Food Redistribution as Optimization

Caleb Phillips, Rhonda Hoenigman, and Becky Higbee

CU-CS-1085-11 August 2011

University of Colorado at Boulder

Technical Report CU-CS-1085-11
Department of Computer Science
Campus Box 430
University of Colorado
Boulder, Colorado 80309

Food Redistribution as Optimization

Caleb Phillips, Rhonda Hoenigman, and Becky Higbee

August 2011

Abstract

In this paper we study the simultaneous problems of food waste and hunger in the context of the possible solution of food (waste) rescue and redistribution. To this end, we develop an empirical model that can be used in Monte Carlo simulations to study the dynamics of the underlying problem. Our model's parameters are derived from a unique data set provided by a large food bank and food rescue organization in north central Colorado. We find that food supply is a non-parametric heavy-tailed process that is well-modeled with an extreme value peaks-over-threshold model. Although the underlying process is stochastic, the basic approach of food rescue and redistribution appears to be feasible both at small and large scales. The ultimate efficacy of this model is intimately tied to the rate at which food expires and hence the ability to preserve and quickly transport and redistribute food. The cost of the redistribution is tied to the number and density of participating suppliers, and costs can be reduced (and supply increased) simply by recruiting additional donors to participate. Our results show that with sufficient funding and manpower, a significant amount of food can be rescued from the waste stream and used to feed the hungry.

1 Introduction

There is a contradiction present in the United States (US) today: between 27% and 50% of food produced for consumption is wasted in some stage of production, distribution, or preparation [10, 19, 18]. Meanwhile 14.7% of americans (1 in 7) are having difficulty finding enough to eat, and 5.7% are going hungry[1]. The populations most at risk of hunger include 36.6% of households with children and single mothers, 27.8% of households with children and single fathers, 26.9% of hispanic households, 24.9% of black households, 21.3% of households with children, and 7.8% of seniors living alone [21]. Globally, food prices have reached unprecedented levels in 2011, and we are currently in the midst of a worldwide hunger epidemic [7]. A clear question arises when studying these statistics: is it possible to recover food from the waste stream and redistribute it to those who are hungry, thereby reducing both waste and hunger?

The idea of food rescue and redistribution is not a new one. Non-profit food rescue and gleaning organizations (e.g., [12, 27, 17]) have been operating on this basic premise for more than 30 years, and there are dozens of organizations of varying size that currently rescue food in some capacity and redistribute it. These organizations recover food that would otherwise be wasted from *donors* (e.g., grocery stores, farms, retailers, restaurants) and redestribute it via *agencies* (e.g. food banks/pantries, soup kitchens, and shelters) to those in need. Recently, a coalition of major grocers and retailers organized under the Feeding America project with the goal of large scale food rescue, redistribution, and documentation [16]. Two popular recent books have studied the problem of systematic food waste in both the U.S. and Europe [34, 2]. Yet, to our knowledge there has been no prior effort to quantify the cost and practicality of the food rescue and redistribution model on a large scale.

Identifying sustainable systems for reducing food waste and hunger, and undestanding their cost, can have a substantial impact on waste reduction and food distribution on a local, national and even global

[1]The United States Department of Agriculture (USDA) defines these two classes of hunger as "low food security" and "very low food security".

scale. In this paper, we make an important first step in this direction by studying the food rescue and redistribution dynamics in a pair of neighboring counties in north central Colorado. We investigate the task of food recovery as a time-sensitive spatial distribution problem involving food supply and demand and the energy cost of redistribution from donors to agencies. To this end, we build an empirical model using data from a large food rescue organization in north central Colorado. Although our results are limited to this region, we contend that it is representative of a large class of similar regions with a mix of rural, small urban, and urban service areas. Using this model, we present an optimization framework for finding low cost solutions to food recovery and redistribution, which can be used to determine average, best case, and worst case bounds on the problem through simulation.

The main contributions of this work are as follows:

- We present the first formal description of the food redistribution problem as a well defined optimization problem.
- We investigate a novel dataset of food rescue and distribution from a large food bank and discover that the food supply process is heavy tailed and can be well modeled with extreme value theory.
- We present initial results on the feasibility of the redistribution approach and find that basic demand can be met easily with available supply, and with small changes and optimizations, a substantial portion of hunger may be mitigated with rescued food.

In the next section we describe the data collected for this study and the statistical models we derive from it. In section 3, we will discuss the balance of our model, including the optimization framework, providing background on similar problems. Section 4 describes the implementation of our simulation framework and design decisions. Section 5 discusses our experimental design and section 6 provides the results of the experiments. Finally, section 7 concludes with a summary and discussion of future work.

2 Measurements

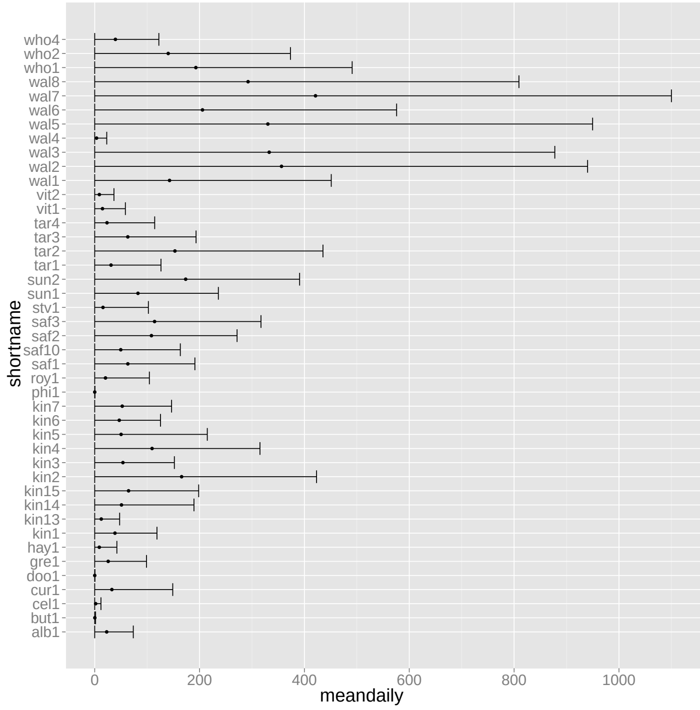
The data we use in this study was supplied by Community Food Share (CFS), the sole food bank for Boulder and Broomfield counties in north central Colorado [31]. This data includes the pounds of food received from each donor on each day for one month: December 2010. Donations from 46 distinct organizations as well as several anonymous donors, comprising 1453 donations and 186,980 lbs of food[2]. This food was distributed to 304 unique agencies which are predominantly homeless shelters, soup kitchens, smaller food pantries, and other organizations that serve at risk populations. Figure 1 plots the mean daily supply from each donor as well as a probability distribution function (PDF) of the daily donations. In the PDF, the donors are divided into three categories: manufacturers, bakeries, and grocers, which comprise the vast majority of donor types. Some donors provide substantially more food than others and the largest grocery stores provide the most food on average. We will look at correlations between store size and amount donated in section 2.2. The PDF hints that the underlying data is heavy-tailed, with zero-supply as a normal case and large supply events occurring with smaller probability. Although each of the categories has its own shape and scale, the same basic heavy-tailed distributional shape seems to dominate.

2.1 Characterizing Supply (Waste)

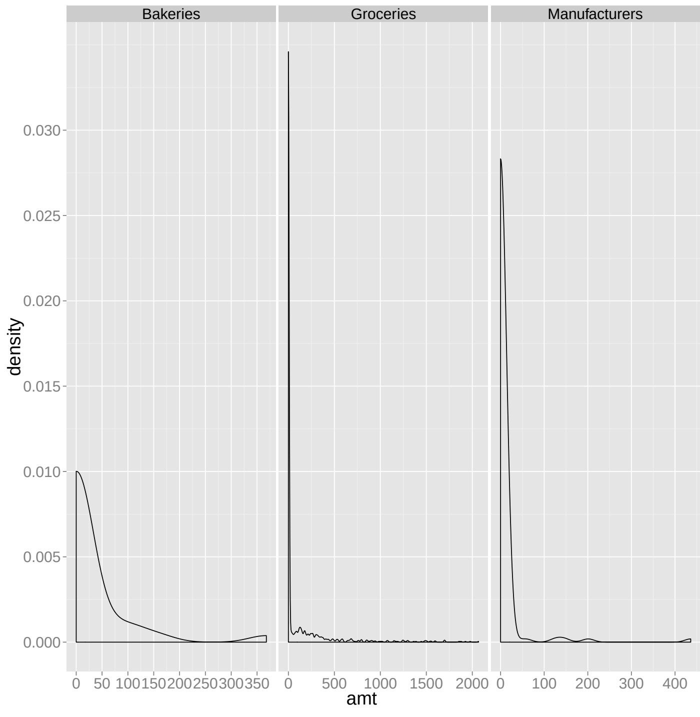
We use the data provided by CFS to charactize the dynamics of the supply and demand processes statistically. On a given day, the data includes some number of donations, or "food supply events". For

[2]Most food rescue organizations weigh their donations so that they can provide documentation to their donors for the purpose of tax-deduction. CFS uses categories based on USDA food "pyramid" groups: grains, vegetables, fruit, protein, dairy, fat/sweets, soup/baby food, and non-food.

[2]Donations are broken up by food category, so an individual pickup may be documented with as many as 7 distinct donations. Some donations are recorded by case and are not weighed. In this case we have taken a lower-bound estimate that one case weighs one pound. In practice these cases may weigh anywhere from 10-15 lbs and sometimes as much as 25. Because we do not have the exact weight available, we have erred on the side of conservativism.



(a) Mean Daily Supply (lbs)



(b) Daily Supply Distribution (lbs) by Category

Figure 1: Daily supply data by donor and category in pounds.

Data	Threshold (θ)	Rate (r)	Location (μ)	Scale (σ)	Shape (ζ)	Mean
All Suppliers	0	0.236	0	374.406	0.077	405.64
Only Grocers	0	0.271	0	394.350	0.054	416.86
Only Manufacturers	0	0.045	0	140.820	-0.010	139.43

Table 1: GPD fit parameters for daily supply and demand distributions in pounds.

a given donor, we make the assumption that on days where food is not picked up, there is zero supply. This data has a distinct heavy tail, where distribution is skewed to the left, indicating that small values are most common, but that with small probability, large and sometimes extremely large values can be observed.

This type of distribution is well modeled by a Peaks Over Threshold (POC) approach, using a fit from a Generalized Pareto Distribution (GPD). The Pareto distribution is a power law probability distribution which features a long heavy tail (i.e., most events are small, but some events can be extremely large). The POC approach to fitting is common in weather modeling, and particularly in modeling the likelihood of extreme weather events [6]. In this approach, there is a threshold value (θ) for an event, that will be selected at some constant rate (r). In our problem the threshold value is zero, and we are looking to model the probability and size of food supply events greater than that threshold. These are the events where the donor has waste that can be recovered. Events larger than the threshold are modeled with a heavy tailed distribution described by location (μ), scale (σ), and shape (ζ) parameters. These parameters describe the basic shape, width, and location shift of our fit. The location parameter, which is the mean in the parametric sense (i.e., it is the "center" of the entire distribution) is zero. The scale parameter describes the uncorrected central tendency of the tail and the shape parameter describes its spread.

