 67.4 |
| % of instances with optimality gap $\leq 0.02$ | 94.2 | 91.8 | 92.7 | 85.7 |
| % of instances with optimality gap $\geq 0.10$ | 0.0 | 0.0 | 0.0 | 0.1 |

### Table 12. Comparing expected waste (in percentage) of the DBH solutions with the expected wastes from the optimal MSRA-w solutions and from the optimal MSRA-e solutions

| Number of agencies | $N = 4$ | $N = 5$ | $N = 6$ | $N = 7$ |
|--------------------|---------|---------|---------|---------|
| Average $H_w - Z^*_w$ | 1.5 | 2.0 | 2.3 | 2.6 |
| Maximum $H_w - Z^*_w$ | 6.2 | 6.9 | 9.5 | 9.9 |
| Average $H_w - \tilde{w}$ | -0.1 | 0.1 | 0.2 | 0.4 |
| Maximum $H_w - \tilde{w}$ | 4.1 | 3.1 | 3.5 | 4.7 |
6.1.1. Constructing initial routes

Clustering.

The clustering formulation in the DBH ((6a) to (6i)) is modified to incorporate travel time through an additional term in the objective function that encourages assignment of nodes to clusters based on travel time, similar to the generalized assignment problem of Fisher and Jaikumar (1981). We define a seed node for each cluster and consider the proximity of each node to the seeds. The first seed selected is the one with the maximum distance to all other nodes. We continue selecting the other \( K - 1 \) seeds by calculating each node’s total distance to the current seeds. Let \( t_{ic} \) denote the travel time between node \( i \in \mathcal{N} \cup \mathcal{M} \) and the seed of cluster \( c \in \mathcal{C} \). Let \( \lambda_T \) denote the weight for the travel time objective such that \( \lambda_V + \lambda_F + \lambda_T = 1 \). We modify the objective function (6a) of the clustering formulation as follows:

\[
\min \lambda_V g(V) - \lambda_F g(F) + \lambda_T g(T),
\]

\[
\text{where} \quad g(T) = \frac{\sum_{i \in \mathcal{C}} \sum_{i \in \mathcal{N} \cup \mathcal{M}} \delta_{i,c} t_{ic}}{K \times T}.
\]

Large values of \( \lambda_T \) are more likely to yield feasible routes in terms of the travel time constraint, which may be at the expense of large optimality gaps in terms of equity, whereas small values of \( \lambda_T \) may lead to large travel time feasibility gaps. Given a particular value of \( \lambda_T \), we solve the clustering model multiple times with different weights of \( \lambda_V \) and \( \lambda_F \) and obtain multiple clustering solutions. For each clustering solution, we then apply the sequencing heuristic below to construct routes.

Sequencing.

In each cluster, donor and agency sequences are constructed separately. Although we solely focus on minimizing travel time in sequencing donors, agencies are sequenced based on demand variability. The steps of the sequencing heuristic are as follows:

1. Agencies are sequenced based on the sequencing heuristic in the DBH.
2. The agency with the highest CV is selected as the first agency visited after donor visits. Given the depot and the first agency, a cheapest insertion heuristic is applied to create the donor sequence in a cluster.

After obtaining routes for each clustering solution, the algorithm proceeds with the solution with the smallest amount of total time infeasibility.

6.1.2. Improving solution feasibility and quality

The initial solution obtained by applying the procedure above may involve infeasible routes with respect to travel time restrictions. First, if all routes in the initial solution are infeasible with respect to the time constraint, we do not continue with the current iteration and reconstruct the initial solution with a different value of \( \lambda_T \). Second, if some of the routes in the solution are infeasible, we apply a set of feasibility moves that try to achieve feasibility. Finally, if all routes in the solution are feasible, we apply a set of improvement moves that attempt to improve solution quality without creating infeasibility.

Feasibility moves.

Feasibility moves attempt to (i) achieve the feasibility of routes or (ii) improve the time feasibility of feasible routes by increasing the time surplus (i.e., the difference between available travel time and route time). The resource allocation problem is not explicitly solved when applying the feasibility moves; however, the route supply and demand levels are considered in making node reassignments. We apply the following feasibility moves sequentially until no further improving moves exist or a maximum number of iterations is performed. After feasibility is ensured for all routes, the algorithm applies improvement moves.

