Econometric Institute Report EI 2001-08

Stochastic approaches for product recovery network design: a case study

Ovidiu Listes*, Rommert Dekker†

Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

Abstract

Increased uncertainty is one of the characteristics of product recovery networks. In particular the strategic design of their logistic infrastructure has to take uncertain information into account. In this paper we present stochastic programming based approaches by which a deterministic location model for product recovery network design may be extended to explicitly account for the uncertainties. Such a stochastic model seeks a solution which is appropriately balanced between some alternative scenarios identified by field experts. We apply the stochastic models to a representative real case study on recycling sand from demolition waste in The Netherlands. The interpretation of the results is meant to give more insight into decision-making for reverse logistics.

Keywords: Location, Reverse Logistics, Uncertainty, Stochastic Programming

1 Introduction

The design of product recovery networks is one of the challenging and actual reverse logistics problems. Stimulated by environmental, legislative and economical reasons several recovery networks have been set up in the last decade. Besides paper and glass nowadays building waste, white-and-brown goods, electric and electronic equipment is being collected for recycling in The Netherlands and other countries in Europe. The set-up of a product recovery network is however, not similar to that of a forward logistic network. Fleischmann et al (2000) state that product recovery not only reverses the product stream with the consequence that there are many supply sources and few demand points, but that the design is severely complicated by the high uncertainty in many factors. This is especially true for those recovery networks that are driven by legislative or environmental reasons (typically for Western Europe). In this case there is large uncertainty where the recovered product should go to and moreover, some supply may leak away to less costly disposal ways. Furthermore often new technology is developed for recycling, with the consequence that realized cost figures can deviate substantially from the foreseen values.

The design of logistic networks has long been the application of location models within operations research (see Geoffrion and Powers (1995)). Several algorithmic developments coupled with increases in computer speed now allow us to solve real-world problems to (almost) optimality with commercial software like GAMS or CPLEX (see Brooke et al. (1988), ILOG (1999)). The expectations on the modeling however, have gone up with these developments and the handling of uncertainty is one of the aspects which needs to be tackled, especially in reverse logistic networks. A standard way to tackle uncertainty is to do a single or multi-parameter sensitivity analysis (see Morgan and Henrion (1990)), which gives insight into how the solution and the costs change if one or more input parameters are varied. This approach can be extended by introducing scenarios for the input parameters and finding the individual solution that performs best over the set of scenarios. The drawback of these approaches, however, is that they stick to solutions which are optimal for at least one set of parameters. Kall and Wallace (1994) claim that stochastic programming techniques offer more flexibility for handling uncertainty and can come up with solutions that can not be found by scenario analysis. The application of stochastic programming to integer location problems is however a challenging research subject. Stochastic integer models are

*e-mail: listes@few.eur.nl

†e-mail: rdekker@few.eur.nl

well-known for their computational intractability, since they combine two types of models which are by themselves difficult to solve. Besides theoretical developments, important algorithmic approaches have also been proposed, but the instances they can address seem still quite limited in size. Moreover, there are only few applications of stochastic programs to location problems. A representative example is a branch and cut procedure proposed by Laporte et al. (1994) for solving a location problem with stochastic demands. This procedure is repetitively applied to a large number of smaller sized instances of the problem, obtained by parametric variation of some other parameters (different than demand), one parameter at a time.

To the best of our knowledge stochastic programming is hardly applied to practical situations of logistic network design. In this paper we present stochastic programming based approaches by which a large-scale location model for product recovery network design may be extended to account for uncertainties. We apply the stochastic models to a case study concerning the recycling of sand from demolition waste, which was done some years ago (see Barros et al (1998)). Although there was quite some uncertainty involved in this case, it was previously tackled by scenario analysis only. As besides all data, we have some real information on the uncertain aspects, the stochastic programming approaches are employed for explicitly taking this information into account. Since the aim is to develop insights for problems with real-world dimension, the construction of the stochastic models deliberately follows a rather simple technique, which may be potentially used to extend in a reasonable manner any large location model in which uncertainty is an issue and a relatively small set of realistic scenarios can be identified.

This paper is structured as follows. In the next section we describe the case study in more detail. In section 3 we discuss the modeling of the problem and indicate how the deterministic location model may be extended using stochastic programming techniques, according to the available information on the uncertain factors. The results of the stochastic programming approach as applied to the presented case are discussed in section 4. Finally, section 5 contains some concluding remarks.

2 Recycling of sieved sand: the case study

The case concerned the re-use of sieved sand originating from crushing and sieving of construction waste in The Netherlands. This sand used to be landfilled, although it could be polluted with poly-aromatic carbonates. This pollution also prevented an easy re-use of the sand. Legislation was adopted forbidding landfilling of the sand and prescribing recycling. In order to achieve this, the syndicate of building waste processors proposed the set-up of a network with depots storing the sand from the crushers. In the depots the sand is tested on its pollution by an independent organization. Three categories of sand are distinguished, clean sand which can directly be used, half-clean sand, which can be applied in secured applications and finally polluted sand, which needs to be cleaned. Next cleaning facilities are envisaged to clean the polluted sand. These involve high investments.

The problem was to get insight into the logistical costs with setting up such a network and to give advise about locations of both the storage depots and the cleaning facilities. A facility location model was set-up and solved by a combination of valid inequalities, heuristic rounding off procedures and iterative exact solving of sub-problems. Excel, Turbo Pascal and GAMS software was used for methods and for input and output of the data. The resulting software ran on a PC and was handed over to the syndicate to update the advise when new data would become available (see Barros et al. (1998)).

This case is representative for many product recovery networks as remarked by Fleischmann et al. (2000). It has many sources, high investments costs for the recycling installation with the implication that only few will be built, not yet tested recycling technology and unclear destinations of the recycled products. Also the stream splitting is typical for reverse logistics. The aspect of two layers is often but not always the case and it merely affects the solution procedures. Similar product recovery networks concern for instance carpet recycling (see Ammons et al. (1999), Flapper et al. (1997)). We therefore think that the results obtained in this paper are of more general nature than just the case discussed here. The title reflects this tendency as the aim is to gain insight into the general issue of product recovery network design through the framework offered by this case.

