22nd International Congress on Modelling and Simulation, Hobart, Tasmania, Australia, 3 to 8 December 2017
mssanz.org.au/modsim2017

Whey reverse logistics network design: a stochastic hierarchical facility location model

Rasul Esmaeilbeigi [a], Ali Eshragh [a], Rodolfo Garcia-Flores [b], Mojtaba Heydar [c]

[a]*School of Mathematical and Physical Sciences, The University of Newcastle, Australia*
[b]*Optimization Applications and Platforms, CSIRO Data-61, Docklands, Australia*
[c]*School of Electrical Engineering and Computing, The University of Newcastle, Australia*
Email: rasul.esmaeilbeigi@uon.edu.au

Abstract: Whey reverse logistics has received a considerable attention due to the recent environmental legis-lation and competitive advantages. In this paper, we study a whey reverse logistics network design problem with stochastic demand, where demand is the amount of raw whey produced by a cheese maker that should be converted into a commercial product. We formulate this problem as a hierarchical facility location problem with two levels of facilities, namely collection centers and plants. Collection centers receive raw whey from cheese makers and convert it to concentrated whey. Then the concentrated whey is sent to plants for further processing. A plant can be established in a potential location if a collection center has already been established in that point. The objective is to determine which facilities to open and to allocate each demand point to the open facilities such that the total expected cost is minimized. The total cost is comprised of transportation cost and fixed cost of opening facilities. The problem with deterministic demand is formulated as a mixed integer linear program. We use this formulation to model the stochastic version of the problem as a two-stage stochastic program, where decisions are made at two sequential stages. At the first stage, we decide which facilities to open and at the second stage, after observing a realization of demands, we decide how to allocate each cheese maker to the open facilities. We use the sample average approximation method to estimate the expected value function. The resulting formulation can be solved using a standard linear solver. Results of a computational study on a set of randomly generated instances will be reported.

Keywords: *Whey reverse logistics, dairy industry, hierarchical facility location, stochastic programming*

1 INTRODUCTION

Due to stricter environmental legislation and economic advantages, recycling of whey, a by-product of cheese production process, has drawn a considerable attention (Nath et al., 2016). Historically, whey was considered a burdensome, environmentally damaging by-product of cheese making. It was considered a nuisance and troublesome to dispose of since the amount of whey nearly equalled 90% of the amount of milk (Bansal and Bhandari, 2016). The most economical options to dispose of whey were spraying onto fields and paddocks, discharge into creeks, rivers and the ocean, treatment through municipal sewage works and use as animal feed (Smithers, 2015).

The developments in science and technology have transformed whey from a troublesome waste product to valuable dairy ingredients. About 50% of the total produced whey is now transformed into value-added products (Yadav et al., 2015). As demand for milk-derived products is increasing, the production of whey is increased accordingly. This can pose a serious waste management problem in countries where this by-product is not processed, but dumped into the environment. Furthermore, it is important to recover whey in an economic manner to remain competitive. The future food insecurity, on the other hand, demands the utilization of nutrient-rich residues from food industries as value-added products (Patel, 2015). These issues emphasize the necessity for an efficient whey Reverse Logistics.

Reverse logistics has been defined as "the process of planning, implementing and controlling backward flows of raw materials, in process inventory, packaging and finished goods, from a manufacturing, distribution or use point, to a point of recovery or point of proper disposal" (de Brito and Dekker, 2004). Firms may adopt reverse logistics by choice because of inherent economic or competitive advantages or by force because of legislation or environmental reasons (Agrawal et al., 2015). Consequently, designing a whey recovery network is an important reverse logistics problem.

Reverse logistics activities are mainly supported by two types of facilities, namely collection centers and recovery/remanufacturing facilities (Melo et al., 2009). Products are first collected at the collection centers and then transported to the recovery facilities for further processing. These facilities are typically hierarchical in the sense that they provide different levels of services. Thus, designing such reverse logistics networks is relevant to the Hierarchical Facility Location Problems (HFLPs) in which facilities provide different levels of services. The reader is referred to Narula (1984), Şahin and Süral (2007) and Daskin (2013) for more information on HFLPs.

