ESTIMATING SURPLUS FOOD SUPPLY FOR FOOD RESCUE AND DELIVERY OPERATIONS

Divya J. Nair, corresponding author
School of Civil and Environmental Engineering, University of New South Wales,
Sydney, NSW 2052, Australia,
d.jayakumarnair@unsw.edu.au, Phone: +61 470 442 793 Fax: (+61 2) 9385 6139

Taha Hossein Rashidi
School of Civil and Environmental Engineering, University of New South Wales,
Sydney, NSW 2052, Australia,
rashidi@unsw.edu.au, +61 (2) 9385 5063

Vinayak V. Dixit
School of Civil and Environmental Engineering, University of New South Wales,
Sydney, NSW 2052, Australia,
v.dixit@unsw.edu.au, (+61 2) 9385 5381

Abstract
Hunger remains a largely hidden social problem in many developed nations. The not-for-profit food rescue organisations aids in alleviating hunger, by rescuing the surplus food from different food providers and re-distributing to people in need. However, surplus food donation is a random process which varies with regard to quantity, time and place. Understanding the dynamics of food recovery and forecasting food donations using historical information has significant importance in inventory management and redistribution, particularly in reducing operational costs and achieving a sustainable and equitable distribution of inventory incorporating uncertainties in supply. This paper uses different modelling techniques including multiple linear regression, structural equation modelling and neural networks to explore the patterns and dynamics of food donation and distribution. A set of significant indicators has been identified to describe the current food donation process, to predict daily average food donated by different food providers and also to anticipate the potential donation from a new donor which may appear in the network in the future. Results suggest that structural equation modelling and neural networks provide improved demand estimation when compared to conventional multiple linear regression. We also discuss the usefulness of these models in sustainable and equitable management of food recovery and redistribution.

Keywords: Relief logistics, Sustainable food recovery and redistribution, forecasting surplus food donation, artificial neural network, structural equation model, multiple linear regression

1 Introduction

Although, most of the industrialized and developed countries produce enough food to feed themselves and the rest of the world, millions of people live with very low food security. Australia, being one of them, produces enough food to feed 60 million people, almost twice as their current estimated population [1]. However, recent research indicate that more than 1.9 million tonnes of food is discarded from the manufacturing and retail sector into landfill each year, and over 2 million people have low food security [2]. Most of the food they throw away is avoidable and could have been consumed if it had been managed better. This food waste and food insecurity problem is tackled by an ever-growing number of food rescue organizations (Foodbanks, Food Rescue, OzHarvest, Secondbite, etc.) which collect surplus food from different food providers and redistribute it to welfare agencies supporting various forms of food relief. Foodbanks are not-for-profit organizations which act as a pantry to the charities and community groups that feed the hungry. They rescue food products, including perishable goods, incorrectly labelled items, etc., from different local sources such as farmers, manufacturers and retailers. These food products are then stored in warehouses, sorted, packed and sometimes processed before being delivered to welfare agencies or to specific delivery points, accounting for the perishability of the products and the requests of agencies. In essence, they function as aggregators and distributors of surplus food rescued from various sources. There are many other food rescue organizations that collect food from these food Foodbanks and different food providers, including groceries, supermarkets, cafes, farmers, wholesalers, small vendors, restaurants, etc., and directly deliver at no charge to agencies providing assistance to vulnerable men, women and children. Due to the perishability of food products collected, they are not stored in the warehouses, but are instead delivered on the same day itself. They operate trucks that visit food providers and agencies daily. The trucks start from a depot, collect food from food providers and deliver it to agencies, before returning back to the depot empty. The sequence of visits to agencies and food providers is determined based on the location of food providers and agencies, the quantity of food rescued and the demand of agencies. The frequency of visits of agencies during a week depends on the average daily availability of different categories of food. The efficiency of these food recovery operations depends on effective utilization of the recovered food with minimum wastage. *A major reason for wastage and inequitable distribution is that the quantity and category of food donated is unknown until observed upon the driver's arrival. The aim of the study is to analyze food donation data to help food rescue organizations to deal with this uncertainty.*

To address the uncertainty issues, it is important to understand how it affects the logistics and operation of food rescue operations. Unlike most logistical organizations, the operations of not-for-profit food rescue organizations are not solely cost driven. These organizations operate in the social interest and are therefore governed by fairness and equity considerations. Each welfare agency has a request (demand), a single product type or a combination of different product types, which is a function of the type of food assistance they provide (breakfast program for kids, community kitchens, food parcels, etc.), the size of the agency (number of people they support), frequency of service, etc. Ideally, agencies should determine and communicate their demand in advance so that food rescue organizations can effectively design the routes and equitably allocate the limited surplus food among the agencies. However, the type and quantity of food available at each food provider is unknown until observed upon the driver's arrival. In practice, in the absence of supply information, the decision maker designs initial routes minimizing the operational cost and the drivers make critical decisions regarding the delivery of rescued food to the agencies. Generally, the driver is expected to: (i) satisfy the agency's demand while reserving the supply for the other remaining agencies on the route and (ii) re-design the initial routes if the food available (different categories) at a food provider is insufficient to meet the demand of the agencies. Thus, uncertainty in the supply often leads to re-routing, higher operating costs, wastage of rescued food and unfair allocation of food. Hence, addressing these concerns are the major goals of many non-profit food rescue organizations.

The aims of the study are to 1) forecast the supply of different categories of food at different food providers, 2) identify parameters explaining the quantity of food supply and 3) investigate the applicability of different forecasting methods to predict the quantity of different types of food supply per day at each food provider. Understanding the pattern and availability of donated food is crucial in order to effectively plan and manage vehicle routes, and allocate different categories of food equitably among welfare agencies. The forecasting models developed in this study would better equip decision makers in anticipating the food availability, well before a journey starts. This would facilitate efficient operations and help in bringing down the operational costs incurred by food rescue organisations. Additionally, estimating the average daily availability of different categories of food would enable decision makers to understand the underruns (supply is less than required) and overruns (supply is greater than required). Thus food rescue organisations could effectively design the frequency of visit and schedule (assign into a particular day) of visit matching the supply and demand and minimizing waste. The model incorporates information related to the observable characteristics of food provider, such as type, size, region (land-use, population and area) and day of donation (Weekday/Weekend). Several modelling specifications have been employed in this study, including Multiple Linear Regression (MLR), Structural Equation Modelling (SEM) and two artificial neural networks, namely, Feed-Forward backpropagation Neural Network (FFNN) and Generalized Regression Neural Network (GRNN). These models are used to estimate the average food donated per day per category per food provider using historical data provided by OzHarvest, Sydney. FFNN and GRNN are two groups of Artificial Neural Networks (ANN) that perform differently based on input variables. Both have their own advantages and disadvantages. While FFNN is sensitive towards the neuron interconnection weights and local minimum, GRNN gives a better approximation when the input variables are continuous. Another advantage of GRNN is fast learning and convergence to the optimal solution as the sample size increases [3]. In the proposed study, along with comparing the estimates of (ANN), MLR and SEM, we intend to identify the best neural network approach for demand forecasting by considering two different types of ANN.

The remainder of the article is structured as follows. Section 2 illustrates the background and a brief literature review. The datasets used in this study are then explained, and the explanatory variables are discussed in Section 3. Section 4 presents a brief description of different forecasting methods. Experimental results from different models are discussed and compared in section 5. Conclusions and future research directions are discussed in the final section.