3 Modeling

As mentioned, the case at hand concerns the design of a two-level network in which storage and cleaning tasks are to be located across a set of potential sites (storage in regional depots and cleaning in treatment facilities). This network is to extend existing supply sources and to direct recycled material to end use demand points. The sources of sieved sand correspond to 33 existing sorting facilities. The lack of information about future projects requiring sand was handled by selecting 10 strategic sites as potential points of demand. These initial elements are represented in Figure 1 (some triangles represent multiple sources). The new facilities that must be created are regional depots for classifying the sand (86 potential sites) and treatment facilities for cleaning the polluted sand (21 potential sites). Transport of sand throughout the network can be done by water between reachable sites, which is much cheaper than by road, the mode employed otherwise. The resulting network structure is schematically given in Figure 2.

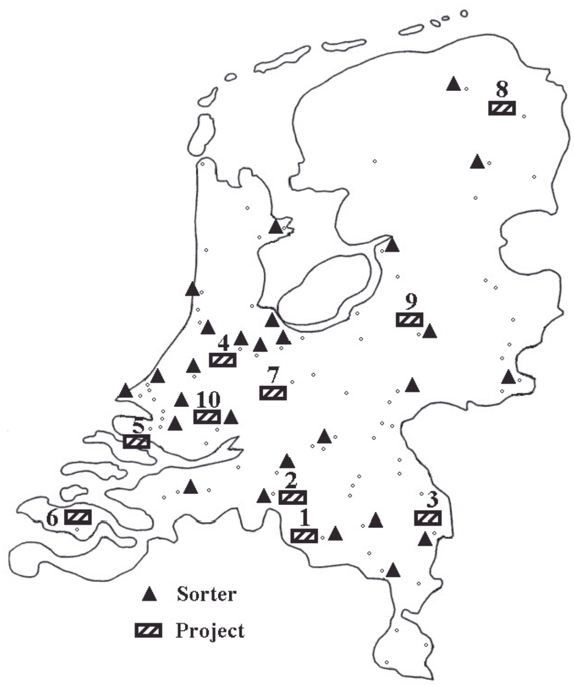

Figure 1: Sorters and potential projects

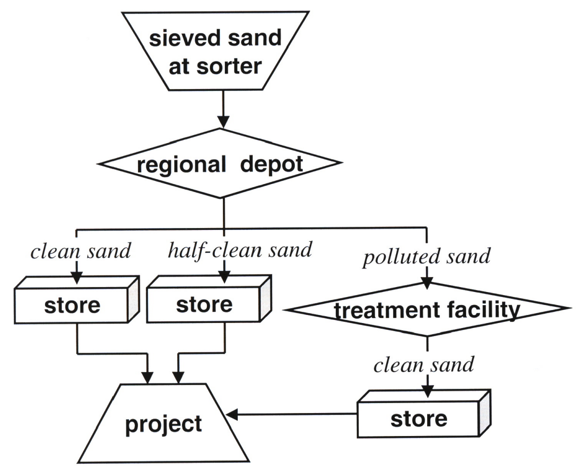

Figure 2: Sand network structure

In the model set-up for this problem we use as objective function the net revenue computed as the value of fees charged for sand entering the network plus the revenues from selling recycled sand to projects minus various costs involved. These costs are fixed costs for opening the new facilities as well as variable processing and transportation costs based on amounts of sand processed and respectively shipped between sites. The objective function is determined on a yearly basis. Consequently annualized fixed costs for opening new facilities are computed by amortizing the total investments over the planning horizon. Based on available technology for cleaning only one type of installation with rather high fixed costs was possible. Nevertheless a case with several cleaning capacities could also be addressed by considering treatment at each capacity as a separate investment option, probably associated with specific sites. For simplicity the fixed costs of regional depots are assumed to vary as an affine function of their storage capacity. Also processing costs depend only on the type of facility and transportation costs are proportional to the amount and the distance over which it is transported. The underlying model can be verbally described as

Maximize ***Net Revenue***
= Fees + Sales Revenue
− Fixed opening costs − Transportation Costs − Processing Costs

Subject to: ***Balances of flows between sites***
(e.g. based on material shipped from depots to facilities)
Context specific constraints
(storage is allowed; demand may be partially satisfied)
Capacity constraints (upper bounds)
(on storage at depots and on processing at facilities)

This formulation slightly differs from the one in Barros et al. (1998) in order to allow for further investigations rather than for setting a marketing perspective. Its mathematical description can be found in the appendix. The stochastic extensions proposed below further concentrate on finding the appropriate location for the new facilities. The focus on the location issue is motivated by the fact that investments are made on long run and are based on uncertain information. Since these networks are expensive and difficult to change, establishing a more robust location for the new facilities becomes a central issue at stake. Therefore the models are primarily configured with variable infrastructure to assess where to locate the new facilities, but they can be as well configured with all the infrastructure fixed to investigate operational aspects of candidate network configurations.

3.1 Data and uncertainty

Important data used in implementations are presented in Table 1 (see Barros et al. (1998) for details).

Description	**Value**
Total available sand at sorters (initial estimation)	992,400 tons per year
Processing capacity of a cleaning facility	150,000 tons per year
Storing capacities of potential depots	from 1,200 to 240,000 tons per year
Fixed costs for opening a cleaning facility	4,057,125 Dfl per year
Fixed costs for opening a depot with 150,000 capacity	670,625 Dfl per year
Transportation costs by road	0.09 Dfl per ton per km
Transportation costs by water	0.02 Dfl per ton per km
Handling costs at regional depots	1 Dfl per ton
Processing costs at cleaning facilities	45,74 Dfl per ton

Table 1: Data used in implementations

A characteristic of the sand problem is that the infrastructure to be created must be able to handle the whole amount of sand coming from sorters. This kind of requirement originates in newly emerged legislation and may apply as well to similar systems with different product features. The amount of material that may be expected becomes therefore a very powerful driver of investments. Nevertheless the amount and quality of the returned flows may be uncertain, because the initial quality mix is unknown and companies may choose to recycle the sand in another way (the law prescribes recycling of sand, but not the way how). In particular the actual supply may turn out to be (much) less than expected. We refer to the initial estimation as the *high supply case* and besides this we also consider a *low supply case* deemed relevant for the given capacity of a cleaning facility. The initial estimation was done for each source separately and as amounts are decreased, it is done uniformly throughout these sources. The percentages of clean, half-clean and respectively polluted sand are different in the two cases and induce different fees charged for sand entering the network.