Whey reverse logistics network design is first addressed by García-Flores et al. (2015) in which the authors study an actual cluster of cheese makers in Minas–Gerais, Brazil. They formulate the problem as a HFLP with collection centers and plants as two levels of facilities. Each cheese maker (demand point) has a specific amount of raw whey (demand) that must be transported to the collection centers. A collection center provides CONCENTRATE service which concentrates the raw whey. This service converts each liter of the raw whey to 0.339 liters of the concentrated whey. The concentrated whey must go to the plants for further processing. Plants provide PROCESS service which converts the concentrated whey to a commercial product. The authors consider 40% demineralized whey powder (40DWP) as the commercial product. In this case, each liter of the concentrated whey is converted to 0.186 liters of 40DWP.

García-Flores et al. (2015) define each service as a sequence of stages. The set of operations in each stage is carried out by a single machine, but there may be more than one available instance of that machine. Different instances of a machine do the same task, but in different rates and operating costs. A valid combination of machines that perform stage operations (tasks) of a service is called a configuration of that service. The authors assume that different facilities may have different configurations. Their proposed model consists of selecting a set of facilities to open, allocating each demand point to the open facilities, and assigning a configuration of machines to each open facility such that the total profit is maximized. The total profit is the difference between the total income from selling 40DWP and the total cost.

In the present paper, we formulate the whey reverse logistics network design problem as a mixed integer linear program. We assume that there is only one configuration available for each level of facilities. The objective is to locate a number of collection centers and plants among the demand points and allocate each demand point to the open facilities such that the total cost is minimized. The total cost consists of the transportation costs and the fixed costs of opening facilities. To tackle the issue of uncertainty, we assume that demands are random variables and use the two-stage stochastic programming approach to address the problem. In this approach, the decisions are made at two different stages. At the first stage, we decide which facilities to open. At the second stage, when a realization of demands is observed, we decide how to allocate each demand point to the open facilities. We use the sample average approximation method to approximate the expected value function.

The present paper is organized as follows: In Section 2 we present a deterministic formulation for the problem. In Section 3 we formulate the problem with stochastic demands as a two-stage stochastic program. In Section 4 we report results of a computational study. The paper is concluded with Section 5.

2 Deterministic formulation

Our proposed model of whey reverse logistics network design problem consists of collection centers and plants, providing CONCENTRATE and PROCESS services, respectively. Every cheese maker is a potential location for collection centers and plants. A plant cannot be established in a particular location unless a collection center has already been established in that location. This property is called a *nested hierarchy*. The problem can be considered as a HFLP with *single-flow* pattern as the raw whey must go from a particular location to the collection centers and then the resulting concentrated whey is transported to the plants. In other words, raw whey cannot go to a plant directly. The problem has a *non-coherent* structure in the sense that the concentrated whey from a particular collection center can go to any plants. Finally, the problem has the *referral* discipline because for each liter of the raw whey only 0.339 liters of the concentrated whey is transported to the collection centers. The reader is referred to Şahin and Süral (2007) for the relevant terminology of HFLPs. Figure 1 shows an example of a feasible solution for the problem. The parameters of the problem are given in Table 1.