2 Background

There is considerable relevant literature discussing the role of forecasting techniques in estimating future demand using historical data in various domains. While most of them focus on areas like transport planning, supply chain management, weather forecasting, sales forecasting, economic forecasting, etc., very few discuss the use of forecasting techniques in estimating blood donation demand and supply [4,5] potential organ donation [6] and scarce resource consumption [7–9]. Despite its wide applicability, forecasting models received little attention in food rescue operations. While the recent few studies focus on optimizing collection and delivery schedules [10–14] and equitable allocation of rescued food [15,16], few studies addressed the need of forecasting the donation amount. Lien et al. 2014 [16] proposed a resource allocation model for a food rescue organisation in Chicago, for effective and equitable allocation of rescued food, considering an egalitarian welfare utility function as an indicator of equity. They compared the performance of their allocation model in the case of uncertain supply with the case where all the supply are known prior to routing and found that the model performance, in terms of maximizing equity and minimizing wastage, can be improved if the supply is known prior to routing.

Phillips et al. 2013[17] proposed an empirical model to estimate the total quantity of food rescued by Food Bank in north central Colorado. The authors described the food donation process using a peak over threshold model, where the events greater than zero were modeled using a Generalized Pareto distribution. The surplus food donated by food providers was modeled as a function of their type (grocer, manufacturer, individual and farm), and size. Their study focused primarily on understanding the gap between demand and supply and strategies to improve the total food rescued. However, they considered only the total amount of food rescued, rather than looking at the nutritional value and category of food rescued.

Davis et al. 2013 [18] analyzed the food rescue operations of Food Bank of central and eastern North Carolina. They discussed the use of time series forecasting techniques, moving average and exponential smoothing to forecast the amount of food donated per category per donor type. The results suggested that exponential smoothing provides a reasonable approximation of food donation compared to the moving average method. Jiang et al. 2013 [19] extended this study by exploring different data mining techniques to study the pattern of donation, impact of frequency of donation by a donor in the total amount of donation, the trend in donation and stochasticity in donation (using Markov Chain analysis). However, these models were limited to estimating average monthly food availability. Recently, Brock and Davis (2015) [20] extended this study by estimating the average daily donation using different forcasting techniques. They compared the estimates of the traditional forecasting method MLR with the data mining approach of multi-layer perceptron neural network (MLP-NN) in predicting the amount of food received from a supermarket by the Food Bank of Central and Eastern North Carolina [20]. They focused mainly on donations from the supermarket, assuming donations were a function of supermarket sales. These were dependent on the purchasing power of community, financial wellness, unemployment, past donation, frequency in donation, week of the year, weekday, etc. Their results suggest that

MLP-MNN models outperform the conventional MLR models. However, these models were limited to estimating the average daily food available at supermarkets and could not identify methods to estimate the average daily food availability at food providers with different characteristics. Although Phillips et al. 2013 and Davis et al. 2013 addressed donations from different food providers, they were limited to total donation amount and the average monthly donation amount respectively. In this paper we evaluate three approximation methods on their ability to estimate the average daily availability of different categories of food at different food providers. Additionally, we propose a second generation multivariate modelling technique, SEM, that enables the model to account for the correlation between different types of food donated.

3 Data Analysis and Variables

In this section, we describe the data and the variables used for the study. The historical data used in the study was provided by OzHarvest, one of the largest food rescue organizations in Sydney. The study area is shown in Fig. 1. OzHarvest was the first perishable food rescue organization in Sydney, Australia, founded in 2004, that rescues 56 tonnes of surplus food every week from different food providers, including groceries, supermarkets, cafes, farmers, wholesalers, small vendors, restaurants, etc. The data includes food donation received from around 200 food providers between March 2014 to February 2014. The data consist of a total of 44,050 records. Each donation record indicates the day of donation, donation amount, category of food and details of contributing food provider. The different categories of food rescued are: (i) fruits and vegetables (fruits, vegetables, packaged salads, etc.), (ii) bread (bread, baked item, sandwiches, pancakes, etc.), (iii) dairy (cheese, egg, pudding, milk, desserts, pastries, etc.) (iv) drystock (breakfast bars, canned foods, snacks, pasta, rice, noodles, etc.), (v) cooked meals and (vi) others (meat products, frozen products, drinks, seasoning, etc.).

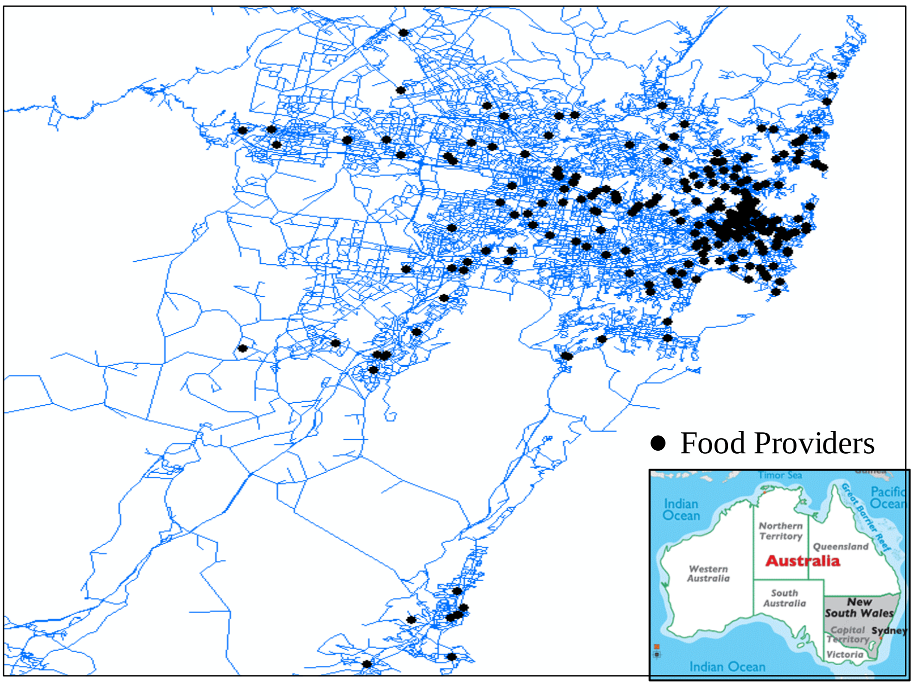
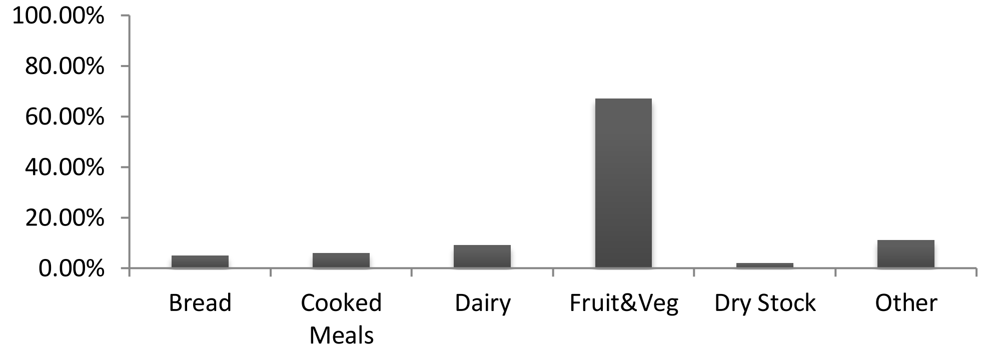