In order to acheive tighter fits, we fit the supply data for all stores as well as manufacturers and grocers separately. Figure 2 shows the quality of the fit for all stores combined. This figure provides the probability plot, quantile-quantile plot, and return level plot, which visually compare the distribution of the observations and the fitted model. In these plots, the closer points are to the diagonal line, the better the model is able to predict the data. The figure in the bottom right shows a histogram of supply events greater than zero and the corresponding linear fit. Table 1 provides the maximum likelihood estimator (MLE) fitted parameters. Overall the fits are remarkably good: the overall data is clearly Pareto, as is the supply of Grocers, who are the predominant type of donor in our data. This is an exciting result because it means that food supply (waste) events can be modeled with, and thought about, using the same models used to predict and forecast extreme weather events.

2.2 Predicting Supply

Using the fit parameters in Table 1, Pareto-distributed daily supply values can be generated:

$$s_{i,t} = \begin{cases} \mu + \frac{\sigma * (U(0,1)^{-\zeta} - 1)}{\zeta} & U(0,1) <= r \\ \theta & o.w. \end{cases} \quad (1)$$

Where $U(0,1)$ is a uniformly distributed random number in $[0,1]$. We use this function in our experiments below to generate random, correctly distributed supply values that can be used in Monte Carlo[3]style simulations.

In order to use this model to predict the supply of stores for which we do not have data, it is necessary to understand how other variables contribute to the shape and magnitude of the supply distribution. To this end, we investigate several other variables that might be correlated with mean daily supply:

[3]Monte Carlo methods are ones that compute or predict the behavior of a (possibly) complex system by releated random sampling. In the classic approach, random samples are drawn from the appropriate probability distribution and then used in a deterministic algorithm to simulate the underlying dynamics of the system stochastically.

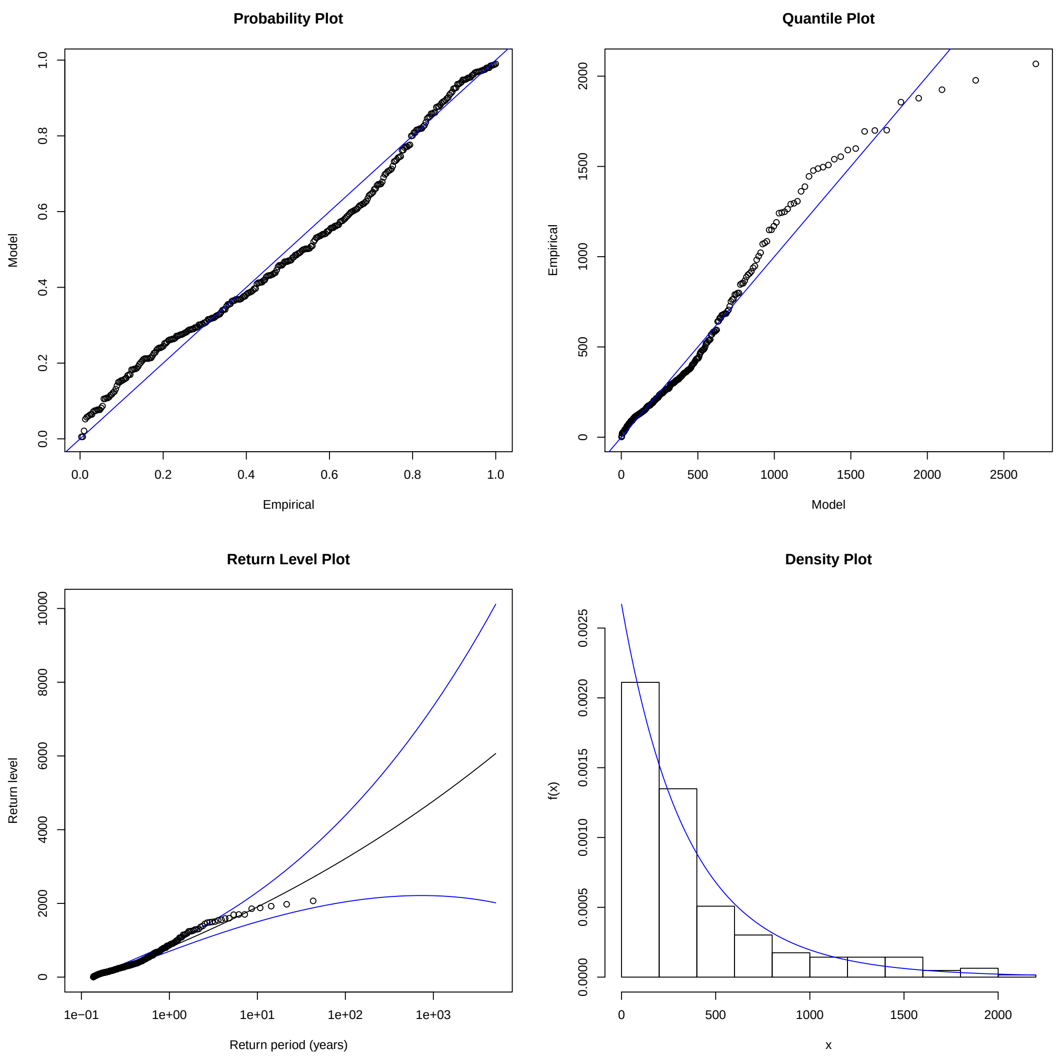


Figure 2: Generalized Pareto fit statistics for supply data.

store size (building square footage), municipal zoning category, and store distance from the warehouse. Zoning categories and square footage were calculated using publicly available geographical information published by the City of Boulder [5]. The distance of each donor from the warehouse was calculated using drive-distance from the MapQuest Directions API [20].

We perform a factorial analysis of variance (ANOVA) to see how each of these factors (individually and in combination) effects the mean daily supply. The GPD mean daily supply[4]($\bar{s}$) for a given supplier is:

$$\bar{s} = \mu + \frac{\sigma}{1-\zeta} \tag{2}$$

To obtain this value for each supplier, we perform a GPD fit to determine the GPD parameters and then calculate the mean. The result of the ANOVA shows that the most important correlating variables are size and zoning category with F-values of 82.32 and 9.68 respectively[5], and p-values well below the conservative significance threshold of $\alpha = 0.01$. For size alone, a Pearson's product-moment correlation test provides a correlation coefficient of 0.623 and a small p-value. Given this, we can say that the mean daily supply (waste) and variability is independently correlated with the size of the donor. This is an important result because it allows us to predict the supply (waste) distribution for a given donor based simply on publicly available information: square footage and municipal zoning category.

We perform a linear least squares fit to the GPD mean daily supply as a function of donor square footage. This produces a fit with slope (m) of 2.554e-03, intercept (b) of 96.22, and adjusted R^2 of 0.372. Figure 3 shows this relationship along with the fit line. Using this information we are able to predict the scale of the supply distribution for a new donor using its square footage (x):

$$\mu + \frac{\sigma}{1-\zeta} = mx + b \tag{3}$$

Since we know that the most important parameter is scale (σ), which by itself is well correlated with supply, we focus on predicting it and draw the remaining unknowns (μ, ζ) from the general fits given above. Solving for σ, and substituting in 0.077 for ζ from table 1, 0 for μ, and the m and b values given above we obtain:

$$\sigma = 0.923(0.002554x + 96.22) \tag{4}$$

2.3 Characterizing Demand (Hunger)

A final modeling task is to fit demand data. Due to privacy concerns for some agencies, it is not possible to determine the per-agency demand from any defining characteristics, such as agency size or surrounding population density. Instead, we fit the aggregate demand each day, which corresponds to the amount of food delivered by CFS or picked up at the warehouse directly by the agency.

During the month of December, CFS distributed food on 17 of 31 days. Figure 4 shows the kernel-smoothed probability density of a given daily demand (in lbs) for those 17 days. Unfortunately, the data is insufficient to provide a strong statistical fit. If we assume the underlying distribution is parametric, which seems reasonable given that total demand is likely a strict function of the number of hungry people, which should not be highly varying in time, we can summarize the distribution with the mean (3939 lbs) and standard deviation (2697 lbs). This is the demand for food in pounds per day.

This empirically derived average value is not a perfect indicator of the actual demand in the CFS service area. The food required to feed all people with food insecurity could actually be much higher. Several reports have detailed food insecurity in the US, reaching different conclusions about the extent of the problem. According to [16], which describes the efforts of the Feeding America program, 5.7 million

[4]This same analysis can be performed using well known parametric statistics (somewhat questionably), such as the arithmetic mean of daily supply. In that case the ANOVA and correlation results presented below still hold, but with slightly less confident p-values.

[5]The F-value is a statistic that describes the ratio between explained variance and unexplained variance. Or, put differently, the ratio of between-group variability to within-group variability.