Sand type	Proj.1	Proj.2	Proj.3	Proj.4	Proj.5	Proj.6	Proj.7	Proj.8	Proj.9	Proj.10
Clean	80,000	50,000	50,000	0	80,000	50,000	80,000	0	80,000	200,000
Half-clean	50,000	50,000	75,000	90,000	60,000	50,000	100,000	50,000	100,000	150,000

Table 2: Demand of sand in tons per year

Estimations of demand of each type of sand for the 10 potential projects considered are given in Table 2. In order to handle the lack of information about the location of the actual projects the following scenarios were established:

Basis scenario: projects 3, 6, 7, 8 (overall)
Scenario 1 : projects 2, 5, 8, 10 (south-west and north-east)
Scenario 2 : projects 3, 8, 9 (along east region)
Scenario 3 : projects 5, 6, 10 (center and south-west)
Scenario 4 : projects 3, 9, 10 (center and south-east)

Scenario 5 : projects 1, 4, 6, 7, 9 (center and south)
Scenario 6 : projects 2, 5, 8, 9, 10 (south-west, center and north-east)

In this way demand varies not only in geographical distribution but also in *total* amounts of clean respectively half-clean sand required. The underlying assumption is that such scenarios may be established by experts in the field as representing reasonable future developments. Construction of scenarios above follows the previous form in Scholten (1995) rather than the one in Barros et al. (1998).

3.2 The stochastic programming approach

Define Φ as the set of all possible scenarios and $\phi \in \Phi$ as a particular scenario. The complete mathematical description of the underlying deterministic model for scenario ϕ is included in the appendix. In order to describe the main ideas of the stochastic approach we use here a concise notation for the deterministic model: all the integer decision variables are included into one vector y of dimension m and all the continuous decision variables are included into one vector x of dimension n. The notation for the coefficients is adjusted accordingly: f is the m-dimensional vector of the fixed costs for opening facilities and c is the vector of dimension n containing the rest of the coefficients in the objective function. Then the concise deterministic model for scenario ϕ can be stated as

$$
\begin{array}{lrcl}
\max & -f\,y + c(\phi)x & & \\
\text{s.t.} & W_0(\phi)\,x & = & b(\phi) \\
 & W_1(\phi)\,x & \leq & d(\phi) \\
 & T\,y \;+\; W_2(\phi)\,x & \leq & 0 \\
 & y \in \{0,1\}^m,\ x \in \mathbb{R}^n,\ x \geq 0 & &
\end{array}
$$

where W_0 is a $n_0 \times n$ matrix, b is a n_0 vector, W_1 is a $n_1 \times n$ matrix, d is a n_1 vector, W_2 is a $n_2 \times n$ matrix and T is a $m\prime \times m$ matrix. Solutions given by this model for each scenario ϕ separately form the basis for scenario analysis (see Barros et al. (1998)).

In a stochastic programming approach probabilities are associated with scenarios and a solution is sought which is suitably balanced against the various scenarios (see Birge and Louveaux (1997), Kall and Wallace (1994)). The stochastic solution is not optimal in general for any of the individual scenarios. We use this approach in the sequel to address specific questions about the impact of demand/supply uncertainty on network design.

3.2.1 Locational uncertainty of demand

We denote by Ω the set of demand scenarios and by ω a particular demand scenario (here Φ coincides with Ω and ϕ with ω). Only the right hand side parameter d is determined to change in demand scenarios.

In order to approach this locational uncertainty of demand we extend the deterministic model above to the following two-stage stochastic model: the first stage corresponds to the investments that must be made for opening facilities prior to knowing the actual realizations of the random parameters and the second stage corresponds to the allocation of flows through the established network after the values of the random parameters become known. Consequently the location variables y are assigned as first stage variables and the allocation variables x are assigned as second stage variables. For each pair (y, ω) the performance measure is given by the second stage program as the maximum revenue that can be achieved in the network given by location y and under scenario ω. The decision y is evaluated this way across all scenarios and the expected net revenue is recorded as the indicator of the decision. The first stage program aims then at maximizing this indicator over the set of all possible first stage decisions. So the two-stage stochastic programming model states as

$$
\begin{array}{llcl}
\max & -f\,y + \mathbb{E}_{\omega}[\,Q(y,\omega)\,] & \quad\text{where}\quad & Q(y,\omega) \;=\; \max\; c\,x \\
\text{s.t.} & y \in \{0,1\}^m & & \text{s.t.} \quad W_0\,x \;=\; b \\
 & & & \qquad\;\; W_1\,x \;\le\; d(\omega) \\
 & & & \qquad\;\; W_2\,x \;\le\; -T\,y \\
 & & & \qquad\;\; x \in \mathbb{R}^n,\; x \ge 0
\end{array}
$$

Clearly once a value of y is fixed in the first stage, second stage decisions depend on scenario and thus variables x are likely to change under different realizations of ω. Therefore when random parameters follow (finite) discrete distributions it is useful to index the second stage variables by ω in order to assess costs and benefits in each situation. This creates a separate set of variables of the form x_ω for each scenario. Using this explicit description of the second stage variables for all scenarios the problem can be stated in the following *extensive form* (see Birge and Louveaux (1997) for this term):

$$
\begin{array}{llcll}
\max & -f\,y + \mathbb{E}_{\omega}\,[\,c\,x_{\,\omega}\,] & & & \\
\text{s.t.} & \qquad\qquad W_0(\omega)\,x_{\,\omega} & = & b & \forall\,\omega\in\Omega \\
 & \qquad\qquad W_1(\omega)\,x_{\,\omega} & \le & d(\omega) & \forall\,\omega\in\Omega \\
 & T\,y \;\;+\;\; W_2(\omega)\,x_{\,\omega} & \le & 0 & \forall\,\omega\in\Omega \\
 & y \in \{0,1\}^m & & & \\
 & x_{\,\omega} \in \mathbb{R}^n,\; x_{\,\omega} \ge 0 \quad \forall\,\omega\in\Omega & & &
\end{array}
$$

Since the expectation involved is in this case just an ordinary sum, the last formulation is a (large scale) mixed integer linear programming model. This form was used for implementing the two-stage stochastic model in two different variants: one with fixed low supply and uncertain demand, the other with fixed high supply and uncertain demand.