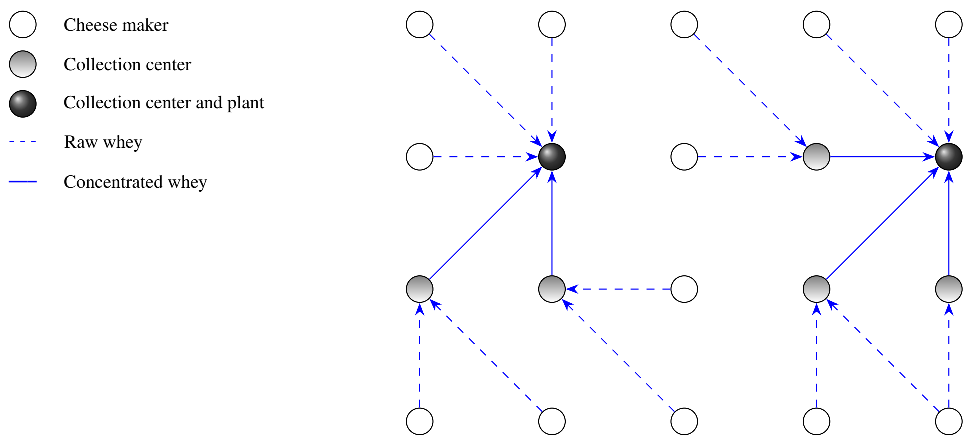

Figure 1. An example of a feasible solution for the problem

Table 1. The parameters of the problem

Notation	Definition
n	The number of nodes (cheese makers)
$\mathcal{N}$	The set of nodes, that is, $\mathcal{N} = \{1, 2, \ldots, n\}$
t_{ij}	The transportation cost from node $i \in \mathcal{N}$ to node $j \in \mathcal{N}$ ($ per liter per day)
d_i	The amount of raw whey at node $i \in \mathcal{N}$ (liter per day)
f_i^1	The fixed cost of opening a collection center at location $i \in \mathcal{N}$ ($ per day)
f_i^2	The fixed cost of opening a plant at location $i \in \mathcal{N}$ ($ per day)
c_1	The capacity of a collection center (liter per day)
c_2	The capacity of a plant (liter per day)
θ	The conversion rate of raw whey to concentrated whey (i.e., $\theta = 0.339$)

We formulate this problem as a mixed integer linear program. Let y_i^1 be a binary variable which takes value one if a collection center is established in node $i \in \mathcal{N}$, and zero otherwise. The decision variable y_i^2 is defined similarly for a plant. We define the continuous variable x_{ij}^1 to represent the amount of raw whey transported from node $i \in \mathcal{N}$ to node $j \in \mathcal{N}$. The decision variable x_{ij}^2 is defined similarly for the concentrated whey. A deterministic formulation of the problem can be represented as follows:

$$
\begin{aligned}
\min \quad & \sum_{i \in \mathcal{N}} (f_i^1 y_i^1 + f_i^2 y_i^2) + \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} t_{ij} \, (x_{ij}^1 + x_{ij}^2) && && (1) \\
\text{subject to} \quad & \sum_{i \in \mathcal{N}} x_{ji}^1 \geq d_j && \forall j \in \mathcal{N} && (2) \\
& \sum_{i \in \mathcal{N}} x_{ji}^2 = \theta \sum_{i \in \mathcal{N}} x_{ij}^1 && \forall j \in \mathcal{N} && (3) \\
& \sum_{i \in \mathcal{N}} x_{ij}^1 \leq c_1 y_j^1 && \forall j \in \mathcal{N} && (4) \\
& \sum_{i \in \mathcal{N}} x_{ij}^2 \leq c_2 y_j^2 && \forall j \in \mathcal{N} && (5) \\
& y_j^2 \leq y_j^1 && \forall j \in \mathcal{N} && (6) \\
& y_j^1, y_j^2 \in \{0, 1\} && \forall j \in \mathcal{N} && (7) \\
& x_{ij}^1, x_{ij}^2 \geq 0 && \forall i, j \in \mathcal{N} && (8)
\end{aligned}
$$

The objective function (1) minimizes the total transportation and fixed costs. Constraint (2) implies that the raw whey produced by the cheese makers must be transported to collection centers. According to the balance Constraint (3), the concentrated whey produced in location $j \in \mathcal{N}$ is proportional to the input raw whey to node $j \in \mathcal{N}$ by the conversion rate of θ. Constraint (4) ensures that raw whey is transported to node $j \in \mathcal{N}$ if a collection center is established in that node. This constraint also guarantees that the incoming raw whey does not exceed the capacity of the collection center. Constraint (5) works similarly for concentrated whey and plants. The hierarchical Constraint (6) states that a plant can be established in node $j \in \mathcal{N}$ if a collection center is also established in that node. Constraints (7) and (8) specify the domain of decision variables. The formulation can be solved by any standard linear solvers such as CPLEX or GUROBI.