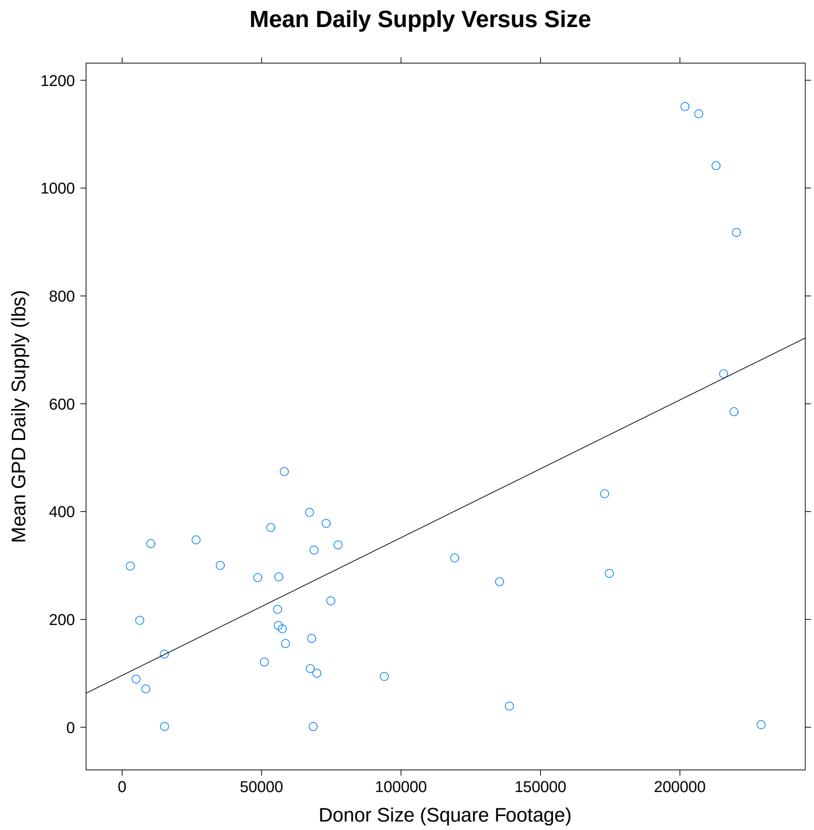


Figure 3: Correlation between mean GPD

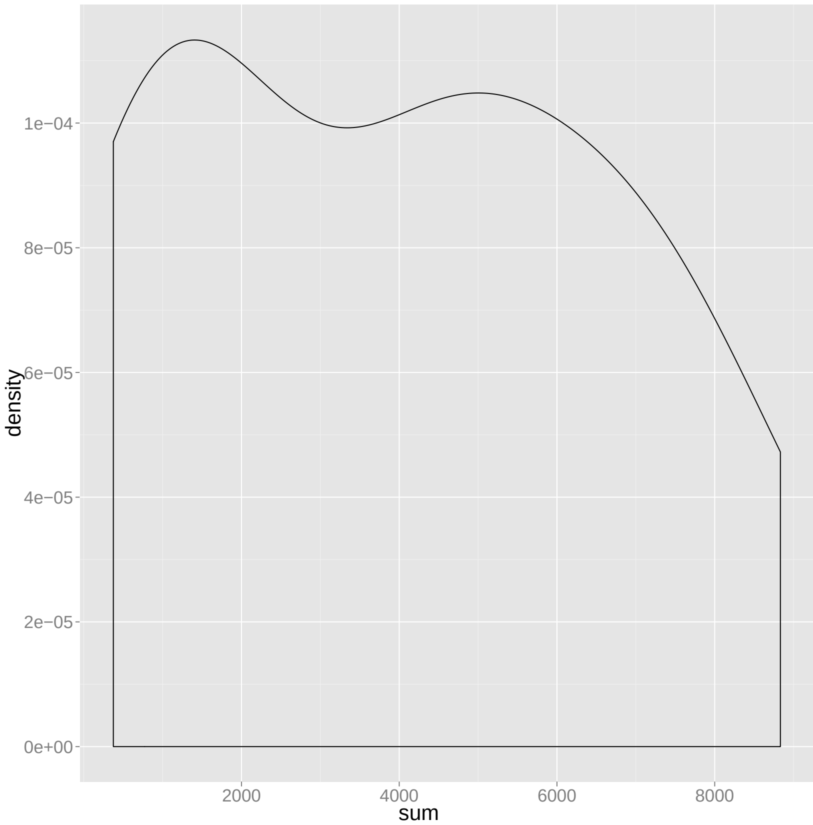


Figure 4: PDF of daily total distribution data in pounds.

unique individuals (or 1.661% of the US population[6]) are served each week by the approximately 37,000 agencies aligned with their program. There are 348,017 individuals in the service area of CFS[7], meaning that, based on national-level statistics, there are approximately 5,781 unique individuals per week that will need food assistance. In [21], Nord et al. show that nationally in 2009, 14.7% of households were food insecure at some point during the year, 9% had low food security, and 5.7% had very low food security[8]. Using the 5.7% figure, that would indicate that 20,490 individuals have very low food security in the service region of CFS. A local study conducted by Feeding America in conjunction with CFS found that approximately 10,800 unique individuals seek food assistance per week in the CFS service region [15]. Using USDA statistics for average consumption of food, a typical American in 2010 consumes approximately 2.85 lbs of food per day of which the majority is meat and protein (0.41 lbs) and grain (0.48 lbs) [23][9]. Given this, if we assume that each individual who has very low food security acquires a third of their daily intake via CFS, the mean daily demand would be between 5,491 lbs (using national Feeding america levels), 10,260 lbs (using local CFS Feeding America service statistics), 19,465 lbs (using USDA very low food security percentile), and 48,600 lbs (using USDA low food security perentile). This is a staggeringly large range that serves to highlight the fact that consensus on hunger and food demand does not exist. Even the lowest daily demand of 5,491 lbs is well below the current food distribution from CFS, indicating that there could be several thousand people with food insecurity who are not receiving food assistance.

In the remainder of this paper, we will treat the high end estimate of 48,600 lbs per day as the gold standard—the equivalent of providing one meal to every person in the service area who experiences some form of food insecurity. Although this would equate to feeding people who, for the most part, have enough money to feed themselves, it would do a great deal towards helping the poorest Americans avoid weighing tradeoffs between nutritional value of purchased foods, and other expenses. Similarly, we will treat 10,260 lbs as the goal, which equates to the net demand that CFS and other area agencies already attempt to fulfill during times of highest demand (i.e., a week when all unique clients come to pick up food).

In the next section we will describe the remainder of the model, which defines the mechanics of supply and demand using an optimization framework. In section 4, we will tie together that optimization model and the statistical models here into a simulator that can be used to study the dynamics and feasibility of the food redistribution problem.

3 Model

In this section we develop a model of food supply, demand, and cost of delivery between areas of supply and areas of demand. The premise of the model is simple: there are some number of potential donors that generate some amount of recoverable waste each day. Performing a pickup at a given donor incurs a cost, which is related to how far away that donor is, and optionally how much food is to be transported. The goal of the model is to determine a schedule for each day that chooses which donors to visit such that food demand (hunger) is met and minimizes the cost of doing a pickup.

We begin by discussing optimization problems in general and related approaches in order to help motivate the direction we have chosen. Then, we describe the general model for our problem and two important extensions to improve its realism that take into account food expiration and food storage at a central warehouse.

[6] The US Census Bureau estimates the population of the US as 307,006,550 in 2009.

[7] The US Census Bureau estimates the population of Boulder and Broomfield counties at 303,482 and 55,990 respectively.

[8] Low food security is defined by the USDA as "Reports of reduced quality, variety, or desirability of diet. Little or no indication of reduced food intake" and very low food security is defined as "Reports of multiple indications of disrupted eating patterns and reduced food intake" [22].

[9] Assuming the weight of dairy products is equivalent to the same volume weight of water and the weight of vegetables and fruit is equivalent to half the weight of the same volume of water. The latest available usage statistics are from 2008.

3.1 Background and Related Problems

Optimization is the process of finding the best option from a set of options using a fitness function to capture the relevant details of the problem [26]. In the food redistribution problem, the relevant details are the food supply from donors, food demand from agencies, and the cost of picking up food from donors and delivering it to agencies.

There are several classes of problems that provide context for the food redistribution problem, and these classes are not mutually exclusive. There are resource allocation problems in which limited resources are allocated to maximize a goal, such as profit, i.e. [14, 30]. There are also minimum cost flow problems such as shortest path, transportation, or maximum flow network, i.e. [35, 33]. In shortest path problems, the objective is to find the minimum distance, time, or cost for a sequence of activities or travel through a network. The maximum flow problem expands on shortest path by adding constraints on the flow capacity between nodes, which limits the flow to some maximum value. In real-world examples, this would be equivalent to restricting the number of cars on a road between cities, or the size of pipelines. In transportation problems, the objective is to determine the minimum cost between centers of supply and demand.

With food distribution, the objective is to find the minimum distance through a network of donors that provides enough food to meet demand. In this way, it is a shortest path problem, where the flow has to be greater than a constraint, which is the demand. A fitness function can be formulated that is used in a solution method. One common method for solving these types of problems is linear programming (LP) [13]. This approach finds the optimal solution on problems where inputs and constraints are linear combinations. Mileage, for example, is linear. In food redistribution, the cost per mile is the same regardless of which path in the network is being travelled. In addition, the decision to visit any given store on that day is discrete — we either go to the store or we do not. If we visit the store, we get 100% of the supply available, but we also incur 100% of the cost of travel.

While LP is common for solving minimum flow problems, it is possible for these problems to also be nonlinear. In food redistribution for example, we can use the LP approach to find the optimal solution using a daily cost calculation. However, formulating the problem to find a multi-day pickup schedule introduces non-linearities that make the LP approach less effective. The multi-day solution generates a pickup schedule based on knowledge of food supply of the current and subsequent days as well as how the food expires over time. The most cost-effective solution could be to only pickup food once a week for any given store. However, if the store has perishable items, one week might be too long to wait. In addition, if a store has a small donation one day, but will have a large donation the next day, the optimal solution might be to batch the pickups—skipping the pickup on the current day and picking up both the current and the next day's supply at one time.

For complex optimization problems that cannot be solved using LP, or other simple solving methods, heuristic and metaheuristics methods are often used to find approximate solutions [13]. These approaches have been used for minimum flow, i.e. [25, 8] and resource allocation problems in which there is a time element that requires a multi-day solution, i.e. [4, 9].

3.2 Food Redistribution Model

Our model is as follows. We start with N donors with food supply, and M agencies with food demand. The aggregate available food supply ($\hat{s}$) on a given day (t) in arbitrary units is the sum of the supply from each individual donor:

$$\hat{s}_t = \sum_{i}^{N} s_{i,t} \tag{5}$$

Similarly, aggregate demand ($\hat{d}$) on a given day (t) in arbitrary units is the sum of demand at each individual agency:

$$\hat{d}_t = \sum_{i}^{M} d_{i,t} \tag{6}$$

Picking up food everyday from every donor would retrieve every pound of available food, but it might not be cost effective. We establish a schedule where everyday, there is a pickup from a subset of donors. This multi-day pickup schedule is controlled by a boolean matrix ($r_{i,t}$), which identifies which suppliers have pickups scheduled on which days:

$$r_{i,t} = \begin{cases} 1 & \textit{pickup at supplier } i \textit{ on day } t \\ 0 & o.w. \end{cases} \tag{7}$$

Each donor is associated with a constant cost of making a pickup (c_i). The total cost ($\hat{c}_t$) on day (t) is the cost to visit the selected donors on that day:

$$\hat{c}_t = \sum_{i}^{N} c_i * r_{i,t} \tag{8}$$

The total supply for that pickup schedule ($\hat{q}_t$) is then:

$$\hat{q}_t = \sum_{i}^{N} s_{i,t} * r_{i,t} \tag{9}$$

In this formulation, the objective is to minimize $\hat{c}_t$, such that $\hat{q}_t \geq \hat{d}_t$ for $\forall t$. This will minimize cost and provide enough supply to meet demand:

$$min \sum_{t}^{T} \hat{c}_t \; s.t. \; \hat{q}_t \geq \hat{d}_t \tag{10}$$

3.3 Food Expiration

Establishing a multi-day pickup schedule for food presents a unique challenge in that food can go bad. Some of the food not picked up on day t will expire by day $t+1$, but the other food will remain. The rate of expiration is related to the way the food is stored, the state of the food, and the type of food. In this model we've made the simplifying assumption that all food expires at the same rate regardless of these conditions. The food supply available on day $t+1$ is the new supply for that day, plus the previous day's supply that was not picked up and did not expire:

$$s_{i,t+1} = s_{i,t+1} + \epsilon(1) * s_{i,t} * (!r_{i,t}) \tag{11}$$

where $\epsilon(\Delta t)$ is $\in [0,1]$ and quantifies the fraction of food not expected to expire over Δt nights. $!r_{i,t}$ is the inverse of the boolean scheduling matrix (and hence is 1 when a pickup does not occur and 0 otherwise). Further extensions can consider more complete definitions of ϵ. This recurrence can be trivially rewritten as a sumation of the previous t days:

$$s'_{i,t} = s_{i,t} + \sum_{a=0}^{t-1} \epsilon(1) * s_{i,a} * (!r_{i,a}) \tag{12}$$

3.4 Food Storage

Another important component to the model is a central warehouse where excess food can be stored after it is picked up. This allows for overages in recovery to be used the following day. The warehouse supply on a given day t is the amount picked up on day t minus the day t demand plus the warehouse supply from the previous day that did not expire:

$$\hat{w}_t = (\hat{s}_t - \hat{d}_t) + \hat{w}_{t-1} * \epsilon(1) \tag{13}$$

It should be noted that the assumption of a central warehouse is not universal. For instance, the largest food rescue organization in the United States, City Harvest in New York City, operates without a central warehouse and instead delivers goods in the same day they are rescued, often as part of the same tour [12]. CFS, on the other hand, operates using a central warehouse in Niwot, Colorado. Each day, three full-time drivers make scheduled pickups at area donors and bring the goods back to the warehouse, where they are stored for later distribution.