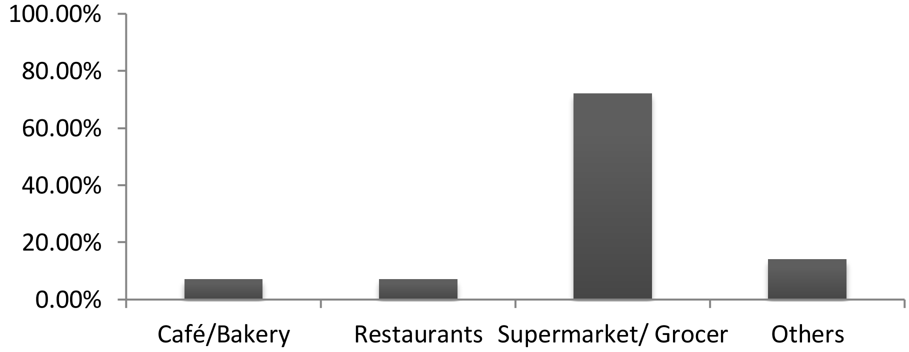
Fig 1. Study Area

3.1 Data Analysis

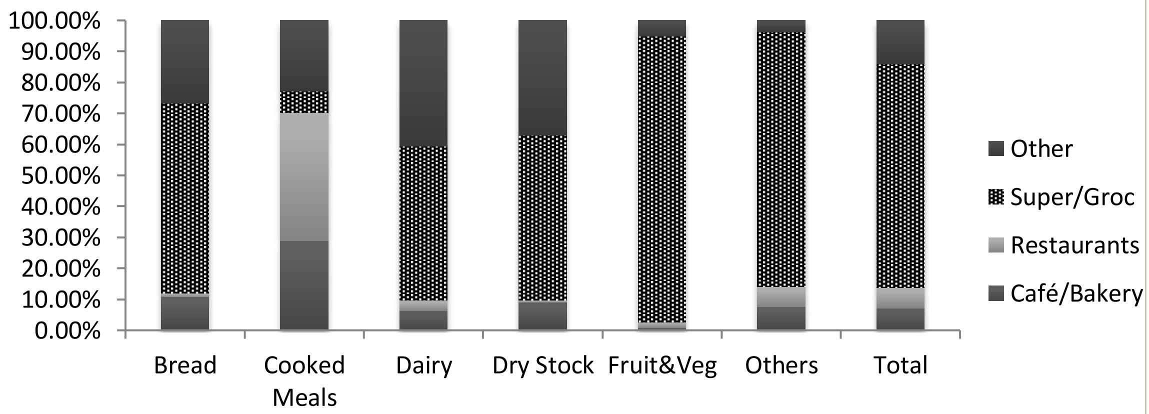
The different categories of food rescued and its statistics are shown in Fig. 2. From the data, the major portion of food donated is fruits and vegetables, accounting for 67% of the total food donated (Fig 2.a). While supermarkets and groceries donate the largest volume of fruits and vegetables, bread, raw meat and drystock, restaurants donate cooked meals and sandwiches, and cafes and bakeries donate dairy and desserts (Fig 2.c). Hence, it is expected that the quantity of different types of food donated depends on the type of food provider. In the case of food providers, the largest contribution comes from supermarket and groceries (Fig 2.b).



2. (a) Different categories of food donated



2. (b) Percentage food donation by different food providers



2. (c) Percentage food donation per category by different food providers

Fig 2. Summary Statistics of Food donation.

The average daily food donated (Fig 3.(a)) and the percentage of food providers (Fig. 3.(b)) during weekdays is about three times that of weekends. Since there is a big deviation in the number of food providers donating food and the quantity of food, we disaggregate the dataset into donations during weekends and donations during weekdays, and develop two different sets of models to forecast weekday and weekend donations.

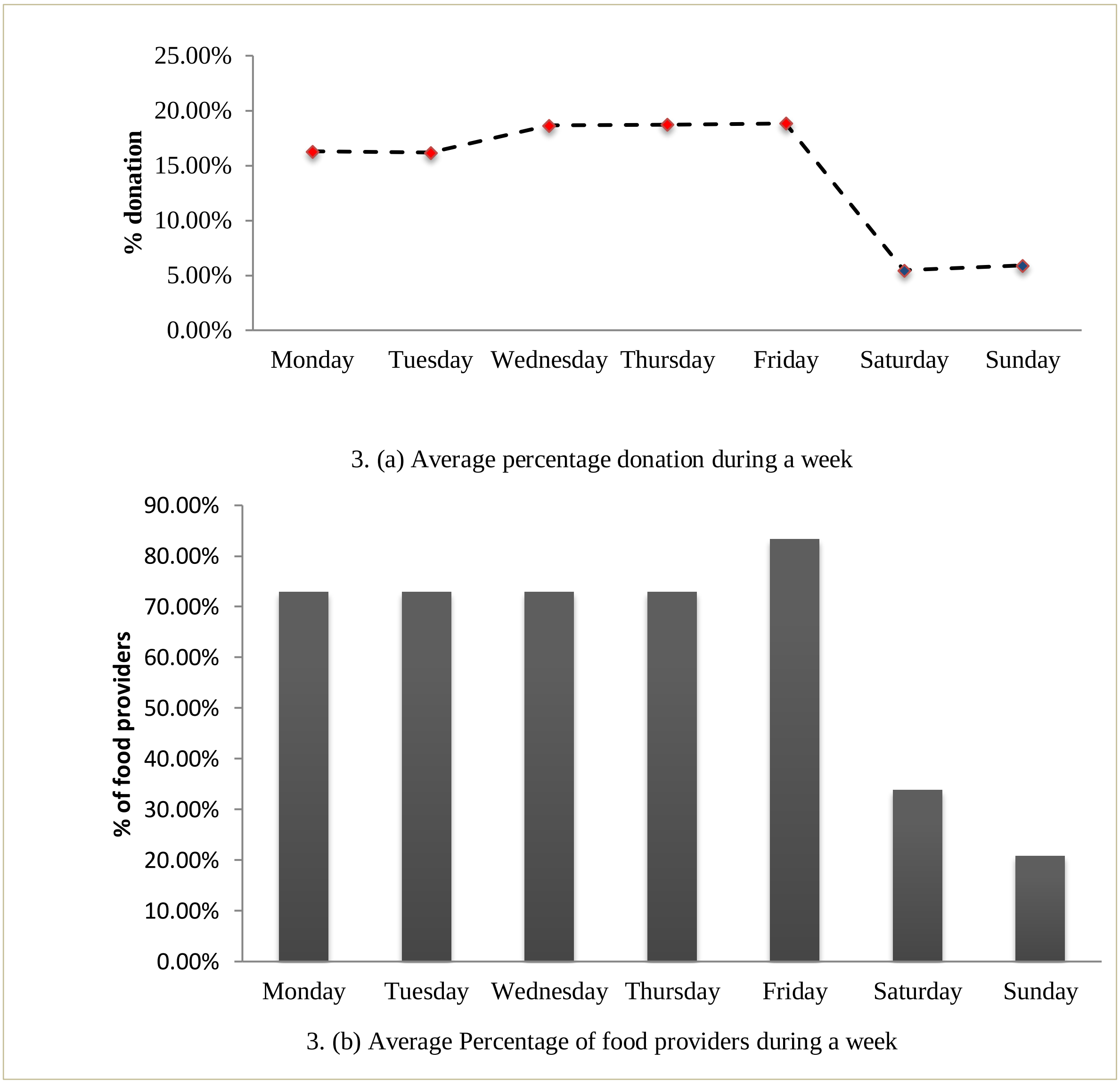


Fig 3. Average donation and food providers during a week

While some food providers donate surplus food once a week, some donate multiple times a week, depending on the type and size of food providers. Similarly, rescued food is delivered to the agencies once or multiple times a week depending on availability of the type of food they request. Understanding the availability of surplus food aids the food rescue organizations to decide the frequency of visits and efficiently group the food providers and agencies that need to be visited on a particular day.