3 STOCHASTIC PROGRAMMING

We are interested in the stochastic version of the problem in which demands are random variables. We use the two-stage stochastic programming approach to address the stochastic problem. In this approach, the decisions are made at two sequential stages. At the first stage, we decide where to install the collection centers and plants. At the second stage, when a realization of demands is observed, we decide how to distribute demands among the facilities. Because at the first stage we do not know the future demands, we minimize the expected total costs, that is, the establishment fixed costs of facilities plus the expected transportation costs. Note that the fixed costs of establishing facilities do not depend on the realization of demands as we fix the location of facilities at the first stage. However, the transportation costs are incurred at the second stage after we know the location of facilities and the values of demands. Therefore, they depend on the realization of demands, hence minimizing their expected values at the first stage.

A two-stage stochastic program is typically solved by constructing scenarios. Let D_i be the random variable representing the demand of node $i \in \mathcal{N}$. Each realization of the random vector $\Delta = (D_1, \ldots, D_n)$ is called a scenario. We assume that random variables $D_1, \ldots, D_n$ are stochastically independent. If we construct scenarios by assigning only three possible values (e.g., low, medium and high) to each component of Δ, then the total number of scenarios is 3^n. Due to such exponential growth in the number of scenarios, it is usually impractical to consider all possible scenarios. We may, however, generate a random sample $\Delta^1, \ldots, \Delta^m$ of m realizations of the random vector Δ to construct m scenarios. This idea is utilized in the so-called Sample Average Approximation Method (henceforth referred to as SAAM), where the weight of each scenario is set to $\frac{1}{m}$. This designation of weights is consistent with the law of large numbers, implying that the sample mean is a good estimation of the population mean (Shapiro, 2008).

We consider SAAM and formulate the resulting stochastic program as a mixed integer linear program. Let $\mathcal{S}_m = \{1, \ldots, m\}$ be the set of scenarios generated through SAAM. The demand of node $i \in \mathcal{N}$ in scenario $s \in \mathcal{S}_m$ (the i^{th} component of Δ^s) is denoted by d_{is}. We assume that the demand d_{is} must be satisfied for each node $i \in \mathcal{N}$ and for each scenario $s \in \mathcal{S}_m$. Depending on the distributions of random variables $D_1, \ldots, D_n$, we may have $\sum_{i\in\mathcal{N}} d_{is} > \min(nc_1, \frac{nc_2}{\theta})$ for some scenario $s \in \mathcal{S}_m$. In this case, we are not able to process the total raw whey through our (already installed) facilities even if we establish a facility in every node at the first stage. Therefore, a recourse action is required to satisfy the additional demands. We assume that the additional demand of node $i \in \mathcal{N}$ is outsourced and costs us w_i dollars per liter.

In order to formulate the problem, we define the decision variable z_{is} to denote the outsourcing demand of node $i \in \mathcal{N}$ in scenario $s \in \mathcal{S}_m$. We also define the continuous variable x^1_{ijs} to represent the amount of raw whey transported from node $i \in \mathcal{N}$ to node $j \in \mathcal{N}$ if scenario $s \in \mathcal{S}_m$ occurs. The continuous variable x^2_{ijs} is defined similarly for the concentrated whey. Using the new decision variables and the binary variables y^1_j and y^2_j, the stochastic program can be formulated as follows:

$$\min \sum_{i\in\mathcal{N}} (f^1_i y^1_i + f^2_i y^2_i) + \frac{1}{m} \sum_{i\in\mathcal{N}} \sum_{j\in\mathcal{N}} \sum_{s\in\mathcal{S}_m} t_{ij}(x^1_{ijs} + x^2_{ijs}) + \frac{1}{m} \sum_{i\in\mathcal{N}} \sum_{s\in\mathcal{S}_m} w_i z_{is} \tag{9}$$