4 Implementation

In this section we tie together the empirical results in section 2 and the model outlined in section 3 to develop a simulation environment that can be implemented and studied in practice. In particular, we will discuss approximations and assumptions that are necessary in order to allow experimental analysis of the problem in a reasonable amount of time.

4.1 Iterative Multiday Solution

Finding the optimal solution using the model described in section 3 is combinatorial and has a worst-case complexity of $O(2^{NT})$ for N suppliers and T days. For all but the smallest N and T, this problem is intractable. However, as discussed in section 3.1, the optimal solution for a single day, with constant costs, can be formulated as a linear program and solved quickly:

$$\hat{c} = c_1 * r_1 + c_2 * r_2 + \cdots + c_n * r_n \tag{14}$$

$$min\ \hat{c}\ s.t.\ \sum_i^N r_i * s_i \geq \sum_i^M d_i \tag{15}$$

A sub-optimal solution for multiple days can be calculated by applying this linear program iteratively, once for each day:

$$C = \sum_t^T\ min\ \hat{c}\ s.t.\ \sum_i^N r_i * s_i \geq \sum_i^M d_i \tag{16}$$

The main failing of this approach is that it is unable to make multi-day decisions, where today's pickup schedule is based on knowing the supply for the following days. However, in the real world, it's unlikely this information would be available anyway. In this sense, the iterative solution may be more realistic than the the general purpose optimal solution.

For each time step (day), an optimal solution is found using a linear program (LP) solver. The constructed LP searches for a schedule (r) that minimizes the total cost subject to the constraint that the amount recovered must be greater than or equal to the amount demanded. To solve this equation on a computer we look to the freely distributed *lp solve* project, which is able to solve mixed integer linear programs [1]. Because our solution vector is actually a binary vector, the mixed integer and binary extensions are necessary. An example problem, presented in the lp language used by *lp solve* is given in figure 5.

```
min: 24.39x0 + 56.59x1 + 35.81x2 + 46.01x3 + 24.64x4 + 29.62x5 + 36.654x6;
c1: 177.99x0 + 552.06x1 + 12.0x2 + 5.2x3 + 2.1x4 + 3.02x5 + 1.2x6;
bin x0 x1 x2 x3 x4 x5 x6;
```

Figure 5: Example lp solve program formulated in the lp language.

4.2 Determining Costs

Constant costs for each supplier must be calculated in order for this problem to be solved as a linear program. The data from CFS provides us with the names of 46 suppliers. We first manually determine the street address of each donor and then calculate the distance of each supplier from the CFS warehouse.

One of our research goals is to understand the supply available at candidate donors that are not currently working with CFS. Therefore, we have attempted to locate all grocers, bakeries, and manufacturers in Boulder and Broomfield counties, that might supply donations. The total list of donors is 114. We cannot claim that this list is exhaustive, but we believe it represents the bulk of grocers, bakeries, and food manufacturers in the region. In future work, we hope to include restuarants and cafes as well, but have excluded them in the present study because CFS does not pick up at them.

To calculate distance, we use the MapQuest Driving API directions [20]. For each supplier in the CFS data, we retrieve driving directions (which presumably use an optimized shortest path, taking into account the actual constraints of the roads) and use the total driving length of the first offered route, in kilometers, as the cost of visiting that supplier.

4.3 Clustering Suppliers

In order to allow for batching pickups at nearby locations, we determine how our 114 suppliers are clustered together. We use the MapQuest Driving API to calculate shortest-route driving distance between all pairs of nodes. This requires 6498 unique queries and takes some time[10]. Figure 6 shows a higherarchical dendrogram of all 114 suppliers based on distance to each other. Arbitrarily drawing a line at 10km in this dendogram we can see that there are 18 clusters formed if we group all nodes within 10km of one another. With the smaller set of 46 suppliers that CFS currently visits, we see 8 clusters at a radius of approximately 10km. Based on this observation, we proceed with a k-means clustering with a goal of $k = 20$ clusters when considering all 114 suppliers and $k = 10$ clusters when considering the smaller set of 46.

In k-means clustering, observations are partitioned into k clusters so that each observation belongs to the cluster whose center (mean location) is nearest. To solve this, we use the method of Hartigan and Wong, as implemented in the R statistics package [11, 29]. The result is 20 points in space that the 114 suppliers are centered around so that the mean distance from any supplier to their point is minimized.

When using the resulting computed clusters, we make the basic assumption that when a pickup is done at any cluster member, all the other cluster members should be visited. In this sense, each supplier is its own cluster and the supply at that cluster is the sum of all members. To determine the cost of visiting a cluster we take the average driving distance of all cluster members from the central warehouse, multiply it by two (to count the return trip), and then add it to the minimum sum of the distances from one node to each other node. Hence, the cost for visiting a given cluster k is:

$$c_k = \frac{2}{N_k} \sum_{i}^{N_k} c_{w,i} + min_j \sum_{i}^{N_k} c_{i,j} \; s.t. \; i, j \in \mathbf{cluster}_k \tag{17}$$

where N_k is the number of members in cluster k, $c_{i,j}$ is the cost of driving from i to j, and w is the index of the warehouse.

[10] Originally, we made use of the Google Maps API, but eventually apandoned it in favor of MapQuest because of restrictions on the API in terms of rate limits and daily request quotas.

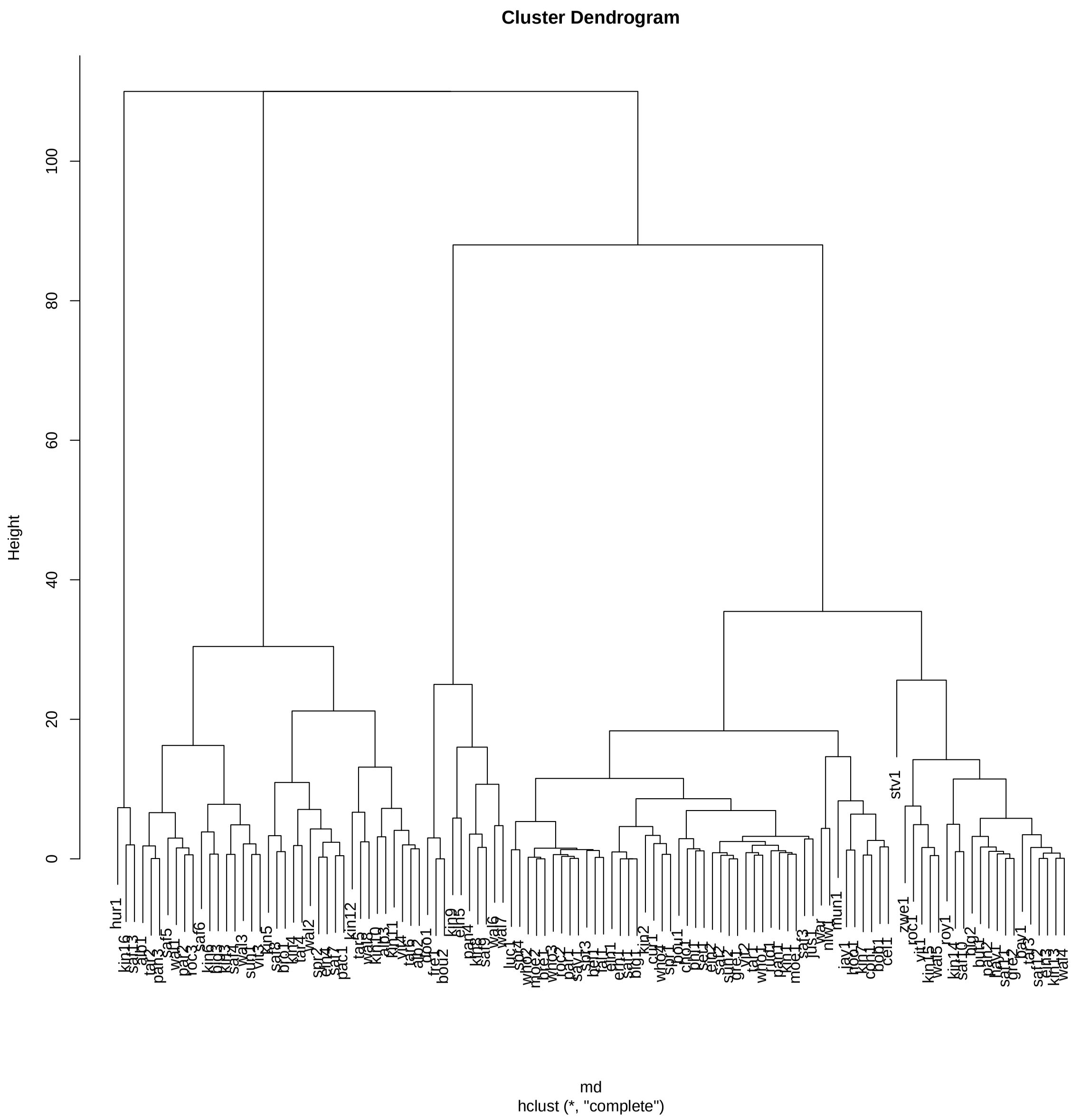


Figure 6: Hierarchical clustering of suppliers based on distance.

Although we cannot claim that this is a perfect solution, it is quick to compute, actually reduces the computational complexity of the problem, and models the way food rescuers actually do pickups. As we will see, it also produces cost values that are very close to the average costs incurred by CFS. Hence, we proceed with this cost function for the remainder of the paper.