3.2.2 Additional uncertainty of supply

We further consider a set Ξ of supply scenarios, independent of demand scenarios, and denote by ξ a particular supply scenario (here $\Phi = \Xi \times \Omega$ and one individual scenario ϕ consists of a possible combination (ξ,ω) of supply–demand realizations). In this case the solution given by one scenario with low supply realization may open few facilities and thus be infeasible for other scenario with high supply realization. On the other hand the solution given by one scenario with high supply may open many facilities and thus be too costly for a scenario with low supply. Therefore scenario analysis is not applicable in such a situation. Moreover a two-stage model is also not appropriate, since it would "protect" the scenarios with high supply, that is it would open 2 cleaning facilities and thus such a solution would be again too costly. A possible remedy could be to consider extra costs for penalizing infeasibility. However such penalty costs are difficult to estimate in a meaningful manner and this is therefore less realistic.

In order to cope with this situation we used a *three-stage* stochastic programming model, in which location decisions are made as well in the first stage as in the second stage, whereas the corresponding flows decisions are reserved for the third stage. Namely we assume that the common supply of low and high scenarios has to be met in the first stage, after which the actual supply is expected to be revealed and consecutive decisions have to be made if the supply is high. In either case the material is processed and sold to projects in the third stage according to the seven demand scenarios, such that the overall expected revenue is maximized. The situation is schematically rendered by the scenario tree in Figure 3. Such a modeling is not unrealistic since the cleaning facilities are likely to be built one by one rather than at the same time. Denoting by z, y the location decisions for the first respectively the second stage and

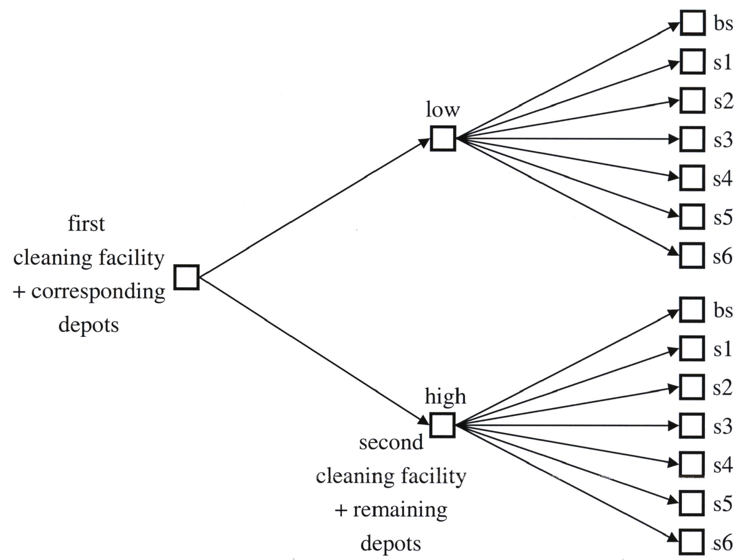

Figure 3: Scenario tree in the three-stage approach

by x the allocation decisions in the third stage, the three stage stochastic programming model states as

$$\begin{array}{ll} \max & -f\,z + \mathbb{E}_\xi[\,Q(z,\xi)\,] \\ \text{s.t.} & z \in \{0,1\}^m \end{array} \qquad \text{where} \qquad \begin{array}{rcl} Q(z,\xi) & = & \max \quad -f\,y + \mathbb{E}_\omega[\,Q_0(z,y,\xi,\omega)\,] \\ & & \text{s.t.} \quad y \in \{0,1\}^m \end{array}$$

and respectively where

$$\begin{aligned} Q_0(z,y,\xi,\omega) &= \max\ c(\xi)x \\ \text{s.t.}\quad W_0(\xi)\,x &= b(\xi) \\ W_1(\xi)\,x &\le d(\omega) \\ W_2(\xi)\,x &\le -T\,(z,y) \\ x \in \mathbb{R}^n,\ x &\ge 0 \end{aligned}$$

Here (z,y) is the $2m$-dimensional vector obtained by concatenating z and y and T is a $m\prime \times 2m$ matrix. The last block of constraints also contains requirements that at each site one type of facility may be opened in at most one stage. As indicated above, the parameters c, W_0, b, W_1, W_2 depend on supply scenarios, whereas d still depends on demand scenarios as before. Indexing y over ξ and x over (ξ,ω) yields the following extensive form, used for implementation:

$$\begin{array}{llll} \max & -f\,z + \mathbb{E}_\xi\,[-f\,y_\xi + \mathbb{E}_\omega\,[\,c(\xi)\,x_{\xi,\omega}\,]\,] & & \\ \text{s.t.} & W_0(\xi)\,x_{\xi,\omega} = b(\xi) & \forall\xi\in\Xi\,,\ \forall\omega\in\Omega \\ & W_1(\xi)\,x_{\xi,\omega} \le d(\omega) & \forall\xi\in\Xi\,,\ \forall\omega\in\Omega \\ & T\,(z,y_\xi) + W_2(\xi)\,x_{\xi,\omega} \le 0 & \forall\xi\in\Xi\,,\ \forall\omega\in\Omega \\ & z \in \{0,1\}^m & \\ & y_\xi \in \{0,1\}^m \quad \forall\xi\in\Xi & \\ & x_{\xi,\omega}\in\mathbb{R}^n,\ x_{\xi,\omega}\ge 0 \quad \forall\xi\in\Xi\,,\ \forall\omega\in\Omega & \end{array}$$

This type of approach is not yet observed in the literature of logistic network design. Yet it arises naturally in our context because decisions must be hedged against very different levels of return flows. As this implies rather high risks if approached in just one step, the decisions have to be split over time based on the assumption that extra information may be gradually acquired. We therefore believe that given the information we have on the uncertainty, the three stage model above is the appropriate approach in this case. Moreover we think that such a multi-stage approach may also shed light on more efficient design of similar product recovery networks as well as in other situations where (forward) logistic networks are to be developed stepwise over time.

3.2.3 Implementation

We used GAMS as modeling language and the mixed integer solver from CPLEX6.5 commercial software for all the variants of the problem (see Brooke et al. (1988) and ILOG (1999)). The original models were extended with several valid inequalities, e.g. expressing logical relations between the continuous flow between two sites and the integral indicators associated to those sites (see Barros et al. (1998)). Priority orders on branching were assigned to the integer variables. In the two stage models a higher priority was assigned to the variables corresponding to the cleaning facilities. In the three stage model four levels were distinguished with priorities decreasing in the following order: cleaning facilities first stage, cleaning facilities second stage, depots first stage, depots second stage.