subject to (6), (7), and

$$\sum_{i\in\mathcal{N}} x^1_{jis} + z_{js} \geq d_{js} \qquad \forall j \in \mathcal{N},\ s \in \mathcal{S}_m \tag{10}$$

$$\sum_{i\in\mathcal{N}} x^2_{jis} = \theta \sum_{i\in\mathcal{N}} x^1_{ijs} \qquad \forall j \in \mathcal{N},\ s \in \mathcal{S}_m \tag{11}$$

$$\sum_{i\in\mathcal{N}} x^1_{ijs} \leq c_1 y^1_j \qquad \forall j \in \mathcal{N},\ s \in \mathcal{S}_m \tag{12}$$

$$\sum_{i\in\mathcal{N}} x^2_{ijs} \leq c_2 y^2_j \qquad \forall j \in \mathcal{N},\ s \in \mathcal{S}_m \tag{13}$$

$$z_{js} \geq 0 \qquad \forall j \in \mathcal{N},\ s \in \mathcal{S}_m \tag{14}$$

$$x^1_{ijs}, x^2_{ijs} \geq 0 \qquad \forall i, j \in \mathcal{N},\ s \in \mathcal{S}_m \tag{15}$$

The objective function (9) minimizes the total expected cost, that is, the establishment fixed costs of facilities plus the expected transportation and outsourcing costs. Constraint (10) implies that the demand of each cheese maker must be satisfied in each scenario. For any given scenario $s \in \mathcal{S}_m$, the Constraints (11)–(13) play the same role as the Constraints (3)–(5), respectively. Constraints (14) and (15) state that the new decision variables are non-negative.

It is interesting to see what it would cost if no plants or collection centers are built. In that case, $z_{is} = d_{is}$ for each $i \in \mathcal{N}$ and $s \in \mathcal{S}_m$ implying that the objective value is

$$\frac{1}{m} \sum_{i\in\mathcal{N}} \sum_{s\in S_m} w_i\, d_{is} = \sum_{i\in\mathcal{N}} w_i \left(\frac{1}{m} \sum_{s\in S_m} d_{is} \right) = \sum_{i\in\mathcal{N}} w_i\, \mu_i,$$

where $\mu_i := \frac{1}{m} \sum_{s\in S_m} d_{is}$ represents the average demand of node $i \in \mathcal{N}$.

4 Computational results

In this section, we report results of a computational experiment on the proposed stochastic program. We use a set of randomly generated instances with $n = 5$ demand points to study the number of required scenarios. We use IBM ILOG CPLEX Optimizer 12.7 with C++ programming language to solve the stochastic programming formulation. All the experiments were run under the Linux operating system with a single-core 2.2 gigahertz Intel processor and 8.0 gigabyte RAM. To generate instances of the problem, we use the discrete uniform distribution as follows: We assume that demand points are located on the xy-plane, where the coordinate of node $i \in \mathcal{N}$ is (X^i, Y^i). We consider the parameter t_{ij} as the distance between the two demand points $i, j \in \mathcal{N}$. We generate X^is and Y^is randomly and independently on the interval $[0, 50]$. Values of the parameters $d_{js}, j \in \mathcal{N}, s \in \mathcal{S}_m$ are generated randomly and independently on the interval $[1, 5]$. Capacity of the facilities, c_1 and c_2, are generated randomly and independently on the intervals $[3, 15]$ and $[2, 6]$, respectively. Fixed costs of the facilities, f^1_i and $f^2_i, i \in \mathcal{N}$, are generated randomly and independently on the interval $[50, 500]$. Values of the parameters $w_i, i \in \mathcal{N}$ are generated randomly and independently on the interval $[50, 100]$. Finally, the value of the parameter θ is given by $\theta = 0.339$.