In-route node exchanges. For an infeasible route, pairwise exchanges are first performed on the donor sequence. The donor exchange that ensures feasibility and leads to the largest time surplus is made. If feasibility is achieved by exchanging donors, agencies are not considered for exchange; otherwise, pairwise exchanges are performed on the agency sequence. For a feasible route only donor exchanges are considered.

Node removal/insertion: This move removes a node from an infeasible route and attempts to insert it to a feasible route. The move is evaluated for each infeasible route, the move that leads to the largest improvement in feasibility is selected in each iteration. For each pair of routes considered for removal and insertion, we calculate a proxy fill rate; i.e., the ratio of total expected demand to total supply. If the infeasible route has a higher proxy fill rate than the feasible route, we attempt to remove a donor from the route; otherwise, an agency removal is considered. After a node is selected for removal, its cheapest insertion heuristic is applied to create the donor sequence in a cluster.

Improvement moves.

Improvement moves are applied after a feasible solution is obtained. First, the resource allocation problem is solved on each route and the objective function value is calculated. The route with the lowest expected minimum fill rate (the worst-performing route) is identified. Then a set of moves is applied to increase the expected minimum fill rate of the worst-performing route. Similar to the DBH, we avoid an extensive local search in the DBH by in applying the improvement moves to limit the number of resource allocation evaluations. In the DBH, the resource allocation problem is solved only for the move that leads to the largest time surplus among the candidate moves,
Table 13. Instances for the evaluation of the DBH$_T$

| Number of nodes | Density | Number of agencies | Number of donors | Number of vehicles | Agency demands | Total supply levels | Donor supplies |
|----------------|---------|--------------------|------------------|-------------------|----------------|---------------------|---------------|
| 20             | Dense   | 8                  | 12               | 5 and 6           | Narrow range   | Scarce              | Narrow range   |
|                | Less dense |                   |                  |                   | Wide range     | Ample               | Wide range     |
|                | Scattered |                   |                  |                   |                |                     |               |
|                | More scattered |           |                  |                   |                |                     |               |
| 25             | Dense   | 10                 | 15               | 6 and 7           | Narrow range   | Scarce              | Narrow range   |
|                | Less dense |                  |                  |                   | Wide range     | Ample               | Wide range     |
|                | Scattered |                  |                  |                   |                |                     |               |
|                | More scattered |         |                  |                   |                |                     |               |

and the move is accepted if it improves the objective. The following improvement moves are applied sequentially for a given number of iterations or until an improved solution cannot be found.

**In-route agency exchanges:** This move checks whether the agency sequence in the worst-performing cluster is based on decreasing CV values and, if it is not, it attempts to restore the CV-based sequencing by exchanging the order of visits to agencies. The move starts from the end of the agency sequence and evaluates all pairwise exchanges to find feasible and improving moves.

**Agency removal/insertion:** The move checks whether an agency from the worst-performing route can be assigned to the best-performing route to improve the objective without violating feasibility. For each agency in the worst-performing route, all insertion positions in the pairing route are evaluated in terms of feasibility. If an improving insertion cannot be found, the move considers agency insertions to other better-performing routes.

**Donor removal/insertion:** This move adds a donor to the worst-performing route if it improves the objective without hurting feasibility. First, all insertion positions for the smallest donor of the best-performing route to the worst-performing route are evaluated for feasibility. If a feasible insertion cannot be found, the effects of removing larger donors from the best-performing route and then on feasibility and objective are evaluated. If an improving removal does not exist, the procedure considers removal of donors from other better-performing routes.

**Inter-route node exchanges:** The move focuses on the worst-performing route and checks whether any agencies (donors) from this route can be exchanged with the agencies (donors) of other routes feasibly while improving the objective. The move considers the best-performing route first as the pairing route and continues with the other better-performing routes as long as a feasible improving exchange cannot be found. Cheapest insertion strategy is used in determining the new positions of nodes in the pairing routes.

### 6.2. Computational analysis of the DBH$_T$

This section evaluates the performance of the DBH$_T$ to solve the MSRA-e with time restrictions.