4.4 Implementation Outline

Algorithm 1 shows the pseudcode implementation of the complete simulation algorithm. For each day, the total demand ($\hat{d}$) is a constant based on the expected demand as discussed in Section 2.3. If we are considering food storage in a warehouse and there is any food available already, this total demand is reduced by that amount to acheive the remaining net demand. Without considering storage the net and total demand will always be equal.

Supply values for each participating donor are drawn stochastically from the GPD distribution derived in Section 2.1. The scale parameter is calcuated for each donor using the square footage as discussed in section 2.2. Determining the square footage for each location is a laborious task due to inconsistent and somtimes unavailable geographic information system (GIS) data. For those suppliers that it is possible, we compute the square footage from municipal data for Boulder County [5]. For those suppliers where this data is unavailable, we use GIS tools and orthoimagery to calculate square footage manually.

Supply accumulates each day at donors if it is not recovered. However, it also expires at a constant rate of $1 - \epsilon$. In our implementation, ϵ is a constant between 0 and 1, but in future work we hope to investigate more realistic epsilon calculation (perhaps by type of food). In some cases, there is simply not enough supply to meet demand, in which case there is no feasible solution to the LP. When this occurs, we assume that all suppliers must be visited to get as much as possible (whatever the cost).

5 Experiments

One of our contributions in this paper is evaluating the feasibility of the redistribution approach for alleviating hunger. In this section we describe a study using the model from Section 3 and stochastically generated data from the empirical distributions described in Section 2 to create simulations of food redistribution scenarios.

Our objectives with these experiments are to

- Demonstrate that the model works in that it reproduces observed food supply from the CFS data.
- Evaluate how model parameters, such as ϵ and warehousing, affect food availability.
- Determine the relationships between food availability, the number of donating stores, and cost.

The approach we take here is a classic repeated measures approach, where the average behavior of a complex system is studied through repeated Monte Carlo simulation and *ex post facto* analysis. Each simulation is run for one year (365 days) and uses the same random seed (so that results can be compared). By studying the problem via simulation, we are able to investigate both the average case and extreme scenarios of food redistribution, which serves to answer whether this approach fills the gap between demand and current supply.

In this first set of experiments, we only consider pickups from the 46 donors that are already participating with CFS. The goal here is to reproduce and understand the dynamics of the food redistribution problem: how much energy (cost) must be expended to meet the worst-case demand, can demand be met reliably, how frequent are underruns ($supply < demand$) and overages (excess pickup), and how the rate of food expiry limits the amount of recoverable food. We include the warehouse, since CFS operates with a warehouse. We use an ϵ of 0.5, indicating that 50% of "waste" food is expected to expire in 24 hours (or, put another way, half of the food will remain on day $t + 1$, and one quarter on day $t + 2$ and so on). Then, using the same set of suppliers, we evaluate how food availability changes as a function of cost and ϵ. These experiments show removing cost restrictions affects how much food could be retrieved, as well as how food expiration contributes to food availability.

Algorithm 1 Multiday Simulator

$D \leftarrow$ constant demand
$\epsilon \leftarrow$ constant epsilon value
$seed \leftarrow$ seed for random number generator
$rate \leftarrow$ constant rate from GPD fit
$shape \leftarrow$ constant shape from GPD fit
$loc \leftarrow$ constant location from GPD fit
$w \leftarrow 0$
$\mathbf{sqft} \leftarrow$ square footage of each supplier
$\mathbf{c} \leftarrow$ constant supplier costs
$\mathbf{olds} \leftarrow 0$
for $i = 1 \rightarrow N$ **do**
    $scale_i \leftarrow 0.923 * (0.002554 * sqft_i + 96.22)$
end for
for $t = 1 \rightarrow T$ **do**
    for $i = 1 \rightarrow N$ **do**
        $r_1 \leftarrow$ random value in $[0, 1]$
        $r_2 \leftarrow$ random value in $[0, 1]$
        if $r_1 \leq rate$ **then**
            $s_i \leftarrow loc + scale_i * \frac{r_2^{-shape} - 1}{shape}$
        else
            $s_i \leftarrow 0$
        end if
        $s_i \leftarrow s_i + olds_i * \epsilon$
    end for
    $w \leftarrow w * \epsilon$
    $d \leftarrow D - w$
    $C, \mathbf{r} \leftarrow lpsolve(\mathbf{c}, \mathbf{s}, d)$
    $S \leftarrow \sum_i^N r_i * s_i$
    if $d - S > 0$ **then**
        $w \leftarrow w + (d - S)$
    end if
    for $i = 1 \rightarrow N$ **do**
        $olds_i \leftarrow (!r_i) * s_i$
    end for
end for

Next, we extrapolate the supply that could be available from stores not currently donating. We include all 114 stores in the CFS distribution area. We examine how cost changes as a function of donating stores by simulating a range of demand goals. We also examine the problem from the other angle — how supply changes for a fixed cost as the number of donating stores changes. Presumably, as more stores donate, the travel distance between donating stores decreases, which lowers the cost for pickups. In all cases, we also evaluate the role of ϵ in food supply by simulating an ϵ of 0 to 1. This includes the full range of food expiration time scales in our model. In these studies, we use a demand of 10,260 lbs per day, which is the demand from the CFS Feeding America service statistics and represents the lower bound on the food required to fully meet the worst case demand in the CFS service area.

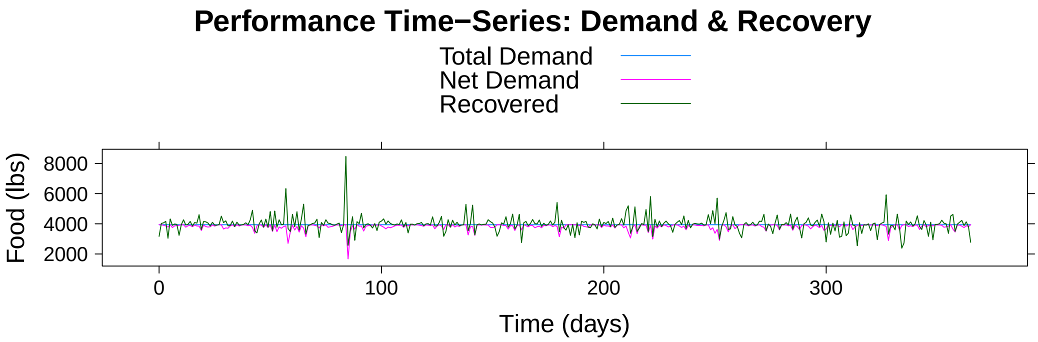
6 Results

Figure 7 provides a first look at the simulation-based results using parameters intended to reproduce the CFS model. Here, we set the daily demand to 3939 lbs, the mean received by CFS in our data set. The top plot shows the supply and demand. The total demand is fixed at 3939, the net demand is each day's demand reduced by the amount present in the warehouse, and the "recovered" line is the amount recovered for each of 365 days. On the majority of days, the demand is met. The bottom plot shows the distribution of excess, where a positive excess reflects an overrage and a negative excess reflects a shortage in food recovered. The excess appears to be normally distributed about the mean of 42.96 lbs, indicating that the recovered food from donors was just slightly higher than the demand. Finally, the middle plot shows the cost on each day and reports the mean cost of 231.23 (kilometers driven). CFS estimates their daily driving distance (sum of three vehicles) to be 212 km, within 10% of the value predicted by our model [32].

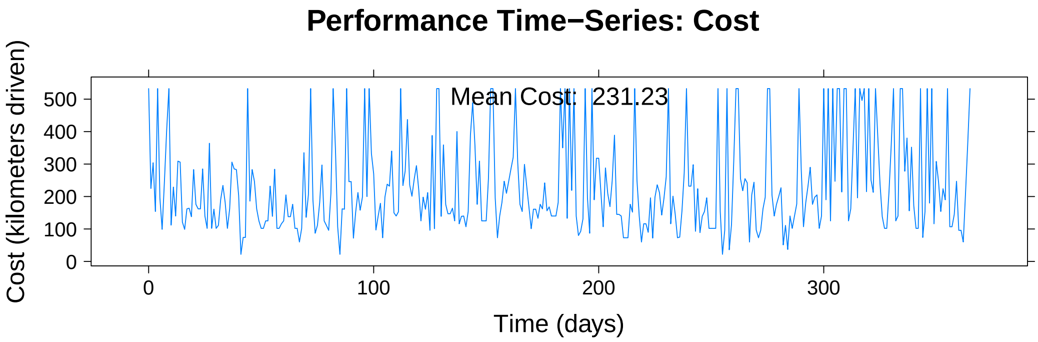
when we raise the target demand from the amount CFS is currently receiving, to the higher goal of 10,260 lbs/day, using the same $\epsilon = 0.5$, underruns are the norm rather than the exception, as shown in Figure 8. A mean cost of 520 km is necessary in order to provide approximately 7,000 lbs/day, which is an average of 3,376 lbs shy of the demand. However, this example leads to another important feature of the model, the ϵ parameter. Figure 9, which shows the same scenario, but with $\epsilon = 0.8$, indicating that food expires at a rate of 20%, instead of 50% as above. In this case, demand is met on most days and the mean excess is -89.43, or in other words, on average there is only a shortage of 89 lbs from the higher demand of 10,260 (nearly a three-fold increase from current supply).

Figure 11, shows this relationship explicitly by plotting the number of days in a 365 day simulation where supply failed to meet demand as a function of the ϵ value. For small ϵ values, the number of underruns is very high, but as ϵ is increased, meeting the higher demand becomes obtainable and the shortages are mitigated. Clearly, the effect of ϵ is not to be underestimated. In a practical sense, this highlights *the importance of proper food storage in the recovery process to smooth supply and mitigate shortages.* Those involved in retail food distribution are well aware of the importance of keeping food in the optimal conditions from farm to shelf, but these results show that it is also an essential consideration in the viability of a food recovery model. For this reason, many food banks focus on non-perishable food items during large fundraising events and food drives. This can be understood in the terms of our model: canned goods have an epsilon near 1.0.