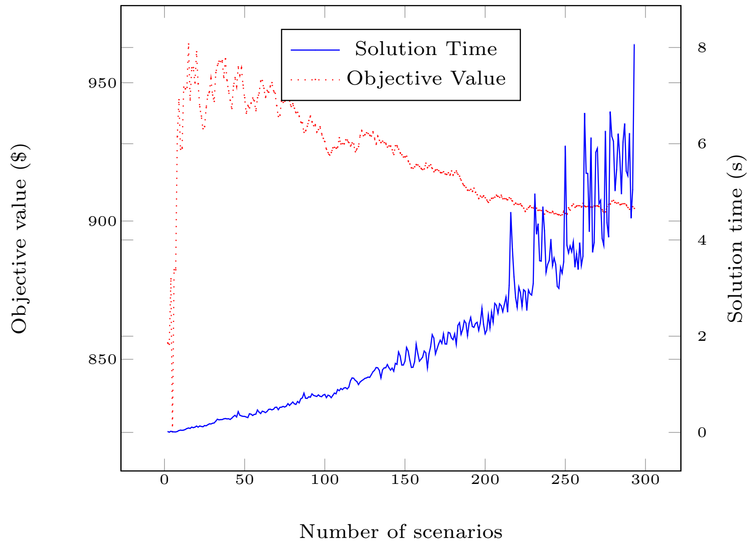

Figure 2: Objective value and solution time of $\mathtt{I}_1$ stopping at $m = 293$.

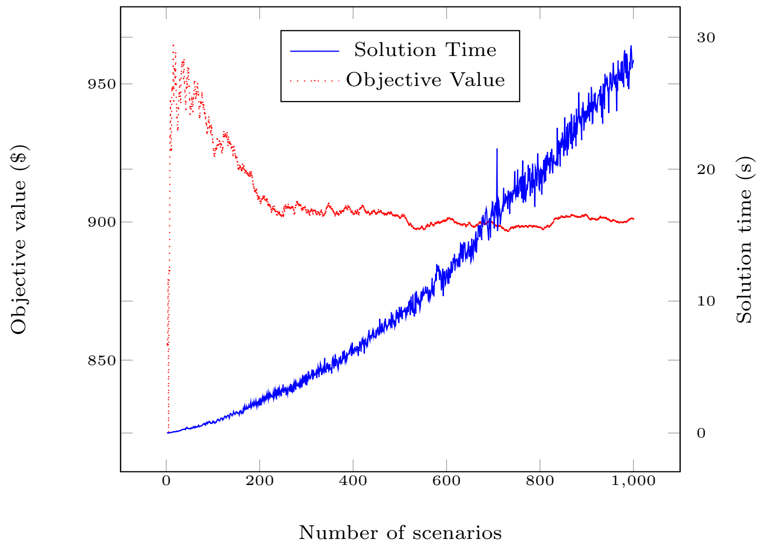

Figure 3: Objective value and solution time of $\mathtt{I}_1$ stopping at $m = 1000$.

In order to find an appropriate number of required scenarios, we randomly generate 100 instances of the deterministic problem. More precisely, for each instance, we generate all parameters of the deterministic formulation (1)–(8) except the demand values. Then, for each instance, we start with $m = 1$ scenarios and increment the number of scenarios until the following stopping criteria are met. In order to explain these criteria, let us define λ_m^k as the optimal value of the objective function (9) for instance $k \in \{1, \ldots, 100\}$ when m scenarios are considered. We stop increasing the number of scenarios if $m \geq 100$ and

$$\frac{\max L_m^k - \min L_m^k}{\min L_m^k} < 0.01, \tag{16}$$

where $L_m^k := \{\lambda_{m-99}^k, \ldots, \lambda_m^k\}$. For each instance, we keep the previous generated scenarios unchanged and add one additional scenario. Figure 2 illustrates the impact of the number of scenarios on the objective value and solution time of the first randomly generated instance (henceforth referred to as $\mathtt{I}_1$). This figure shows that the objective value converges to around \$900 by incrementing the number of scenarios. For this instance, the stopping criteria are satisfied at $m = 293$. In Figure 3 we ignore the stopping criteria and increment the number of scenarios up to 1000 for this particular instance. The figure displays that 293 is an appropriate number of scenarios for this instance. Surprisingly, our experiments on 100 randomly generated instances indicate that on average 307 scenarios are required to satisfy the stopping criteria.