### 6.2.1. Test instances

We evaluate the performance of the DBH$_T$ on 256 test instances. Table 13 summarizes the characteristics of the instances. We consider 20- and 25-node networks to evaluate the solution quality of the DBH$_T$, differing in node densities. Nodes are randomly located in each network. Since the travel time limits the number of feasible routes, we are able to solve larger instances to optimality than without travel time restrictions. Each vehicle can travel 7 hours for collection and distribution of food. The service time at a donor/agency is 20 minutes. We consider instances with five and six vehicles for the 20-node networks and six and seven vehicles for the 25-node networks.

Agency demand means and CVs are randomly drawn from uniform distribution; agency demand means are from U[80, 120] and U[40, 160] for narrow and wide settings, respectively. Demand CVs are drawn from U[0.8, 1.2] and U[0.4, 1.6] for narrow and wide settings, respectively. In instances with scarce supplies, the total supply is 80% of the total demand means, whereas in the settings with ample supplies, the total supply is set 20% larger than the total demand means. Supply amount at each donor is set randomly at wide range and narrow range.

The heuristic is run with four values of $\lambda_T$ from $10^{-4}$ to $10^{-2}$ and the solution with the best objective function value is selected. For each value of $\lambda_T$, the clustering model is solved for $\theta = 5$ times with different values of $\lambda_F$ and $\lambda_V$.

### 6.2.2. Solution quality and speed of the DBH$_T$

Table 14 presents optimality gaps and solution times for the instances. The solution quality of the heuristic generally improves when the number of vehicles is increased, since the heuristic can find feasible solutions easier when the travel time is less restricted. Similarly, gaps tend to be small in denser networks, in which number of feasible routes are larger. For the instances with the dense networks, the optimality gaps are close those of the unconstrained problem.

The best solutions are mostly found when $\lambda_T$ is small, as expected. The best solution corresponds to the solution found by using the smallest value of $\lambda_T$ in 50% of the instances, whereas only in 22.3% of instances can the best solution be found when $\lambda_T$ is set at its largest value. The
Table 14. Solution accuracy and times of the DBH

| Network type       | Number of nodes | Number of vehicles | Dense | Less dense | Scattered | More scattered |
|--------------------|-----------------|--------------------|-------|------------|-----------|---------------|
| Average optimality gap (%) | 20              | 5                  | 1.9   | 2.7        | 2.2       | 3.0           |
| Maximum optimality gap (%)   | 3.2             | 5.8                | 5.3   | 6.7        |           |               |
| % of instances with optimality gap ≤ 0.02 | 62.5            | 43.8               | 62.5  | 37.5       |           |               |
| % of instances with optimality gap ≥ 0.10 | 0.0             | 0.0                | 0.0   | 0.0        |           |               |
| Average solution time (CPU seconds) | 234.5           | 221.3              | 237.6 | 161.9      |           |               |
| Maximum solution time (CPU seconds) | 770             | 783                | 822   | 654        |           |               |
| Average optimality gap (%) | 20              | 6                  | 1.3   | 1.2        | 1.3       | 2.0           |
| Maximum optimality gap (%)   | 3.3             | 2.1                | 3.9   | 4.4        |           |               |
| % of instances with optimality gap ≤ 0.02 | 81.3            | 93.8               | 93.8  | 50.0       |           |               |
| % of instances with optimality gap ≥ 0.10 | 0.0             | 0.0                | 0.0   | 0.0        |           |               |
| Average solution time (CPU seconds) | 407.4           | 440.9              | 398.2 | 351.8      |           |               |
| Maximum solution time (CPU seconds) | 969             | 1074               | 1017  | 1002       |           |               |
| Average optimality gap (%) | 25              | 6                  | 2.8   | 4.0        | 3.1       | 4.5           |
| Maximum optimality gap (%)   | 7.5             | 7.3                | 5.6   | 10.4       |           |               |
| % of instances with optimality gap ≤ 0.02 | 50.0            | 6.3                | 25.0  | 18.8       |           |               |
| % of instances with optimality gap ≥ 0.10 | 0.0             | 0.0                | 0.0   | 6.3        |           |               |
| Average solution time (CPU seconds) | 373.9           | 336.9              | 355.3 | 384.8      |           |               |
| Maximum solution time (CPU seconds) | 654             | 581                | 696   | 870        |           |               |
| Average optimality gap (%) | 25              | 7                  | 2.1   | 3.9        | 3.1       | 3.5           |
| Maximum optimality gap (%)   | 4.1             | 7.2                | 8.3   | 8.1        |           |               |
| % of instances with optimality gap ≤ 0.02 | 56.3            | 25.0               | 37.5  | 25.0       |           |               |
| % of instances with optimality gap ≥ 0.10 | 0.0             | 0.0                | 0.0   | 0.0        |           |               |
| Average solution time (CPU seconds) | 698             | 675.9              | 653.1 | 642.0      |           |               |
| Maximum solution time (CPU seconds) | 983             | 1015               | 887   | 1053       |           |               |