Models	Number of variables		Number of constraints		Running time
	integer	continuous	initial	valid inequalities	
original deterministic	107	6,574	439	5,064	5 min
two-stage { low supply	107	46,018	3,073	35,448	3 h 30 min
two-stage { high supply					11 h
three-stage	214	92,036	6,146	70,896	28 h

Table 3: Dimensions of the models and running times

The underlying deterministic model for each scenario and the two-stage stochastic models in extensive form were solved on a Windows NT-based 450MHz Pentium III PC with 128MB of memory. The average running time for separate scenarios was about 5 minutes. The solution time for the two-stage model was about 3 and a half hours in the low supply case and about 11 hours in the high supply case. The three-stage model was solved on a Sun Enterprise250 UltraSPARCII-400MHz UNIX system with 1GB of memory (shared resources). The CPU running time was about 28 hours. The second computational system features a lower speed, but it was employed in the three-stage case for its capacity to handle larger sets of data during the computation. Besides the models sizes, the tendency to extensive search is to large extent explainable by the *capacitated* feature of the problem and by a relative symmetry of the potential sites in terms of the costs they incur. Details on the implemented models are given in Table 3.

4 Results

The results presented in the sequel are based on a sales price of 27 Dfl per ton of clean sand and 16 Dfl per ton of half-clean sand, which are our estimations. Since investments are mainly driven by legislation, and to a less extent by marketing perspective, variations in sales prices have less influence on network layout. Accordingly less variations in the net revenues are expected (from this point of view our investigations differ from those reported by Newton et al. (1999)). The charged fees depend on quality of supplied sand and significantly increase as quality gets worse. Equal probabilities were associated with demand scenarios as no available information justified another distribution. Variations of these (subjective) probabilities and the corresponding outcomes may be addressed in a separate discussion.

4.1 High supply case

The amount of supplied sand equals in this case the initial estimation of 992,400 tons per year from which 40% is clean (396,960 tons), 40% is half-clean (396,960 tons) and 20% is polluted (198,480 tons). The charged fee corresponding to this quality amounts to 21,50 Dfl per ton (our estimation). Given the cleaning capacity of 150,000 tons per year it turns out that 2 cleaning facilities have to be opened. Yet the number of regional depots and the location of all the facilities are determined taking the seven demand scenarios into account. The stochastic configuration differs from any of the optimal configurations of individual scenarios on both regional depots and cleaning facilities. It is illustrated in Figure 4. The stochastic solution yields a network with 19 regional depots, from which 9 are accessible by water. Given their rather small fixed costs, the regional depots handle locally the flows of sand. However, besides the sources of sand, their location is now influenced by all the projects with the weights induced together by the seven scenarios. Therefore mostly the sites which by their size and position are able to serve efficiently several projects are retained in the stochastic solution. This effect is also illustrated by the fact that more than 99% of the storage capacity of the opened depots is used. The optimal location of the cleaning facilities for the individual scenarios and the stochastic solution are presented in Table 4.

	Scenarios							**Stochastic**
Cleaning facilities	**bs**	**s1**	**s2**	**s3**	**s4**	**s5**	**s6**	**solution**
Alblasserdam		•		•	•			
Dongen							•	
Oosterhout								•
Utrecht	•		•			•		
Westervoort							•	
Almelo	•	•	•	•	•	•		•

Table 4: Individual scenarios solutions and stochastic solution in high supply case

When evaluating the stochastic configuration over each of the scenarios the following pattern is observed. The cleaning facility located in Oosterhout (center west) handles scenario 6, where the two facilities handle comparable yearly amounts, about 100,000 tons either. While the facility in Almelo is also frequent in scenarios, the one in Oosterhout is not reproduced by any scenario. Instead it makes a necessary trade-off between the sites from center or center-west regions separately chosen in the scenarios.

Scenarios	**Optimal net revenue**	**Stochastic net revenue**	**Optimal−Stochastic net revenue**	**Stochastic = % of Optimum**
bs	2,720,640	2,338,988	381,652	86.0
s1	7,727,527	7,627,766	99,761	98.7
s2	492,454	235,472	256,982	47.8
s3	6,718,106	6,436,475	281,631	95.8
s4	7,710,768	7,573,833	136,935	98.2
s5	6,463,536	6,095,327	368,209	94.3
s6	14,091,545	13,822,662	268,883	98.1
Expected	6,560,653	6,304,360	256,293	96.1

Table 5: Optimal values for scenarios and stochastic solution in high supply case

Scenarios	**Optimal net revenue**	**Worst net revenue**	**Optimal−Worst net revenue**	**Worst = % of Optimum**
bs	2,720,640	2,125,853	594,787	78.1
s1	7,727,527	7,161,612	565,915	92.6
s2	492,454	33,892	458,562	7.0
s3	6,718,106	6,253,260	464,846	93.0
s4	7,710,768	7,117,634	593,134	92.3
s5	6,463,536	5,802,199	661,337	89.7
s6	14,091,545	13,279,936	811,609	94.2
Expected	6,560,653	5,967,769	592,884	90.9

Table 6: Worst case analysis for high supply case

Table 5 shows the differences between the revenue given by the optimal solution of each scenario and the revenue given by the stochastic solution for the same scenario. The expected revenue generated by the stochastic solution is 96.1% from the weighted average of individual optima (the hypothetical average with perfect information). The worst case location for one scenario is the configuration taken from the seven scenarios which generates the lowest revenue for that scenario. The corresponding values are given in Table 6. The expected worst case revenue is 90.9% from the weighted average of individual optima. The solution from the scenarios that achieves the maximum expected revenues (the least expected regret) corresponds to scenario 3. The stochastic solution generates better revenues for all the other scenarios (except scenario 3) and an improvement of 0.83% in expectation (yearly) over the solution of scenario 3. The improvement in terms of expected revenues looks rather small in this case. Nevertheless the more uniform distribution of revenues over scenarios as given by the stochastic solution has to be remarked. Sticking to the individual solution of one scenario may yield high revenues for some scenarios, but also poor results for others as indicated in the worst cases table for scenario 2.