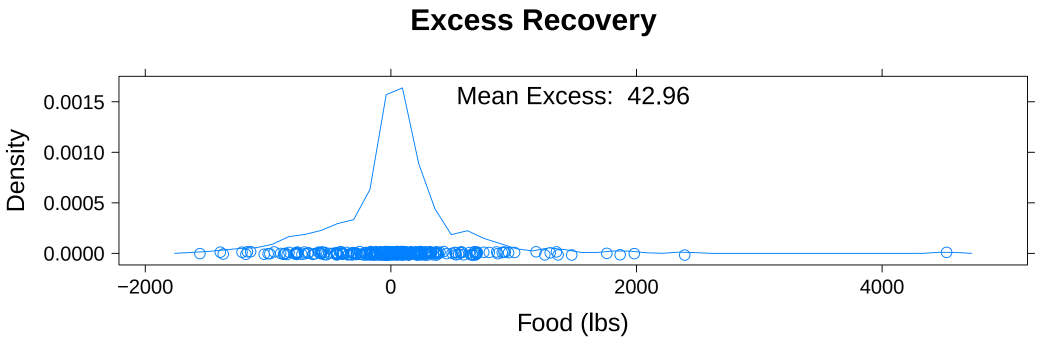
A final consideration of the basic model is intentional overage. Figure 10, shows the same scenario as above, with $\epsilon = 0.5$, but purposely attempting to retrieve 20% more food each day than is actually required. The logic here is that we can capitalize on freak supply "floods" to prepare for potential "droughts" in supply. We can see that in this case, there is a marginal improvement in mean excess, and actually a small increase in cost. From this, we're able to say that this strategy is not an effective one—*picking up extra food now cannot protect from shortages in the future.* Presumably this stems from the fact that food expires at the same rate whether it is left at the supplier, or transfered to the warehouse. In the case where that is not true (for instance, the supplier does not have room to store expiring "waste" for pickup and discard anything not picked up immediately), then this strategy may be more effective.



(a) Demand and Recovery

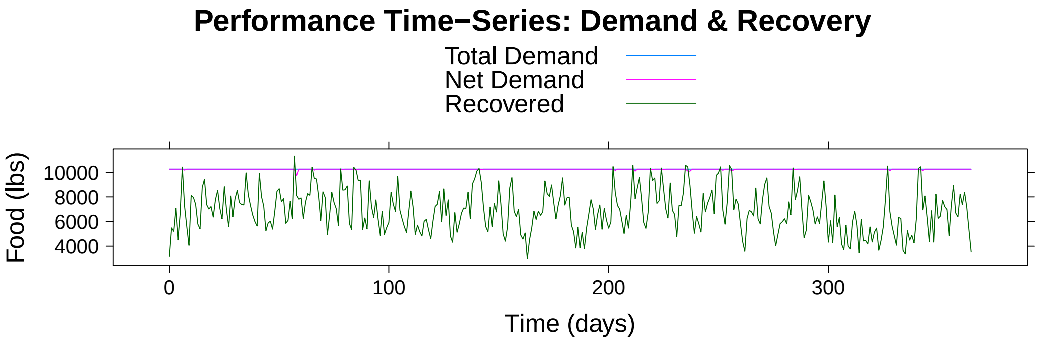


(b) Cost

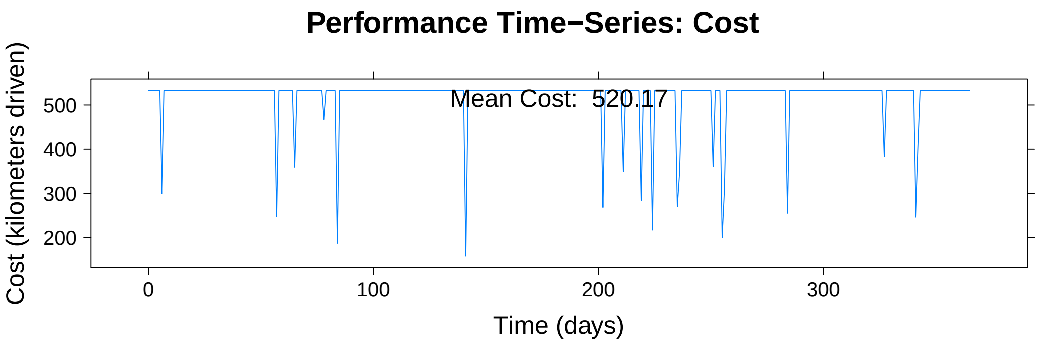


(c) Excess

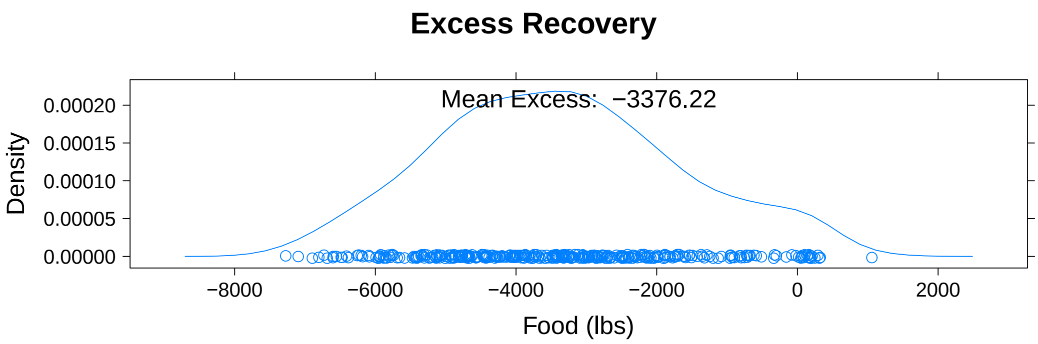
Figure 7: Performance time-series for one year simulation using the existing CFS donors, $\epsilon = 0.5$, and using a central warehouse. Excess is the difference between net demand and recovered. Hence, a positive excess indicates demand has been met.



(a) Demand and Recovery

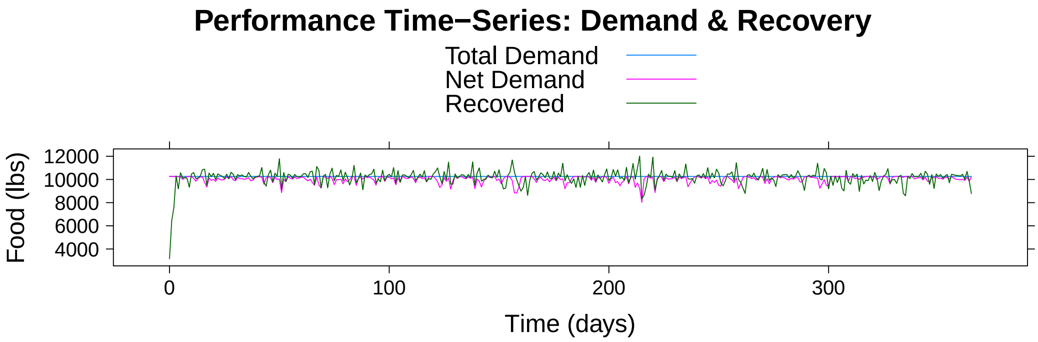


(b) Cost

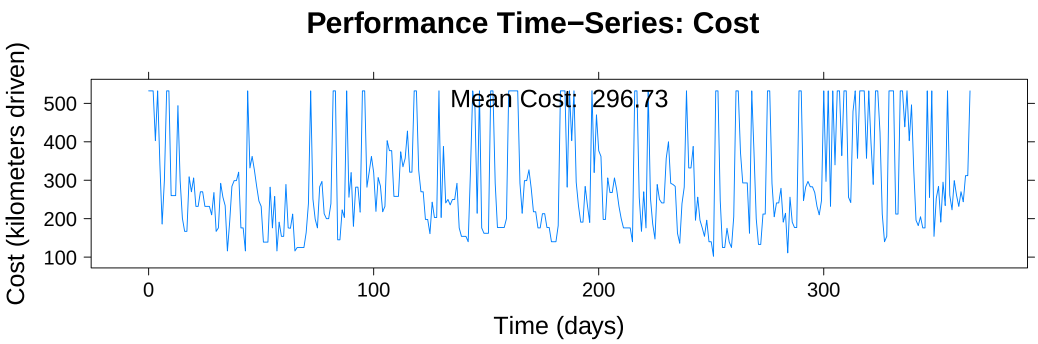


(c) Excess

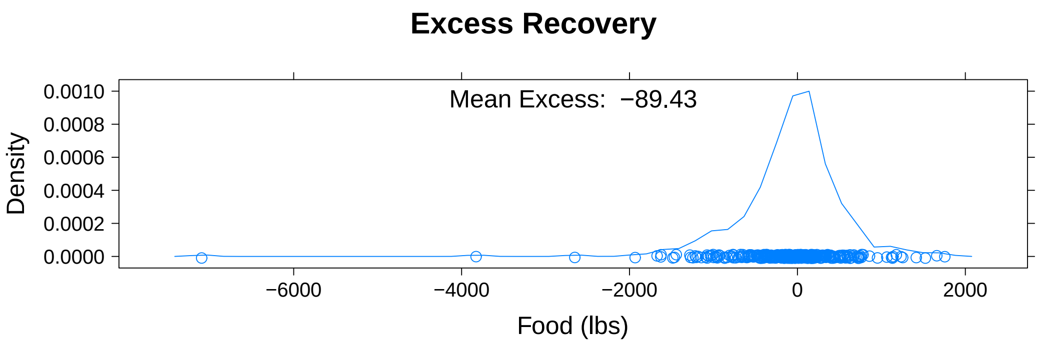
Figure 8: Performance time-series for one year simulation using the existing CFS donors, $\epsilon = 0.5$, and using a central warehouse. The target demand is set at the larger rescue threshold of 10,260. Excess is the difference between net demand and recovered. Hence, negative excess values indicate a shortage.