results of the computational analysis show that the DBH can achieve high-quality solutions within reasonable times for different types of networks and demand/supply settings.

7. Conclusions and future work

In this article we address vehicle routing and resource allocation decisions in a nonprofit distribution system. Motivated by the FRP operations in the Chicago region, we study a multi-vehicle sequential allocation problem (MSRA-e) that incorporates two critical objectives for food distribution: providing equitable service and minimizing waste. We show that the MSRA-e achieves high levels of equity and resource utilization. Furthermore, the numerical analysis shows that the route sizes and compositions under different objectives can be substantially different. Specifically, the MSRA-e generally spreads the variability among different routes, whereas the benchmark problem MSRA-w pools the demand variability along the same route. In MSRA-e, the benefits of inventory pooling are higher when supply is ample and demand has low variability. In this case, using a smaller number of vehicles in food collection/distribution is better than using a large fleet. Based on the observations for the MSRA-e behavior, we develop an efficient heuristic for the problem that can handle an additional constraint on route length and obtain high-quality solutions in terms of equity and waste.

In this article, we enumerate all feasible routes to solve the MSRA-e optimally. An interesting area for future research would be to consider branch and price as an alternative approach to solve the set partitioning problem. Given the complexity of the allocation and sequencing problems, the pricing problem will not likely be able to be solved exactly, and future work could focus on the development of heuristic approaches to the pricing problem. Furthermore, the single food commodity assumption can be relaxed in future research. The single commodity model would be sufficient when there is only one high-valued commodity. However, when there are multiple critical items, choosing an appropriate objective function becomes important. For example, one could maximize the minimum fill rate over all commodities at an agency and include weights for fill rates by commodity type, if appropriate. Insights from the single commodity problem can support the development of objective functions that promote equity among commodities and agencies. Moreover, possible dependencies across the demands for commodities under limited budget and storage capacities can be considered. Future work may also consider alternative operating policies that do not
visit all donors first but rather integrate donor and agency visits. Note that in such a policy the allocation problem becomes more challenging because all donations are no longer known before allocation decisions must be made. Finally, as shown by the robustness analysis in this article, supply fluctuations at donors could affect the performance of FRP operations; therefore, relaxing the assumption of deterministic supplies can be considered in future research.

Acknowledgments

The authors thank Lisa Koch, Robert Matlosz, and Eric Knepper from the Greater Chicago Food Depository and Matthew Knott from Feeding America.

Funding

This research has been supported by grant CMII-0654398 from the National Science Foundation.

References

Balcić, B., Iravani, S. and Smilowitz, K. (2010) A review of equity in nonprofit and public sector: a vehicle routing perspective, in Wiley Encyclopedia of Operations Research and Management Science, J.J. Cochran, L.A. Cox, P. Keskinocak, J.P. Kharoufeh, and J.C. Smith, eds., Wiley, New York.

Campbell, A., Vandenbussche, D. and Hermann, W. (2008) Routing for

References

Funding

This research has been supported by grant CMII-0654398 from the National Science Foundation.