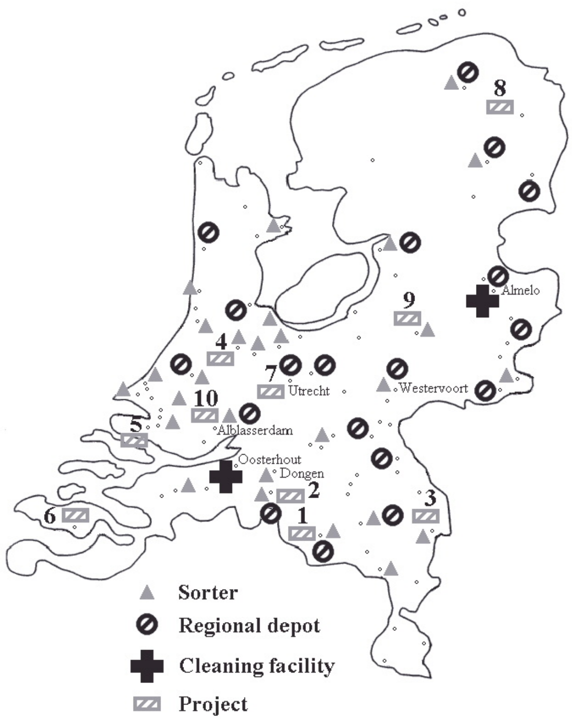

Figure 4: Stochastic solution in high supply case

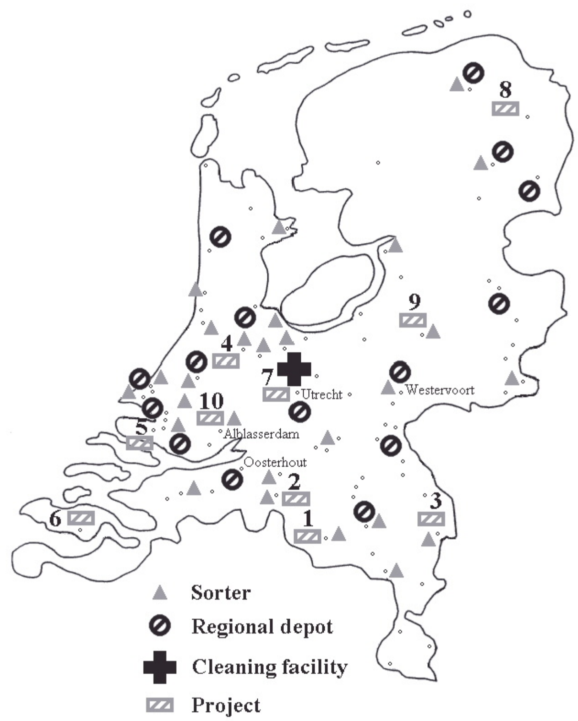

Figure 5: Stochastic solution in low supply case

4.2 Low supply case

In this case the supplied sand amounts to 496,200 tons per year, that is half of the initial estimation. From this 20% is clean (99,240 tons), 50% is half-clean (248,100 tons) and 30% is polluted (148,860 tons). A fee of 23,25 Dfl per ton corresponding to this quality is charged (our estimation). Clearly in this situation only one cleaning facility is necessary, which is to work at (almost) full capacity. Also in this case the stochastic configuration does not coincide with any of the optimal configurations for individual demand scenarios. It is depicted in Figure 5.

	Scenarios							Stochastic solution
Cleaning facilities	**bs**	**s1**	**s2**	**s3**	**s4**	**s5**	**s6**	
Westervoort			•			•		
Oosterhout	•							
Utrecht								•
Alblasserdam		•		•	•		•	

Table 7: Individual scenarios solutions and stochastic solution in low supply case

The stochastic solution yields now a network with 16 regional depots, from which 13 are reachable by water. The more reduced number of depots required in this case forces each regional depot to serve a larger subregion. Consequently the transportation increases and more depots are opened at sites accessible by water, often with a higher capacity. The stochastic solution typically retains this kind of sites and moreover those of them properly balanced between scenarios, such that again more than 99% of the storage capacity of the opened depots is used. The optimal location of the cleaning facilities for each of the scenarios and the stochastic solution are presented in Table 7. A stronger balancing effect is remarked here as the location of the only cleaning facility in Utrecht (center of the country) makes a clear trade-off between the sites in center-west or east regions preferred in the separate scenarios. Similarly to the previous case, values generated by the stochastic solution and the worst cases are presented in Table 8 and respectively Table 9. In this case the expected revenue generated by the stochastic solution is 95.6% from the weighted average of individual optima, whereas the expected worst case revenue is 86.5% from the same amount. The solution taken from the scenarios that achieves the maximum expected revenues (the least expected regret) corresponds in this case to scenario 4. The stochastic solution generates better revenues for all the scenarios except scenarios 3 and 4 and an improvement of 2.31% in expectation (yearly) over the solution of scenario 4. The improvement in terms of expected revenues is higher in this case, but still limited through the nature of the investigated factors. The main gains remain a more robust configuration and a more uniform distribution of revenues over scenarios as given by the stochastic solution.

Scenarios	**Optimal net revenue**	**Stochastic net revenue**	**Optimal−Stochastic net revenue**	**Stochastic = % of Optimum**
bs	3,389,774	3,316,781	72,993	97.8
s1	5,797,840	5,555,551	242,289	95.8
s2	1,322,112	1,095,451	226,661	83.0
s3	5,457,753	5,148,180	309,573	94.3
s4	5,723,589	5,485,855	237,734	95.8
s5	4,829,696	4,681,254	148,442	96.9
s6	5,874,306	5,680,799	193,507	96.7
Expected	4,627,867	4,423,410	204,457	95.6

Table 8: Optimal values for scenarios and stochastic solution in low supply case

Scenarios	**Optimal net revenue**	**Worst net revenue**	**Optimal−Worst net revenue**	**Worst = % of Optimum**
bs	3,389,774	2,845,208	544,566	83.9
s1	5,797,840	5,133,495	664,345	88.5
s2	1,322,112	482,576	839,536	36.5
s3	5,457,753	4,750,406	707,347	87.0
s4	5,723,589	5,203,795	519,794	90.9
s5	4,829,696	4,303,279	526,417	89.1
s6	5,874,306	5,319,311	554,995	90.5
Expected	4,627,867	4,005,438	622,429	86.5

Table 9: Worst case analysis for low supply case

4.3 Three-stage approach

As explained in the modeling section, this case is based on the assumption that after opening some facilities in a first stage, extra information about the actual supply can be acquired, based on which further location decisions are to be made in a second stage. Namely we assume that at least a yearly amount of 496,200 tons (20% clean, 50% half-clean, 30% polluted) will be supplied and the remaining amount up to 992,400 tons (40% clean, 40% half-clean, 20% polluted) will be supplied or not with equal chances. Therefore the overall problem states now differently: where to locate the first cleaning facility in the first stage such that the second treatment facility will be opened or not with 50%–50% probability in the second stage. Simultaneously with each cleaning facility a number of depots are opened, such that facilities opened in the first stage are able to handle the low amount and facilities opened in both stages can handle the high amount.