(a) Demand and Recovery

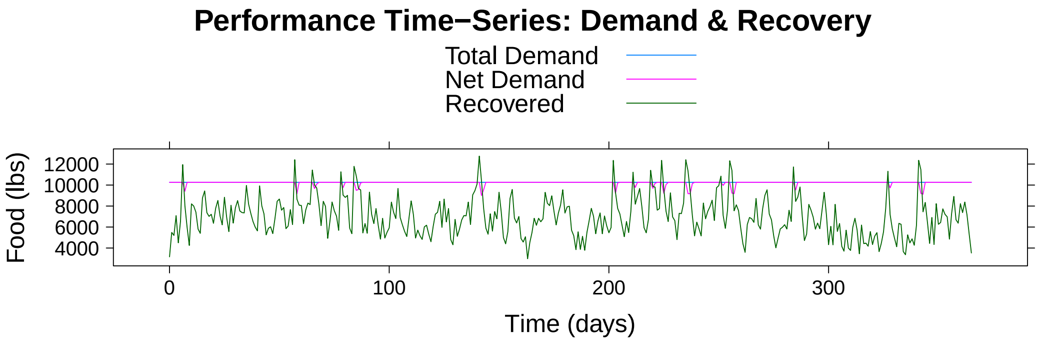


(b) Cost

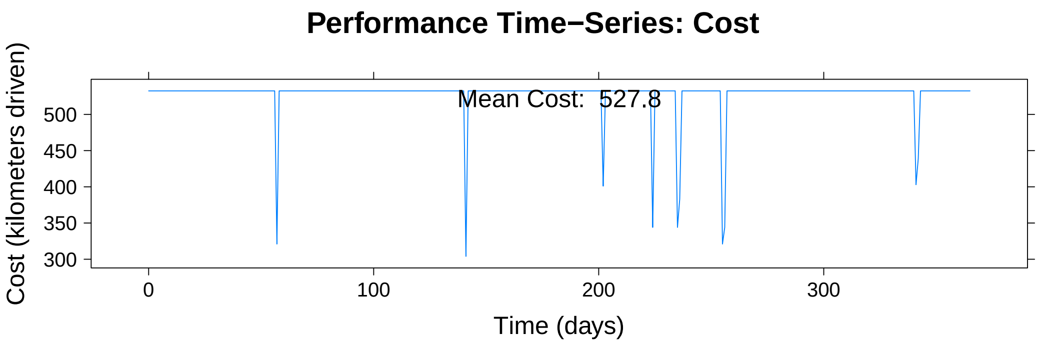


(c) Excess

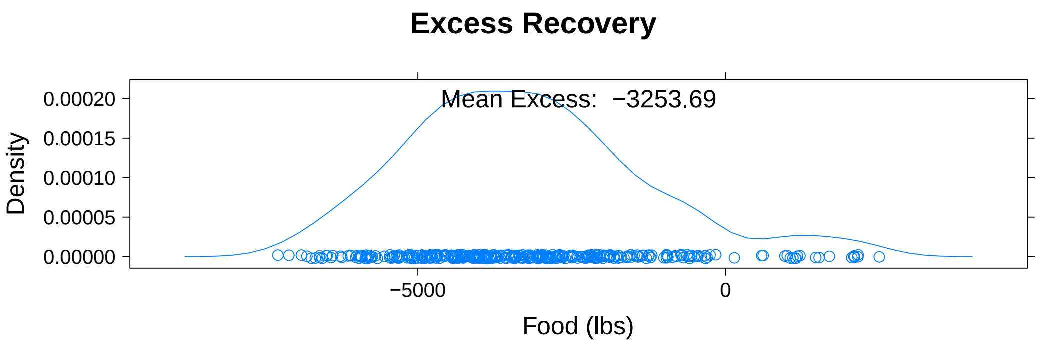
Figure 9: Performance time-series for one year simulation using the existing CFS donors, $\epsilon = 0.8$, and using a central warehouse. The target demand is set at the larger rescue threshold of 10,260. Excess is the difference between net demand and recovered. Hence, negative excess values indicate a shortage.



(a) Demand and Recovery



(b) Cost



(c) Excess

Figure 10: Performance time-series for one year simulation using the existing CFS donors, $\epsilon = 0.5$, and using a central warehouse. The target demand is set at the larger rescue threshold of 10,260. Excess is the difference between net demand and recovered. Hence, negative excess values indicate a shortage.

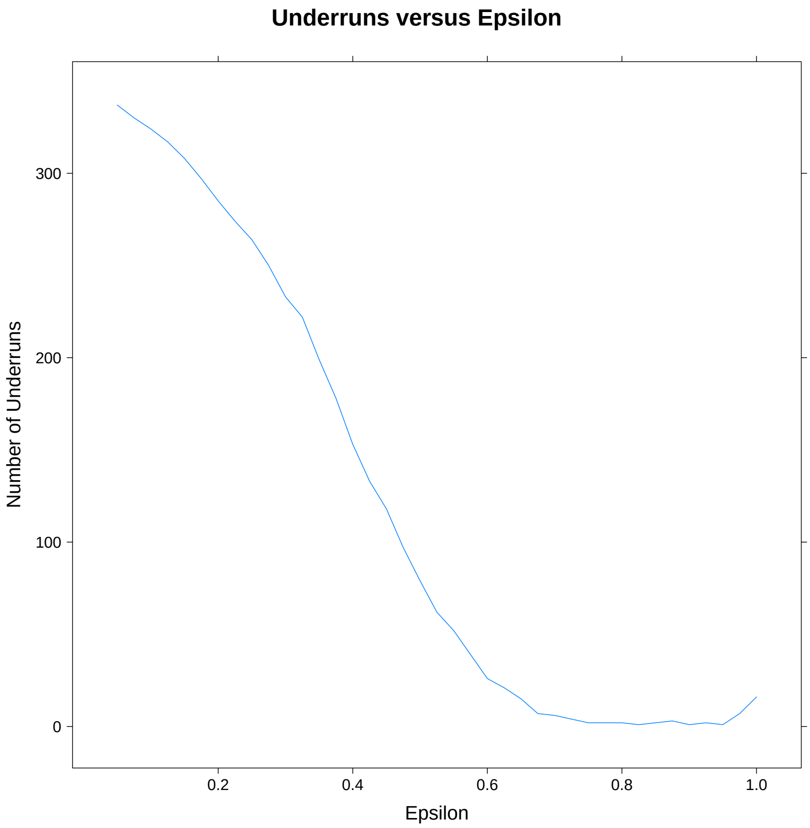


Figure 11: Number of underruns (for 365 day simulation) as a function of the epsilon value. Results are provided with a central warehouse using the set of 46 suppliers.

Next, we add in the rest of the suppliers that are not currently donating food, but could be, increasing the total set of suppliers from 46 to 114. Figure 12 shows the cost vs. food recovery relationship for both the set of 46 suppliers and the set of 114 suppliers. In both plots, there is a saturation point, where there is no more recoverable food. For the set of 46 suppliers this maximum is around 7,000 lbs as we saw above. For the larger set of suppliers, the maximum is closer to 14,000 lbs. There are two conclusions that can be drawn from this result. First, *despite the complexity of the underlying model, the relationship between the amount of food available for rescue and the number of donors participating appears to be linear.* Second, and perhaps more importantly, with sufficient resources, and more participating donors, CFS may be able to comfortably meet their current worst-case demand. To meet this goal, they would only need to drive approximately 600 km a day, which is just less than three times their current expendature. This indicates that *provided sufficient funding is available, and a large number of businesses are participating as donors, the food rescue model can successfully feed the area's hungry using only food that would otherwise be wasted.* Admittedly, the demand of 10,260 lbs is well below the gold standard of 48,600. For that to succeed, according to our model, CFS would need to have sufficient resources to drive approximately 11,000 km per day.

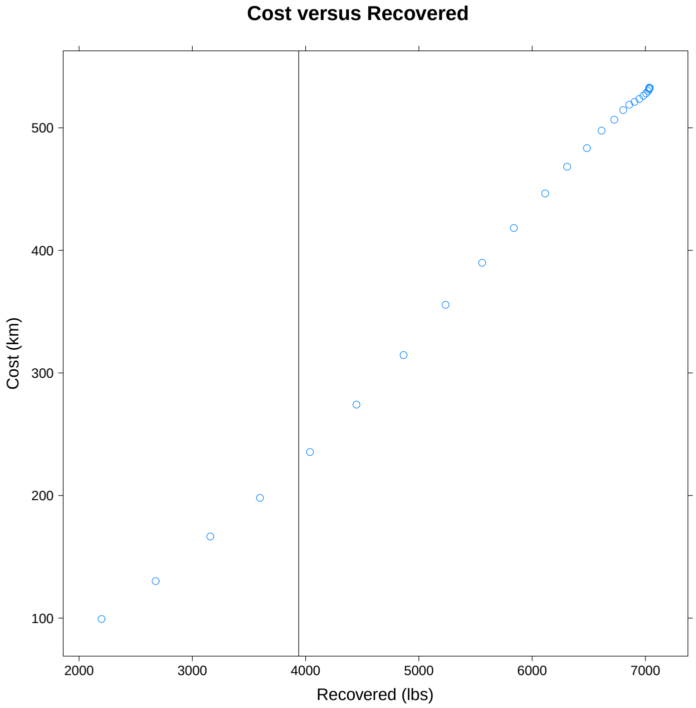
The other interesting thing to note about Figure 12 is shape of the curve seems to be linear for fewer suppliers, and trend towards exponential as the number of suppliers increases. We hypothesize that this is because the greater number of suppliers in the 114 case increases the densities of stores and allows for cost-saving optimizations when the demand goal is small. But, as the demand increases towards the saturation point, there are simply more stores that must be visited, driving up cost quickly. Figure 13 reinforces this point further. In this figure the cost is plotted as a function of the demand goal for each set of suppliers. We can see that up until the crossing point, the larger set of suppliers is able to satisfy the same demand with a smaller cost. This indicates that *the cost of the food redistribution problem can be reduced simply by increasing the number of particpating donors.* These plots plateau at the point that they are no longer to fulfill demand (and hence, have a horizontial asymptote at the cost of visiting all donors).

7 Conclusion

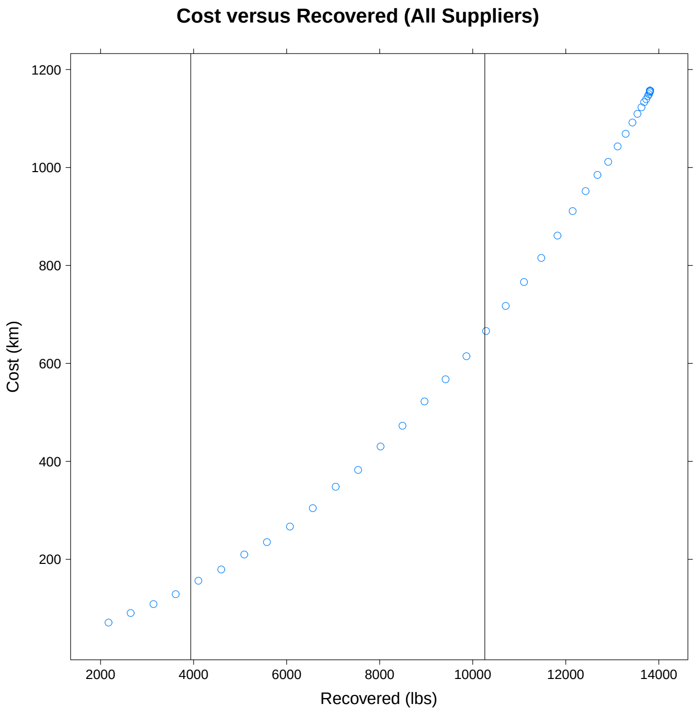
In this paper we have provided the first formal investigation of the food recovery and redistribution problem. To this end, we have developed a novel model that can be used for Monte Carlo style simulation using fitted empirical parameters. For our own experiments, we have made use of data from a large food bank in northern central Colorado. We have shownt that this model is able to reproduce the dynamics of the fundamental food redistribution problem faced by CFS. While we believe that this data is representative of a large class of similar regions with a mix of rural and urban environments, we are careful to remind the reader that our conclusions may not apply in disparite environments (i.e., dense urban or sparse rural). In future work, we hope to integrate additional data in order to broaden our conclusions and perform a more rigorous validation of our model.