Appendix A

The MSRA-w

The MSRA-w focuses on minimizing the total expected waste and makes clustering, sequencing, and allocation decisions based on a waste-minimization objective. The MSRA-w can easily be formulated by only modifying the objective function (1a) of the set-partitioning model in Section 3.2. Let \( w_r \) represent the expected waste of route \( r \in \mathcal{R} \) found by applying waste-minimizing allocation and sequencing policies on route \( r \in \mathcal{R} \). Given \( w_r \) values for all \( r \in \mathcal{R} \), \( Z_w \) is the total expected waste of the system. The objective function of the MSRA-w, which minimizes the total expected waste, is formulated as follows:

\[
\min Z_w = \sum_{r \in \mathcal{R}} w_r y_r. \tag{A1}
\]

Similar to the MSRA-e, only the optimal route from each cluster is selected to the feasible route set \( \mathcal{R} \) in the MSRA-w. Specifically, for each cluster, the optimal waste-based sequencing and resource allocation policies are evaluated as follows:

1. **Sequencing.** Neither agency nor donor sequencing impact the MSRA-w objective function: expected waste is only dependent on the total supply and demand in the cluster. Thus, generating a single feasible sequence is sufficient.

2. **Resource allocation.** Given a sequence \( 1 \to 2 \to \cdots \to N_c \), and supply \( s_i \) upon arrival at agency \( i \) with demand \( d_i \), the optimal allocation amount at agency \( i \) is \( x_i^* = s_i \wedge d_i \). The optimal expected waste, \( w_r \), is calculated as

\[
w_r = \mathbb{E}_{D_1,\ldots,D_N_c}\left[ s_0 - \sum_{i \in N_c} x_i^* \right]. \tag{A2}
\]

Appendix B

TND allocation heuristic

To solve the allocation problem on a given route \( r \) with \( N_c \) agencies visited in order \( 1 \to 2 \to \cdots \to N_c \), apply the following steps.

0. Initialize \( i = 1, s_i = s_0 \) and \( \rho^0_{min} = 1 \).

1. Determine supply allotment for the two-node problem (for nodes \( i \) and \( i+1 \)), \( \hat{s}_i \), using \( \hat{s}_i = s_i \frac{\mu_i + \mu_{i+1}}{\sum_{j \neq j} \mu_j} \).

2. Determine the allocation amount at node \( i \), \( \hat{x}_i \), for the two-node problem (for nodes \( i \) and \( i+1 \)) after observing \( d_i \): \( \hat{x}_i = \min\{H(\hat{s}_i, d_i), \beta_{\min}^{(i+1)} d_i\} \), where \( \hat{m}_i \): medianof
Multi-vehicle allocations in nonprofit organizations

\[
\text{demand at node } i, \delta_{i+1} = \frac{\hat{m}_i - \hat{m}_{i+1}}{\text{avg}(\hat{m}_i, \hat{m}_{i+1})}, \text{ and } \hat{H}(\hat{s}_i, d_i) = \frac{\hat{s}_i d_i + \hat{d}_i (\hat{m}_i + 1 + \sqrt{\hat{s}_i})}{2}.
\]

3. Update \( \beta_i^{\ell \min} = \beta_i^{\ell - 1 \min} \wedge \frac{\hat{s}_i}{d_i} \) and \( s_{i+1} = s_i - \hat{s}_i \).

4. If \( i + 1 < N_c \), then set \( i = i + 1 \) and go to Step 1. Else, \( i + 1 = N_c \); STOP and set \( \hat{x}_{N_c} = s_{N_c} \wedge d_{N_c} \).

Biographies

Burcu Balci is an Assistant Professor of Industrial Engineering at Ozyegin University. She received her Ph.D. in Industrial Engineering from the University of Washington and was a Postdoc at the Industrial Engineering and Management Sciences at Northwestern University. Her research primarily focuses on developing mathematical models and solution methods to improve the design and management of humanitarian relief chains and other public/non-profit systems. Dr. Balci is a recipient of the BAGEP Research Award of Bilim Akademisi-the Science Academy, Turkey.

Seyed Iravani is a Professor of Industrial Engineering and Management Sciences at Northwestern University. His research interests are in the applications of stochastic processes, game theory, and queueing theory to the design and control of manufacturing, service operations, health care, supply chains, and non-profit systems.

Karen Smilowitz is an Associate Professor in Industrial Engineering and Management Sciences at Northwestern University. She studies modeling and solution approaches for logistics and transportation systems, in both commercial and nonprofit applications, working with transportation providers, logistics specialists, and a range of nonprofit organizations.