3.2 Variables

The response variables and explanatory variables used in the multivariate forecasting models presented in section 4 are discussed below. Forecasting methods are used to estimate the (i) average total daily food donation from a food provider during weekends and weekdays and (ii) average daily donation of different categories of food from a food provider during weekends and weekdays. Donations are explained as a function of observable characteristics of the food provider and the day of donation. A summary of all response variables and explanatory variables are provided in Table 1. We assume that the amount of donation from a food provider depends on observable characteristics such as (i) type, (ii) size, and (iii) region. (i) Type is a categorical variable, with values such as Cafe/Bakery, Supermarkets/Grocer, Restaurants and Others. Since the data regarding the exact square feet area of food providers was not available for the study, we represented (ii) size as a categorical variable: small, medium and big (based on surface area). (iii) Region variables indicate the location of food providers in terms of land-use, population and area (Sqkm), where population and area are continuous variables and land-use is a categorical variable (urban industrial and commercial area, urban residential area, conservation area, transportation and other corridors and others). The descriptive statistics of the data used in Weekday and Weekend models are presented in Table 2.

Table 1 Variables Used For The Study

Category	Sub-category	Variable	Notation
Weekend-	Dependent Variables	Total Donation	total_w
		Fruit and Vegetables	fruveg_w
		Dairy products	dairy_w
		Bread	bread_w
		Dry stock	dry_w
		Cooked Meals	cmeal_w
Weekday-	Dependent Variables	Total Donation	total_wd
		Fruit and Vegetables	fruveg_wd
		Dairy products	dairy_wd
		Bread	bread_wd
		Dry stock	dry_wd
		Cooked Meals	cmeal_wd
	(i) Type of food provider	Cafe/Bakery	cafe_bak
Independent Variables		Supermarkets/Grocer	super_groc
		Restaurants	Restaurant
		Others	others_t
	(ii) Size of food provider	Small	size_s
		Medium	size_m
		Big	size_b
	(iii) Region: a. Land-use	Urban, industrial and commercial area	urb_com
		Urban residential area	urb_res
		Conservation area	Cons
		Transportation and other corridors	trans_corr
		Others	others_l
	b.	Population	Popu
	c.	Area	Area

Table 2 Descriptive statistics of the model variables

Descriptive statistics of Dependent Variables				
Weekday Model	Min	Max	Mean	Standard Deviation
Total_wd	0.23	230.31	118.20	27.91
Fruveg_wd	0	230.31	111.96	24.13
Dairy_wd	0	32.83	1.94	3.13
Bread_wd	0	24	1.78	2.45
Dry_wd	0	27.5	1.44	2.42
Cmeal_wd	0	28.33	3.06	4.3
Weekend Model				
Total_w	1.3	183.71	96.98	40.26
Fruveg_w	0	136.67	84.33	35.06
Dairy_w	0	5	1.47	0.94
Bread_w	0	24.29	2.93	2.71
Dry_w	0	26.57	1.8	3.89
Cmeal_w	0	17.43	3.79	1.88
Descriptive statistics of categorical independent variables				
Variable	Weekday Model	Weekend Model		
Type of food provider				
cafe_bak	24.50%	15.63%		
super_groc	44.80%	64.06%		
Restaurant	15.63%	15.63%		
others_t	15.10%	4.69%		
Size of food provider				
size_s	51.56%	31.25%		
size_m	13.54%	14.06%		
size_b	34.90%	59.06%		
Land-use				
urb_com	47.70%	39.38%		
urb_res	32.50%	38.75%		
Cons	9.38%	14.06%		
trans_corr	8.33%	6.25%		
others_l	2.08%	1.56%		
Descriptive statistics of continuous independent variables				
Weekday Model	Min	Max	Mean	Standard Deviation
Popu_wd	1178	88715	24187	14714
Area_wd	0.44	40.15	15.76	4.79
Weekend Model	Min	Max	Mean	Standard Deviation
Area_w	0.439125	36.16477	20.02359	6.303969
Popu_w	1178	88175.74	25176.49	88175.74

4 Forecasting Methods

In this section we describe the forecasting methods used to estimate food donation

4.1 Multiple Linear Regression

MLR is an approach to model the relationship between the donation per category of food and the explanatory variables discussed in section 3, by fitting a linear equation to observed data using the least square approach. The population regression line for l explanatory variables is defined using Eq (1) which explains how the mean of quantity of food donated varies with the explanatory variables.

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots \beta_l X_{li} + u \tag{1}$$

Where, Y_i is the donation/category/day, X_{ji} is explanatory variable, β_j is the estimated parameter which explains the expected change in Y_i for a one unit change in X_{ji} and u is the unobserved error term. In MLR, donation is assumed to have a linear relationship with the explanatory variables and the unknown model parameters are estimated from the data using linear predictor functions.

4.2 Structural Equation Model – Path Analysis

SEM- Path analysis model [21] is a statistical model used to test the causal relationships between the variables. SEM used in this study is a second generation multivariate modelling technique that enables the model to account for the relationships among multiple independent and dependent variables simultaneously [22]. It is an advanced version of MLR which enables the model to account for the relationship between the donations of different categories of food (dependent variables). The parameters are estimated by solving the regression equations simultaneously. The SEM approach can handle a large number of dependent and explanatory variables simultaneously. We develop two models, for weekday donation and weeend donation, and they are presented in Fig. 4 and 5.