Cleaning facilities	Three-stage approach		Stochastic	Stochastic
	first stage	second stage	low supply	high supply
Almelo		•		•
Utrecht			•	
Alblasserdam	•			
Oosterhout				•

Table 10: Three-stage approach solution and two-stage solutions

The stochastic solution given by the three-stage model is presented in Table 10, where also the stochastic solutions from low and high supply cases are included for comparison (cleaning facilities only). Geographical distribution of the three-stage solution is depicted in Figure 6. The cleaning facility opened in the first stage (Alblasserdam) makes a clear trade-off between the cleaning facility in low supply case (Utrecht) and the cleaning facility in the high supply case which works mostly at full capacity (Osterhout). The treatment facility opened in the second stage is the same as the one working mostly at one third of its capacity in high supply case (Almelo). The regional depots opened in the first stage may be regarded as divided in two groups: one group with higher capacities concentrate in the central part of the west region and other group with smaller capacities is spread along the eastern border. The regional depots opened in the second stage are mainly spread in the center and the central-east region. This configurations reflect ones again a generally unbalanced distribution of facilities in the sand problem.

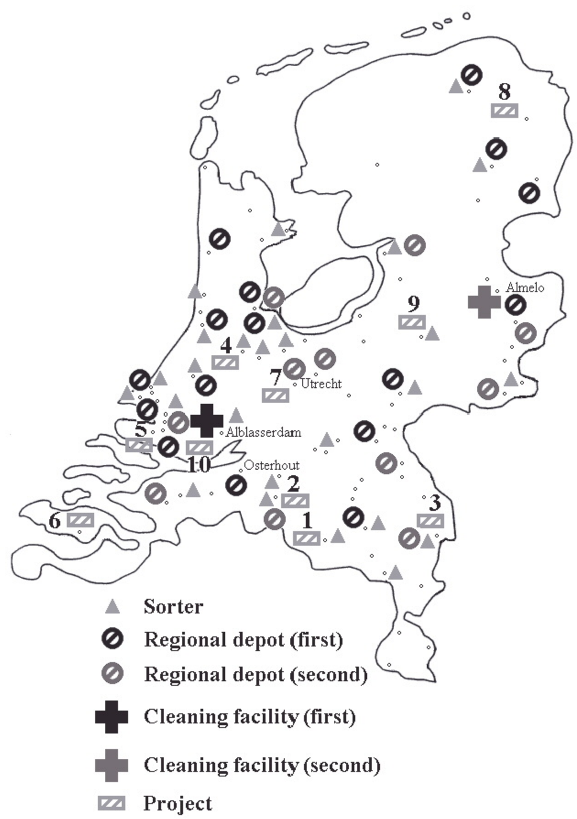

Figure 6: Stochastic solution in three-stage approach

If afterwards the supply turned out to be low, the investments made in the first stage would on one hand yield an expected (with respect to demand) revenue of 4.372.300 Dfl, that is only 51.110 Dfl less than expected revenue based on low supply assumption, but on the other hand would avoid an expected loss of 1.672.600 Dfl that occurred if excessively investing directly based on the high supply assumption (all figures are on a yearly basis). In case afterwards the supply turned out to be high, the investments made in both stages would on one hand generate an expected revenue of 6.273.100 Dfl, that is only 31.260 Dfl less than expected revenue based on direct high supply assumption, but on the other hand would avoid a situation with highly insufficient capacities based on low supply assumption (the avoided loss could be assessed in such a case by means of penalties).

5 Conclusions

Generally the solutions obtained by stochastic programming approach for the sand case perform well based on interpreting the results through the initial premises of the problem: the recycling system is mainly driven by legislation, the cleaning options are restricted through available technology to a single capacity process and transportation can also be done by water at low costs. The first means in particular that the whole amount of available material must be processed rather than collecting economically attractive amounts. Therefore the *number* of newly opened facilities is mostly decided by the amount and the quality of the incoming flows. On the other hand the *location* of the new facilities is twofold influenced by sources and by demand points in a rather non-intuitive way. Moreover, since these factors include uncertainties, finding a configuration which is likely to be more robust on a long run becomes the central investigated issue. At high material volumes the network layout is more flexible with respect to demand location and the improvement in terms of expected revenues based on the stochastic solution seems rather small. However an essential aspect which must be emphasized is the *capacitated* feature of the problem. In particular we deal with a treatment capacity which appears rather restrictive, especially relative to the high supply figures, leading in this case to investments in considerable unused cleaning capacity. To this end it is worth mentioning that future development of new technology enabling treatment at several capacities (and possibly adjustment of the working levels resulted here) is highly desirable. At low material volumes the network configuration seems more dependent on demand location and a balanced solution as generated here is particularly suitable. When more accurate information on material volumes is assumed to be stepwise revealed, splitting decisions over time is proposed as an effective long run strategy. Especially in this case the solution approach needs to be improved to allow for solving larger models in reasonable running times. We have to remark that even in the cases where the explicit improvement in the objective function is not spectacular, the stochastic approaches have the ability to generate *qualitative different* solutions, which are particularly insightful.

Given the representative character of this case, the trends discussed are likely to apply more generally to product recovery networks designed under similar conditions. Moreover the effectiveness of the type of approaches presented may be more striking under some circumstances, such as a larger area over which the potential sites are spread out, several processing capacities or higher transportation costs.

The results presented in this paper point out that given the existing computational power and using an adequate modeling it is nowadays possible to apply stochastic programming techniques to practical situations of logistic networks design. Important issues that this study uncovers are designing and implementing more efficient algorithms for the arisen stochastic models, which may lead to a significant improvement of the running times. These issues may form the subject of future research.

References

Ammons, J.C., Realff, M.J., Newton, D., 1999. Carpet recycling: Determining the reverse production system design, The Journal of Polymer-Plastics Technology and Engineering 38 (3) 547–567.