5 Conclusion

In this paper, we studied a whey reverse logistics network design problem under demand uncertainty. We first formulated the deterministic version of the problem as a hierarchical facility location model. Then a two stage stochastic programming approach was taken to address the problem with stochastic demand. We utilized the sample average approximation method to estimate the expected value function. We solved the resulting stochastic programming formulation with a standard linear solver for small instances of the problem with 5 nodes. Our experiments on 100 randomly generated instances indicated that on average 307 scenarios were required to satisfy our proposed stopping criteria. A future direction of this research is to employ more advanced algorithms to solve real-world instances of the problem with as many as 50 nodes.

References

Agrawal, S., R. K. Singh, and Q. Murtaza (2015). A literature review and perspectives in reverse logistics. *Resources, Conservation and Recycling 97*, 76 – 92.

Bansal, N. and B. Bhandari (2016). Functional milk proteins: Production and utilization—whey-based ingredients. In P. L. H. McSweeney and J. A. O'Mahony (Eds.), *Advanced Dairy Chemistry: Volume 1B: Proteins: Applied Aspects*, pp. 67–98. New York, NY: Springer New York.

Daskin, M. S. (2013). Extensions of location models. In *Network and Discrete Location: Models, Algorithms, and Applications, Second Edition*, Chapter 8, pp. 362–479. New Jersey: John Wiley & Sons, Inc.

de Brito, M. P. and R. Dekker (2004). A framework for reverse logistics. In R. Dekker, M. Fleischmann, K. Inderfurth, and L. N. Van Wassenhove (Eds.), *Reverse Logistics: Quantitative Models for Closed-Loop Supply Chains*, pp. 3–27. Berlin, Heidelberg: Springer Berlin Heidelberg.

García-Flores, R., R. Martins, O. V. de Souza Filho, M. González, C. Mattos, A. Rosenthal, and P. Juliano (2015). A novel facility and equipment selection model for whey utilisation: A brazilian case study. *Computers and Electronics in Agriculture 117*, 127–140.

Melo, M. T., S. Nickel, and F. Saldanha-Da-Gama (2009). Facility location and supply chain management–a review. *European Journal of Operational Research 196*(2), 401–412.

Narula, S. C. (1984). Hierarchical location-allocation problems: a classification scheme. *European Journal of Operational Research 15*(1), 93–99.

Nath, A., B. Verasztó, S. Basak, A. Koris, Z. Kovács, and G. Vatai (2016). Synthesis of lactose-derived nutraceuticals from dairy waste whey—a review. *Food and Bioprocess Technology 9*(1), 16–48.

Patel, S. (2015). Emerging trends in nutraceutical applications of whey protein and its derivatives. *Journal of Food Science and Technology 52*(11), 6847–6858.

Şahin, G. and H. Süral (2007). A review of hierarchical facility location models. *Computers & Operations Research 34*(8), 2310–2331.

Shapiro, A. (2008). Stochastic programming approach to optimization under uncertainty. *Mathematical Programming 112*(1), 183–220.

Smithers, G. W. (2015). Whey-ing up the options yesterday, today and tomorrow. *International Dairy Journal 48*, 2 – 14.

Yadav, J. S. S., S. Yan, S. Pilli, L. Kumar, R. Tyagi, and R. Surampalli (2015). Cheese whey: A potential resource to transform into bioprotein, functional/nutritional proteins and bioactive peptides. *Biotechnology Advances 33*(6, Part 1), 756 – 774.