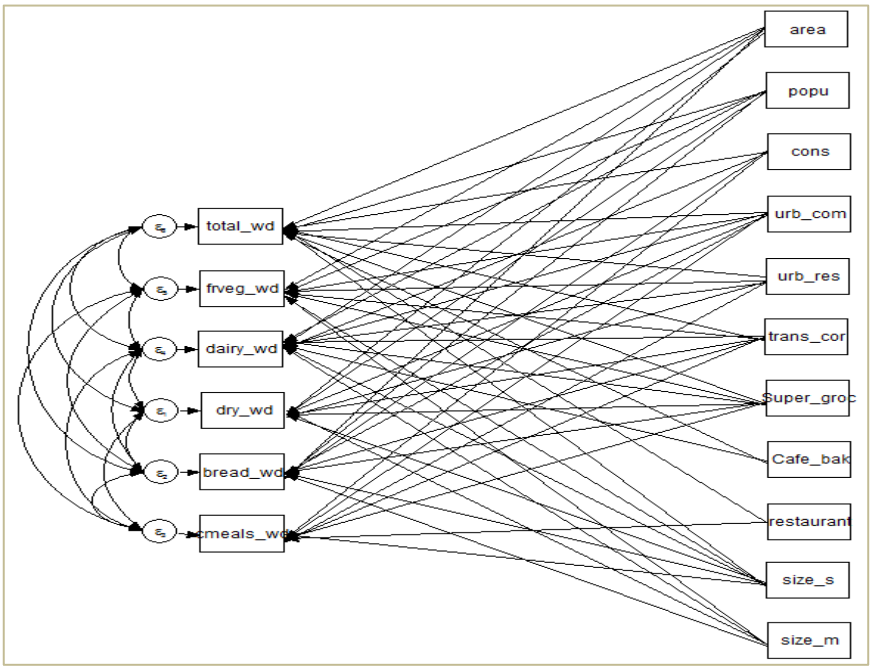


Fig 4. SEM Framework –Weekday

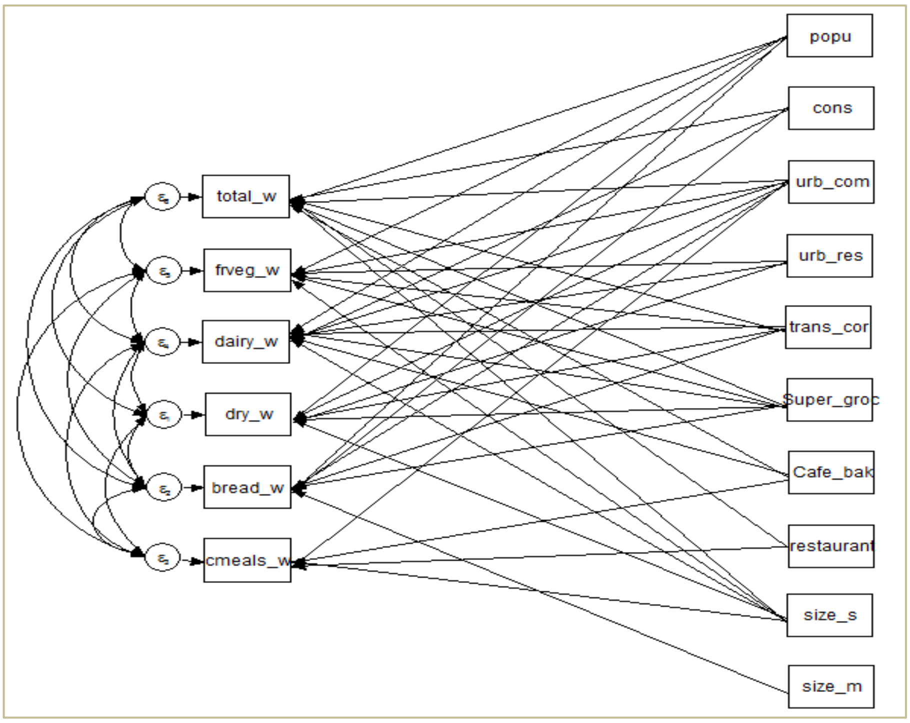


Fig 5. SEM Framework - Weekend

The models presented in the Fig. 4 and Fig. 5 are the final models which was obtained upon dropping insiginificant variables.

4.3 Feedforward backpropagation Neural Network

Artificial neural networks (ANN) [23] are statistical models inspired by the biological nervous system that can be used for pattern recognition, data classification, function approximation, fitness approximation, etc. ANN is a system of highly interconnected neurons passing information to each other and work together to solve a specific problem. Feedforward backpropagation Neural Network (FFNN) is a common type of neural network used for demand forecasting. Fig. 6 shows the *l*-*m*-*n* (*l* input neuron, *m* hidden neuron, *n* output neuron) structure of FFNN with all the three layers.

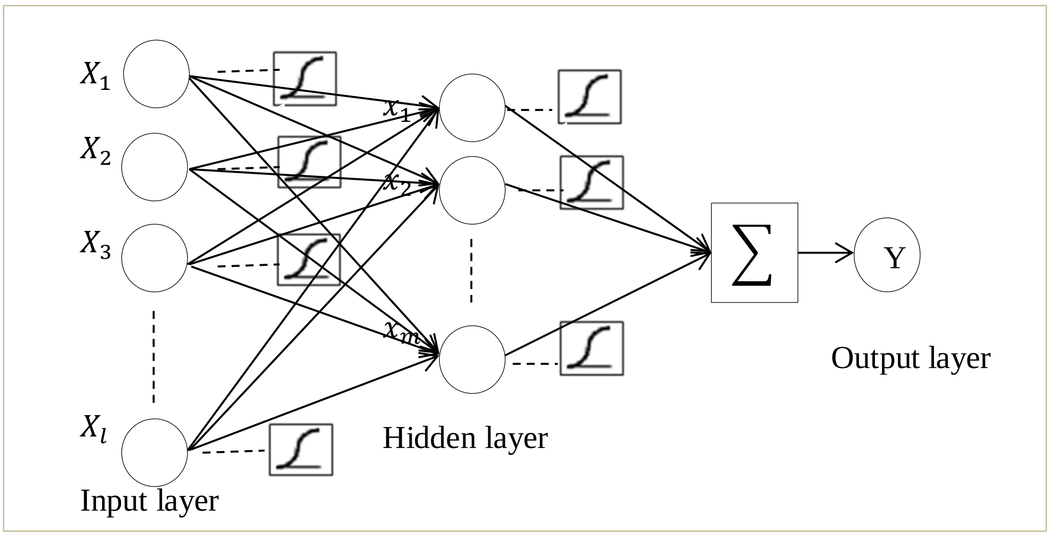


Fig 6. Architecture Of Feedforward Back Propagation Neural Network.

A feedforward neural network begins with an input layer. The input layer is connected to the hidden layer. FFNN with one hidden layer is generally used for function approximation. Hence, for the proposed study, we use an FFNN with single hidden layer. The number of neurons in the input layer and output layer is equal to the number of input and output variables. The neurons in the input layer are used for receiving information regarding the explanatory variables (Table 1) that defines the amount of donation and the neuron in the output layer is used to predict the donation/day. There is no exact method or formula available to determine the number of neurons in the hidden layer. In general, rule-of-thumb methods are adopted to determine the number of neurons. The number (i) should be greater than the sum of 2/3 of the total neurons in the input layer and the total neurons in the output layer [24], (ii) should be less than twice of the number of neurons in the input layer [25] and (iii) should be between the size of input layer and the output layer [26]. Using a lower number of neurons than required to represent the data may lead to underfitting, and a higher number of neurons may lead to overfitting. Hence, it is important to optimize the number of neurons in the hidden layer. To identify the number of neurons (m) in the hidden layer we use a search algorithm.

The search algorithm (Table 3) identifies the optimum number of hidden neurons by looking at the root mean absolute error (RMSE) obtained. Considering the three rules, the lower and upper limits of the number of hidden neurons are fixed as, $(2l/3)+1$ and $2l$. m corresponding the lowest RMSE value is taken as the number of hidden neurons.

Table 3 Pseudocode of Search Algorithm

Algorithm to find the number of neurons in hidden layer m

```
Let m- and m+ (integer values) be the lower and upper limit of m
Let Z_{m*}  be the performance evaluated (RMSE) for the model that produces the
lowest MSE when  the number of neurons in the hidden layer is m*
m- = (2l/3)+1, m+ = 2l , Z*_{m*} = ∞
i = 1, m* = m-
StopAlgorithm = false\
While (StopAlgorithm=false) Do
        Calculate Z_{m*}
        If  Z_{m*}  <   Z*_{m*}
                Z*_{m*} = Z_{m*} , m = m*
                i = i+1, m* = m- +1
                If  m* > m+
                        StopAlgorithm =true
                End if
        End if
End While
m = m*, RMSE =  Z*_{m*}
```

In a feedforward neural network, neurons are connected only forward and the information moves only in one direction, forward, from the input layer to hidden layers and then to the output layer. The input information is modified by interconnection weight, known as weight factor w_{ij}^{k-1}, which represents the interconnection of i^{th} node of the $k-1^{th}$ layer to j^{th}

node of the k^th^ layer. The Output of a j^{th} neuron in k^{th} layer of the network depends on the weight factor and neurons in the previous layer and can be written as Eq (2).

$$Y_j^k = f\left(\sum_{i=0}^{m_{k-1}} w_{ij}^{k-1} Z_i^{k-1}\right) \tag{2}$$

Where, Z_i^{k-1} represent i^{th} neuron in $k\text{-}1^{th}$ layer. Log- sigmoid transfer function “logsig” is used to transfer the information and is calculated by Eq (3).

$$f(x) = \frac{1}{e^{-\sum_{i=0}^{m_{k-1}} w_{ij}^{k-1} Z_i^{k-1}}} \tag{3}$$

Since the forecasting technique discussed in the study uses historical data, we use a supervised training algorithm backpropagation [27] to train the neurons. Backpropagation uses a gradient descent search method to adjust the interconnection weight. In the supervised training method, the network must be provided with both input (explanatory variables) and output (response variable). During training, the backpropagation algorithm compares the desired outputs and the predicted output obtained through the feedforward neural network and then calculates the mean square error (MSE). If MSE is greater than a prescribed limit value, it is back propagated backwards from output to input, continuing until the MSE value is within the limit.