Our chief experimental findings in this work are:

- Food supply (waste) events are heavy-tailed and can be well modeled with extreme value theory "peaks over threshold" models and the generalized Pareto distribution.
- The efficacy of the food redistribution approach hinges on the ability to keep rescued food from perishing. Hence, refridgerated storage and timely transportation and processing are crucial to the success overall.
- Despite the underlying heavy-tailed process and complexity of the model, the basic obtainable supply appears to be a linear function of the number of participating donors. Hence, doubling of the number of participating donors is likely to double the amount of food available.
- The overall cost of food recovery can be reduced substantially, simply by increasing the number of participating donors (and therefore creating more opportunity for food supply events to occur, when they are required by demand).



(a) 46 Suppliers



(b) 114 Suppliers

Figure 12: Cost of recovering food for each of two sets of suppliers. The vertical lines indicate demand reference points: 3939 lbs, and 10,260 lbs.

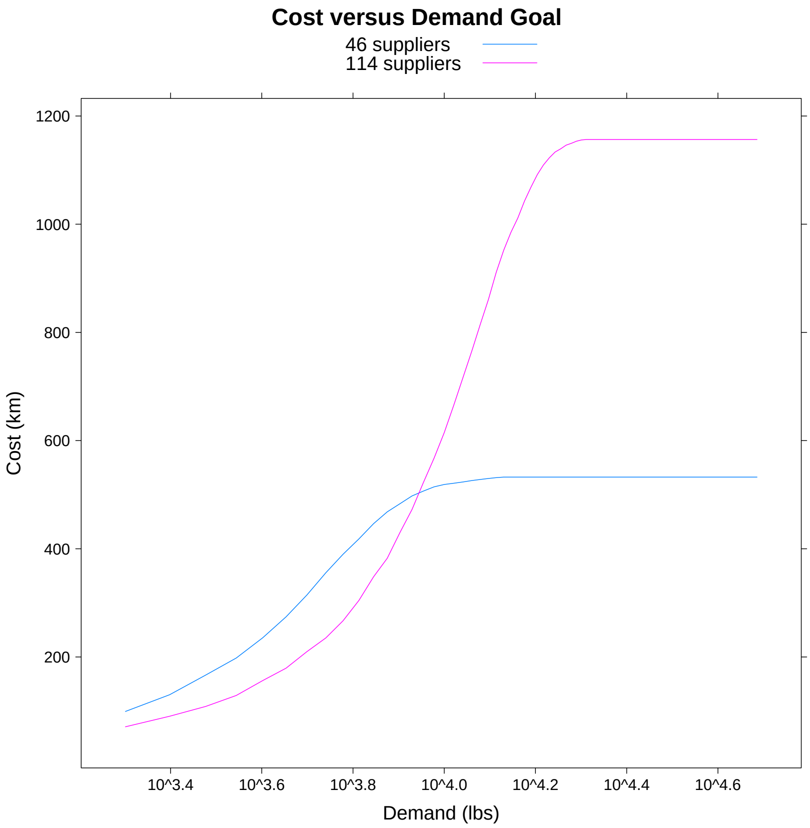


Figure 13: Relationship between cost and demand goal for each of the supplier sets.

- At least in the scenario we have studied, using the data from CFS, we have shown that the food redistribution approach is a feasible method of both subtantively reducing waste and hunger. There is room for growth within the current paradigm, but it requires an increase in available funding.

In future work, we will expand this work to address some additional questions. In particular we are interested in the question of whether this model can be scaled up to a state-wide or national level. It stands to reason that dense urban and sparse rural environments will produce substantially different cost and supply dynamics. However this is an open question. In 2008, there were approximately 85,200 grocery stores in the US [24], which works out to 0.3 chain supermarkets, 0.22 non-chain supermarkets, 3.04 grocery stores, and 1.8 convenience stores per zip code [28]. In this work, we haven't considered smaller potential donors such as restaurants. One reason for this is that CFS is a Feeding America partner, which means that they are unable to accept food donations that are not in their original packaging. Yet, this is a large potential food supply source: in [2], Bloom suggests that the typical food waste associated with a restuarant is on average 3000 lbs per employee per year (or 123 lbs/day for a 15 employee restaurant). There were approximately 566,020 food service places in the US in 2007 [3]. Clearly, there is no shortage of potential donors, the important question is whether they are well positioned for recovery and redistribution and whether the cost of rescue is acceptable.

An additional question is one of nutrition. In our current study, we have looked at bulk lbs of food without concern for the type. Clearly, this is a large simplification that has bearing both on the economics of the problem (supply and demand) as well as the basic expiry of the food. Currently, 88% of grocery stores donate some dry goods, 51% donate produce, 31% donate prepped food and meat [2]. Fresh and healthful foods are hardest for food banks to get since they have a limited shelf life (small ϵ), which are negatively impacted by transportation and stocking time, and pickup limitations (how many pickups per week are possible). By understanding the nature of the problem, optimizing pickup strategies, and sufficiently funding food rescue organizations so that they have the resources to pickup food when it is available, this problem might be mitigated. Feeding America's refridgerated trucks are a good start, but cannot captialize on smaller food waste events, which contribute substantively to the efficacy of the model as a whole. A complete solution may require rescue and redistribution at multiple scales and with varying technologies.

Although preliminary, our work here is an important first step towards understanding the dynamics and limitations of the food rescue and redistribution problem. By formulating the problem for optimization and studying it via simulation, we have been able to draw out fundamental aspects of the underlying problem. In the end, we can present the positive result that this approach to food rescue and redistribution, despite its underlying complexity, can be considered a stable process where obtaining additional food is simply a function of having sufficient participating donors and funding to perform pickups. We hope that this work will help to spur interest in the area, equally among researchers who might be able to bring additional insight into the problem, businesses who can agree to donate their food waste, and policy makers who posess the ability to procure needed funding and resources for food rescue organizations to succeed.

Acknowledgements

We would like to thank Community Food Share, and in particular Tom Reed, who generously provided us their data for analysis and worked with us on reformatting it to suit the needs of this study. We would also like to thank the numerous food rescue organizations and their tireless volunteers, whose work helped to inspire this research.

References

[1] lp solve reference guide. http://lpsolve.sourceforge.net/, July 2011.

[2] J. Bloom. *American Wasteland: How America Throws Away Nearly Half of Its Food (and What We Can Do About It).* Da Capo Lifelong Books, 2010.

[3] U. C. Bureau. 2007 economic census: Sector 72: Accommodation and food services: Industry series: Preliminary summary statistics for the united states: 2007. http://factfinder.census.gov/servlet/IBQTable?_bm=y&-geo_id=&-ds_name=EC0772I1&-_lang=en, July 2011.

[4] I. Christou, A. Zakarian, J.-M. Liu, and H. Carter. A two-phase genetic algorithm for large-scale bid-generation problems at Delta Air Lines. *Interfaces*, 29(2):51–65, 1999.

[5] C. City of Boulder. emaps resources. http://gisweb.ci.boulder.co.us/website/pds/pds_eMapLink/, April 2011.

[6] S. Coles. *An Introduction to Statistical Modeling of Extreme Values.* Springer, 2001.

[7] T. Economist. The 9 billion-people question: A special report on feeding the world. *The Economist*, February 2011.

[8] L. FLeischer and K. Wayne. Fast and simple approximation schemes for generalized flow. *Mathematical Programming*, 91(2):215–238, 2002.

[9] M. Gorman. An application of genetic and tabu searches to the freight railroad operating plan problem. *Annals of Operations Research*, 18:51–69, 1998.

[10] K. D. Hall, J. Guo, M. Dore, and C. C. Chow. The progressive increase of food waste in america and its environmental impact. *PLoS ONE*, 4(11):e7940, 11 2009.

[11] J. Hartigan and M. A. Wong. A k-means clustering algorithm. *Applied Statistics*, 28:100–108, 1979.

[12] C. Harvest. Rescuing food for new york's hungry. http://www.cityharvest.org/, April 2011.

[13] F. S. Hillier and G. J. Lieberman. *Introduction to Operations Research.* McGraw Hill, New York, New York, 2005.

[14] T. Holloran and J. Bryn. United airlines station manpower planning system. *Interfaces*, 16(1):39–50, 1986.

[15] M. P. R. Inc. Hunger in america 2010 / community food share. Technical report, Feeding America, January 2010.

[16] M. P. R. Inc. Hunger in america 2010: National report prepared for feeding america. Technical report, Feeding America, January 2010.

[17] S. G. Inc. Helping feed the hungry for 34 years. http://www.seniorgleaners.org/, April 2011.

[18] T. W. Jones. The corner on food loss. *BioCycle*, pages 2–3, July 2005.

[19] L. Kantor, K. Lipton, and A. Manchester. Estimating and addressing america's food losses. *Food Review*, pages 2–12, 1997.

[20] MapQuest. Mapquest directions api service. http://developer.mapquest.com/web/products/dev-services/directions-ws, April 2011.

[21] M. Nord, A. Coleman-Jensen, M. Andrews, and S. Carlson. Household food security in the united states. Technical report, United States Department of Agriculture, 2009.

[22] U. S. D. of Agriculture. Food security in the united states: Definitions of hunger and food security. http://www.ers.usda.gov/Briefing/FoodSecurity/labels.htm, April 2011.

[23] U. S. D. of Agriculture: Economic Research Service. Loss-adjusted food availability. http://www.ers.usda.gov/Data/foodconsumption/FoodGuideSpreadsheets.htm, April 2010.

[24] U. S. D. of Labor: Bureau of Labor Statistics. Career guide to industries, 2010-11 edition: Grocery stores. http://www.bls.gov/oco/cg/cgs024.htm, July 2011.

[25] J. Oldham. Combinatorial approximation algorithms for generalized flow problems. *Journal of Algorithms*, 38(1):135–169, 2001.

[26] C. Papadimitriou. *Combinatorial Optimization: Algorithms and Complexity*. Prentice-Hall, Englewood Cliffs, New Jersey, 1982.

[27] Philabundance. Driving hunger from our communities. http://www.philabundance.org/, April 2011.

[28] L. M. Powell, S. Slater, D. Mirtcheva, Y. Bao, and F. J. Chaloupka. Food store availability and neighborhood characteristics in the united states. *Preventive Medicine*, 44:189–195, 2007.

[29] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2011. ISBN 3-900051-07-0.

[30] A. Roy, E. DeFalomir, and L. Lasdon. An optimization-based decision support system for a product mix problem. *Interfaces*, 12(2):26–33, 1982.

[31] C. F. Share. http://www.communityfoodshare.org/, April 2011.

[32] C. F. Share. Personal electronic communication, February 2011.

[33] T. Sharkey and H. Romeijn. A simplex algorithm for minimum-cost network-flow problems in infinite networks. *Networks*, 52(1):14–31, 2008.

[34] T. Stuart. *Waste: Uncovering the Global Food Scandal*. W.W. Norton and Company, Inc., 2009.

[35] X. Tang and R. T. M. Wong. Minimizing wire length in floorplanning. *IEEE Transactions on Computer-aided Design of Integrated Circuits*, 25(9):1744–1753, 2006.