4.4 Generalized Regression Neural Network

GRNN [28], one of the most popular neural networks, is a non-linear regression method mainly used for function approximation. Its performance is not sensitive towards the randomly assigned interconnection weights or the number of iterations in the training process, like in FFRN. The structure is very simple and is presented in Fig. 7.

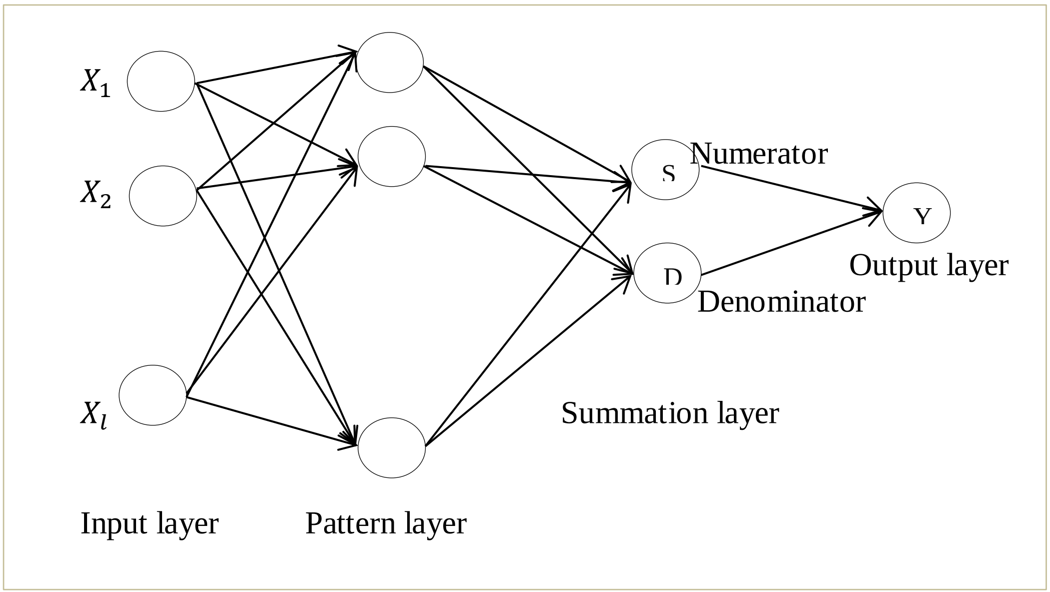


Fig 7. Architecture Of Generalized Regression Neural Network.

It has four layers: input, pattern, summation and output layer. As in FFNN, the number of neurons in the input and output layers of GRNN is fixed and equal to the number of input and output variables. The input layer is connected to the pattern layer, where each neuron represents a training pattern. In each pattern neuron *i*, a Gaussian PDF would be applied to the network input such that predicted output $\widehat{Y_i}(X)$ (Eq (4)),

$$\widehat{Y_i}(X) = \frac{\sum_i Y_i e^{-(\frac{(X-v_i)^2}{2\sigma^2})}}{\sum_i e^{-(\frac{(X-v_i)^2}{2\sigma^2})}} \quad (4)$$

Where, $\widehat{Y_i}$ is the output(donation/day) from i^{th} pattern neurons, X is the input explanatory variables (Table 1), v_i is training vector stored in the i^{th} pattern neuron, and σ is "spread" or "smooth parameter". The value of the spread is calculated experimentally for the problem under study. The pattern layer is connected to the summation layer that has two neurons, summation 'S' neuron and Summation 'D' neuron, which calculates the sum of the weighted and unweighted outputs of the pattern neuron and passes it to the neurons in the output layer. The neurons in output layer supply the predicted output by dividing the output of 'S' neuron with the output of 'D' neuron.

5 Results and Discussion

MLR model parameters are estimated by minimizing the sum of the squares of the errors (least square approach). Initially we considered all the explanatory variables presented in Table 1, for estimating the average daily donation and we then removed the insignificant variables while modelling. The results for parameter estimation of MLR are presented in Table 4. While the sign of the estimates is almost similar for weekend and weekday donation, it varies in different categories of food. The coefficient of the population has a positive sign for all the models which shows that the average daily donation increases as population increases. However the impact of population in average daily food donation is very small. In general, coefficients of supermarket/grocery, urban commercial/industrial and urban residential have positive signs for total donation models and fruit/vegetable donation models. This shows that the average daily donation is greater from supermarkets (consistent with the observation Fig 2.(b) and Fig. 2.(c)) and from urban residential and commercial areas.

The coefficient of the conservation area has a negative sign, showing that the average daily donation is lower from conservation areas. Size_s, which represent that the size of food provider is small, has a negative sign for all the donation models except average daily cooked meal donation (both weekday and weekend). The coefficient of restaurants for cooked meals has a positive sign for both weekend and weekday donation. This shows that the average daily donation of cooked meals is more from restaurants (consistent with the observation Fig. 2.(c)). Average donations of bread, dairy and drystock from restaurants were negligible when compared to other food providers, and hence the coefficient of the restaurant has a negative sign for weekend and weekday models.

Table 3 Modelling Results for MLR

Variable	β	p value	Variable	β	p value	Variable	β	p value
Weekend donations								
total_w			**bread_w**			**fruveg_w**		
Popu	0.0044	0.000	Popu	0.0005	0.000	Popu	0.0005	0.006
Cons	-28.553	0.000	Cons	-8.630	0.000	trans_corr	10.702	0.000
trans_corr	-16.693	0.003	trans_corr	-6.145	0.001	urb_com	17.761	0.020
urb_com	8.007	0.108	urb_com	6.448	0.001	urb_res	24.593	0.000
super_groc	41.795	0.000	urb_res	5.342	0.003	super_groc	6.427	0.000
cafe_bak	37.751	0.000	super_groc	7.382	0.000	size_s	-6.302	0.038
Restaurant	45.230	0.000	cafe_bak	6.980	0.000			
size_s	-12.986	0.001	restaurant	-7.758	0.000	**dairy_w**		
			size_m	1.209	0.021	Popu	0.036	0.000
dry_w						Cons	-33.056	0.000
Cons	-3.417	0.013	**cmeal_w**			trans_corr	-33.373	0.000
trans_corr	-3.479	0.019	urb_com	0.679	0.063	urb_com	33.510	0.000
urb_com	3.707	0.012	cafe_bak	1.950	0	urb_res	32.019	0.000
urb_res	2.538	0.067	restaurant	2.206	0	super_groc	-14.055	0.000
super_groc	10.593	0.000	size_s	0.431	0.16	cafe_bak	-10.829	0.000
cafe_bak	9.194	0.000				restaurant	12.420	0.000
Restaurant	-9.430	0.000				size_s	-1.380	0.060
size_s	-1.232	0.004						
Weekday donations								
total_wd			**bread_wd**			**fruveg_wd**		
Popu	0.027	0.000	Popu	0.003	0.000	Popu	0.009	0.029
Cons	-12.190	0.000	Cons	-0.933	0.000	trans_corr	8.836	0.000
urb_com	7.464	0.005	trans_corr	0.959	0.002	urb_com	6.155	0.001
urb_res	8.264	0.001	urb_com	0.693	0.021	urb_res	18.098	0.000
super_groc	3.142	0.020	super_groc	-1.675	0.000	super_groc	14.176	0.000
cafe_bak	-4.081	0.033	cafe_bak	-0.439	0.060	size_s	-5.245	0.001
Restaurant	-3.142	0.034	restaurant	-0.939	0.000	size_m	3.081	0.074
size_s	-11.274	0.000	size_s	-2.214	0.000			
size_m	3.406	0.011	size_m	1.550	0.000	**dairy_wd**		
						Popu	0.00011	0.000
dry_wd			**cmeal_wd**			Cons	-14.479	0.000
Popu	0.00005	0.000	Popu	0.001	0.024	trans_corr	-17.046	0.000
trans_corr	-0.992	0.001	Cons	-3.470	0.000	urb_com	16.317	0.000
urb_com	-0.643	0.022	trans_corr	-3.663	0.000	urb_res	-14.676	0.000
urb_res	0.931	0.000	urb_com	2.876	0.000	super_groc	4.085	0.000
super_groc	-0.491	0.017	urb_res	-3.393	0.000	cafe_bak	0.684	0.074
size_s	-0.949	0.000	super_groc	-0.966	0.000	size_s	-4.635	0.000
size_m	0.955	0.000	restaurant	1.696	0.000	size_m	-3.770	0.000
			size_s	0.458	0.003			

Table 5 compares the goodness of fit of MLR and neural network models, FFNN and GRNN. The most difficult to forecast were bread (weekday) and cooked meals. These two are the categories that had the least frequent donation and the least total donation amount. We identified the *l*-*m*-*n* structure of FFNN using a search algorithm as discussed in section 4.3. The lower bound was taken as 8 and upper bound as 22 since the number of input neurons was 11. The optimal *l-m-n* structure and goodness of fit, R^2value obtained from the models are presented in Table 5.

Table 5 Comparing The Goodness-of-fit of MLR, FFNN and GRNN

Donation Variable	MLR		FFNN			GRNN	
	R^2	RMSE	(l-m-n)	R^2	RMSE	R^2	RMSE
total_wd	0.229	27.89	(11-17-1)	0.744	15.60	0.644	18.38
fruveg_wd	0.268	23.12	(11-20-1)	0.781	12.97	0.66	16.23
dairy_wd	0.198	6.62	(11-20-1)	0.63	4.49	0.63	4.49
bread_wd	0.115	3.36	(11-13-1)	0.44	2.86	0.43	2.88
dry_wd	0.0736	3.43	(11-9-1)	0.28	3.03	0.281	3.03
cmeal_wd	0.1633	2.64	(11-16-1)	0.79	2.87	0.601	3.95
total_w	0.342	33.85	(11-17-1)	0.57	28.28	0.452	31.98
fruveg_w	0.271	29.08	(11-16-1)	0.503	22.38	0.39	24.80
dairy_w	0.518	6.23	(11-9-1)	0.663	3.92	0.663	3.92
bread_w	0.29	4.86	(11-9-1)	0.493	3.56	0.43	3.99
dry_w	0.37	3.4	(11-9-1)	0.654	2.13	0.654	2.13
cmeal_w	0.139	2.75	(11-20-1)	0.424	1.89	0.22	2.19

Comparing the goodness of fit and RMSE it can be seen that ANN outperforms MLR in forecasting the food donation which is similar to the findings of Brock III and Davis (2015) and in many other fields. This shows that the average daily donation has a nonlinear relationship with observed variables. Also, one cannot expect all the supermarkets or all the restaurants to behave in a similar way. FFNN and GRNN provide a better fit as a result of pattern recognition and generalization made by the network. They are also capable of accounting for the interaction between the observable variables. The RMSE values are comparatively lower and the goodness of fit is high for FFNN when compared to GRNN. This is due to the larger number of categorical variables in the input. GRNN works well with continuous variables rather than categorical variables.

As discussed in section 4.2, in structural equation models we assume a correlation between the response variables. The covariances between the response variables are presented in Table 6.

Table 6 Modelling Results For SEM

Weekend			
Variable		Covariance	p value
total_w	fruveg_w	915.6336	0.000
	dairy_w	61.35311	0.000
	bread_w	71.67085	0.000
	dry_w	28.58984	0.000
fruveg_w	dairy_w	10.03858	0.093
	bread_w	41.10933	0.000
	cmeal_w	-6.89138	0.000
dairy_w	bread_w	2.670572	0.013
	dry_w	9.528874	0.000
bread_w	dry_w	4.229868	0.000
RMSEA	0.026 (lower bound =0.008, upper bound =0.041, pclose =0.998)		
CFI	0.997		
Overall R^2	0.77		

Weekday			
total_wd	dairy_wd	80.28104	0.000
	bread_wd	33.26586	0.000
	dry_wd	25.93783	0.000
	fruveg_wd	561.7875	0.000
dairy_wd	bread_wd	8.656263	0.000
	dry_wd	7.6785	0.000
	fruveg_wd	16.41125	0.000
	cmeal_wd	-2.87507	0.000
bread_wd	dry_wd	1.938019	0.000
	fruveg_wd	11.30472	0.000
	cmeal_wd	-1.32863	0.000
dry_wd	cmeal_wd	2.067621	0.000
fruveg_wd	cmeal_wd	-7.27885	0.000
RMSEA	0.019 (lower bound =0.010, upper bound =0.027, pclose =1.00)		
CFI	0.998		
Overall R^2	0.601		

In the weekend donation model we could not identify any significant correlation between cooked meals and other categories of food. Also, in the weekday model, cooked meals are negatively correlated with dairy products and bread. This is due to the fact that cooked meals are generally supplied by restaurants and dairy and bread by supermarkets, cafe/bakery and other sources.

Table 7 Comparing the overall goodness of fit of MLR, SEM, FFRN and GRNN

	Weekday		Weekend	
	R^2	RMSE	R^2	RMSE
MLR	0.261	19.19	0.439	21.97
SEM	0.601	15.31	0.770	18.57
FFNN	0.781	9.31	0.612	15.09
GRNN	0.656	10.44	0.583	15.72

The R^2value for weekend model is 0.77 and weekday model is 0.601 which shows that the SEM models provide a better approximation when compared to MLR by correlating the response variables. From Table 7, the overall goodness of fit- R^2 and the RMSE (weekend -18.57 and Weekday-15.31) obtained for donation estimates of SEM models are comparable with the R^2 value (weekend - 0.612 and weekday - 0.781) and RMSE (weekend - 15.09 and weekday - 9.31) obtained for FFNN models. Hence SEM models can be effectively used for policy implications because the relationship between variables is explained in the mathematical formulation in a trackable way. Furthermore, the SEM structure facilitates the assessment of hypothetical situations, while NN methods require a simulation process for policy appraisal. The results also suggest that the SEM and NN provide a better demand estimation of different categories of food at different food providers. Since no forecasting methods are perfect, this estimated demand may be considered as partial information, and the operations manager may use this partial information to design the food relief routes with objective functions aimed at obtaining redistribution of recovered food in a sustainable and equitable way.

6 Conclusion and Future Direction

This paper addresses the uncertainty issues in food rescue operations designed to end food insecurity and hunger, and helps food rescue organizations estimate the approximate amount of different types of food available at food providers. Understanding the pattern and availability of donated food helps food rescue organizations effectively plan and manage the storage and equitable distribution of food in a sustainable way. This paper evaluates the impact of three forecasting techniques: multiple linear regression, structural equation modelling and artificial neural networks to explore patterns in the food donation process. Models are used to forecast the average daily donation amount/category/food provider/day. The results suggest that structural equation models and neural networks provide improved demand estimation when compared to conventional multiple linear regression. Due to the simplicity, usefulness and relatively high R^2 value of SEM, it can be effectively utilized for policy implications. Many studies have been conducted in different fields, comparing the efficiency of ANN over conventional MLR and our results agree with their findings that ANN outperforms simpler models when dealing with highly complex data. Food donation is a random process which can be explained as a function of many factors. The study identified a few variables, namely the type of food provider, the size of food provider and region where the food provider is located (land-use, population and area), that have a significant nonlinear impact on food donation. Further improvements in the model include considering the effect of social and demographic characteristics of people living in the region under study. We are also interested in checking the possibility of scaling the model to a national level.

Additionally, we have also conducted a *Needs Assessment Survey* to understand the demand (request of different types of food) of welfare agencies in Sydney. The data collected includes agency name, location, type of food assistance provided, type of customers assisted, and preference and expectation (quantity) of different types of food. Another extension of the paper would be to analyse the data collected through the survey, understand the underrun and overrun pattern and study the impact of supply forecasting model on operational decisions and equitable allocation of rescued food using the routing and allocation models proposed by Nair et al. 2016 [13, 14]

This study directly addresses the research across non-profit sectors to enable societal transformation to enhance sustainability and wellbeing. With the increased shortage of food and poverty, capturing safe and nutritious food that would have been wasted and directing them to vulnerable people through innovative transportation methods addresses the research priorities. Specifically, estimating different types of in-kind food supplied at food providers enables food rescue organizations to design collection and delivery networks by optimizing the total transportation cost and equity in the distribution of food. Furthermore, being able to use these models developed in this research will provide non-profit organizations and government with the tools that would help to enhance sustainable food management systems.

These models are also envisioned to help government agencies to closely track food waste generation and evaluate the environmental impact associated with it.

Acknowledgements

We would like to express our sincere appreciation to Gopi Krishnan ,OzHarvest for their contribution towards this research. This research has been supported by grant ARC LP150101266 from Australian Research Council.

Reference

[1] World Population Statistics 2013. http://www.worldpopulationstatistics.com/australia-population-2014.

[2] Foodbank. End Hunger Report 2014. http://www.foodbank.org.au/wp-content/uploads/2014/10/Foodbank-Hunger-Report2014.pdf.

[3] Specht DF. A General Regression Neural Network. IEEE Trans Neural Netw 1991;2:568–76. doi:10.1109/72.97934.

[4] Drackley A. How Generous Are We? Forecasting and Demographic Correlates of Blood Donation. thesis. 2010.

[5] Drackley A, Newbold KB, Paez A, Heddle N. Forecasting Ontario's blood supply and demand. Transfusion (Paris) 2012;52:366–74. doi:10.1111/j.1537-2995.2011.03280.x.

[6] A Neural Network-based Approach for Predicting Organ Donation Potential - ProQuest n.d. http://search.proquest.com/openview/ef3f0fed37aa70dc56a285ca24d78373/1?pq-origsite=gscholar (accessed September 18, 2015).

[7] Firat M, Yurdusev MA, Turan ME. Evaluation of Artificial Neural Network Techniques for Municipal Water Consumption Modeling. Water Resour Manag 2008;23:617–32. doi:10.1007/s11269-008-9291-3.

[8] JAIN A, ORMSBEE LE. Short-term Water demand forecast modeling techniques—CONVENTIONAL METHODS VERSUS AI. J Am Water Works Assoc 2002;94:64–72.

[9] Nasr GE, Badr EA, Younes MR. Neural networks in forecasting electrical energy consumption: univariate and multivariate approaches. Int J Energy Res 2002;26:67–78. doi:10.1002/er.766.

[10] Davis LB, Sengul I, Ivy JS, Brock III LG, Miles L. Scheduling Food Bank Collections and Deliveries to Ensure Food Safety and Improve Access. Socioecon Plann Sci 2014;48:175–88. doi:10.1016/j.seps.2014.04.001.

[11] Gunes C, Hoeve W-J van, Tayur S. Vehicle Routing for Food Rescue Programs: A Comparison of Different Approaches. In: Lodi A, Milano M, Toth P, editors. Integr. AI Tech. Constraint Program. Comb. Optim. Probl., Springer Berlin Heidelberg; 2010, p. 176–80.

[12] Solak S, Scherrer C, Ghoniem A. The Stop-and-Drop Problem in Nonprofit Food Distribution Networks. Ann Oper Res 2012;221:407–26. doi:10.1007/s10479-012-1068-7.

[13] Nair D. J, Rey D, Dixit V, Valenta T. Models for Food Rescue and Delivery: Routing and Resource Allocation Problem, Transportation Research Board Annual Meeting 2016 Paper #16-2584.

[14] Nair D. J, Grzybowska H, Rey D, Dixit V. A Heuristic Algorithm for Periodic Unpaired Pickup and Delivery Vehicle Routing Problem Transportation Research Record (accepted)
[15] Balcik B, Iravani S, Smilowitz K. Multi-Vehicle Sequential Resource Allocation for a Nonprofit Distribution System. IIE Trans 2014;46:1279–97. doi:10.1080/0740817X.2013.876240.
[16] Lien RW, Iravani SMR, Smilowitz KR. Sequential Resource Allocation for Nonprofit Operations. Oper Res 2014;62:301–17. doi:10.1287/opre.2013.1244.
[17] Phillips C, Hoenigman R, Higbee B, Reed T. Understanding the Sustainability of Retail Food Recovery. PLoS ONE 2013;8:e75530. doi:10.1371/journal.pone.0075530.
[18] Davis LB, Jiang SX, Morgan SD, Harris C. Forecasting Donated Food Goods at a Local Food Bank, Charleston, SC: 2013.
[19] Jiang S, Davis LB, De Mleo HT, Terry J. Using Data Mining to Analyse Donation Data for a Local Food Bank, Las Vegas, NV: 2013.
[20] Brock III LG, Davis LB. Estimating Available Supermarket Commodities for Food Bank Collection in the Absence of Information. Expert Syst Appl 2015;42:3450–61. doi:10.1016/j.eswa.2014.11.068.
[21] Wright S, Correlation and causation. Journal of Agricultural research 1921; 20-557-585
[22] Anderson JC, Gerbing DW. Structural equation modeling in practice: A review and recommended two-step approach. Psychol Bull 1988;103:411–23. doi:10.1037/0033-2909.103.3.411.
[23] McCulloch WS, Pitts W. A Logical Calculus of the Ideas Immanent in Nervous Activity. Bull Math Biophys 1943;5:115–33. doi:10.1007/BF02478259.
[24] Boger Z, Guterman H. Knowledge extraction from artificial neural network models. 1997 IEEE Int. Conf. Syst. Man Cybern. 1997 Comput. Cybern. Simul., vol. 4, 1997, p. 3030–5 vol.4. doi:10.1109/ICSMC.1997.633051.
[25] Berry MJ, Linoff G. Data Mining Techniques: For Marketing, Sales, and Customer Support. New York, NY, USA: John Wiley & Sons, Inc.; 1997.
[26] Blum A. Neural Networks in C++: An Object-Oriented Framework for Building Connectionist Systems. 1st ed. New York, NY, USA: John Wiley & Sons, Inc.; 1992.
[27] Rumelhart DE, Hinton GE, Williams RJ. Learning Representations by Back-Propagating Errors. Cogn. Model., vol. 323, MIT Press; 1986, p. 533–6.
[28] Chevaleyre Y, Dunne PE, Endriss U, Lang J, Maudet N, Rodríguez-Aguilar JA. Multiagent resource allocation. Knowl Eng Rev 2005;20:143–9. doi:10.1017/S0269888905000470.