Barros, A.I., Dekker, R., Scholten, V., 1998. A two-level network for recycling sand: a case study, European Journal of Operations Research 110 (2) 199–214.

Birge, J.R., Louveaux, F., 1997. Introduction to stochastic programming, Springer Series in Operations Research, Springer Verlag, New York.

Brooke, A., Kendrick, D., Meeraus, D., 1988. GAMS, a user's guide, The Scientific Press, Redwood City, CA.

Drezner, Z., 1995. Facility Location: A Survey of Applications and Methods, Springer Series in Operations Research, Springer Verlag, New York.

Flapper, S.D.P., Kip, B.J., Louwers, D., Peters, G., Souren, F., 1997. A location-allocation model for re-use of carpet materials, Technical Report TUE/TM/BDK/LBS/97-11, Eindhoven University of Technology, Eindhoven, The Netherlands.

Fleischmann, M., Krikke, H.R., Dekker, R., Flapper, S.D.P., 2000. A classification of logistics networks for product recovery, Omega 28 (6) 653–666.

Geoffrion, A.M., Powers, R.F., 1995. Twenty years of strategic distribution system design: An evolutionary perspective, Interfaces 25 (5) 105–127.

ILOG, 1999. ILOG CPLEX 6.5 User's manual.

Kall, P., Wallace, S.W., 1994. Stochastic Programming, John Willey and Sons, Chichester, UK.

Laporte, G., Louveaux, F., van Hamme, L., 1994. Exact solution to a location problem with stochastic demands, Transportation Science 28 (2) 95–103.

Morgan, M.G., Henrion, M., 1990. Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis, Cambridge University Press, New York.

Newton, D.J., Realff, M.J., Ammons, J.C., 1999. Carpet recycling: The value of cooperation and a robust approach to determining the reverse prodution system design. In: Flapper, S.D.P., de Ron, A.J., (Eds.), Proceedings of the Second International Working Seminar on Re-use, Eindhoven, The Netherlands, pp. 207–216.

Scholten, V., 1995. An economical model for the location of depots for sieved sand, Master's thesis, Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands (in Dutch).

Appendix

Mixed integer linear programming model for sand network design

We include here the underlying deterministic MILP model for the design of the sand network. The formulation below slightly differs from the initial form in Barros et al. (1998). However the differences come from technical reasons only and do not change the problem context. Therefore we refer to Barros et al. (1998) for further details on modeling issues.

For the mathematical description of the model the following notation is used. First, the indices used to distinguish the elements of the model are defined in Table 11. Table 12 describes the notation used for the parameters of the model. Similarly, the notation for the decision variables is given in Table 13.

Index	Description	Range
i	sources of sieved sand	$1, \ldots, m$
j	potential regional depots sites	$1, \ldots, n$
k	potential cleaning facilities sites	$1, \ldots, p$
l	projects	$1, \ldots, q$
s	type of sand: clean, half-clean, polluted	$1, 2, 3$

Table 11: Indices used in the model

Parameter	Description
o_i	supply of sand at the source i (in tons)
d_{ls}	demand of sand type s = 1,2 of project l (in tons)
t_s	percentage from supply of sand type s
fee	fee charged for sand entering the network (in Dfl per ton)
pr_s	price of sand type s sold to projects (in Dfl per ton)
f_j	fixed cost of opening regional depot j (in Dfl)
g_k	fixed cost of opening cleaning facility k (in Dfl)
a_{ij}	sum of transportation costs between source i and regional depot j and processing costs at this depot (in Dfl per ton)
b_{jk}	sum of transportation costs between regional depot j and cleaning facility k and processing costs at this treatment facility (in Dfl per ton)
c_{kl}	transportation costs between cleaning facility k and project l (in Dfl per ton)
e_{jls}	transportation costs of sand type s =1,2 between regional depot j and project l (in Dfl per ton)
H_j	maximal storage capacity of regional depot j (in tons)
R_k	maximal processing capacity of cleaning facility k (in tons)

Table 12: Parameters of the model

Variable	Description	Type
y_j	$= \begin{cases} 1 & \text{if regional depot } j \text{ is opened} \\ 0 & \text{otherwise} \end{cases}$	integer $\{0,1\}$
z_k	$= \begin{cases} 1 & \text{if cleaning facility } k \text{ is opened} \\ 0 & \text{otherwise} \end{cases}$	integer $\{0,1\}$
u_{ij}	amount of sieved sand shipped from sorting facility i to regional depot j (in tons)	continuous ≥ 0
v_{jk}	amount of polluted sand shipped from regional depot j to treatment facility k (in tons)	continuous ≥ 0
w_{kl}	amount of clean sand shipped from cleaning facility k to project l (in tons)	continuous ≥ 0
x_{jls}	amount of sand of type s = 1,2 shipped from regional depot j to project l (in tons)	continuous ≥ 0

Table 13: Decision variables

Define Φ as the set of all possible scenarios and $\phi \in \Phi$ as a particular scenario. Some parameters may have different values for different scenarios. Therefore using the above notation the deterministic mixed integer linear programming model for a particular scenario ϕ states as

$$\begin{aligned}
\max \quad & \sum_{i=1}^{m}\sum_{j=1}^{n} fee(\phi) u_{ij} + \sum_{j=1}^{n}\sum_{l=1}^{q}\sum_{s=1}^{2} pr_s x_{jls} \\
& - \sum_{j=1}^{n} f_j y_j - \sum_{k=1}^{p} g_k z_k - \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} u_{ij} - \sum_{j=1}^{n}\sum_{k=1}^{p} b_{jk} v_{jk} - \sum_{k=1}^{p}\sum_{l=1}^{q} c_{kl} w_{kl} - \sum_{j=1}^{n}\sum_{l=1}^{q}\sum_{s=1}^{2} e_{jls} x_{jls} \qquad (SP)
\end{aligned}$$

$$\text{s.t.} \quad \sum_{j=1}^{n} u_{ij} = o_i(\phi)\,, \quad i = 1, \ldots, m \tag{1}$$

$$\sum_{j=1}^{n} x_{jl1} + \sum_{k=1}^{p} w_{kl} \leq d_{l1}(\phi)\,, \quad l = 1, \ldots, q \tag{2}$$

$$\sum_{j=1}^{n} x_{jl2} \leq d_{l2}(\phi)\,, \quad l = 1, \ldots, q \tag{3}$$

$$\sum_{k=1}^{p} v_{jk} = t_3(\phi) \sum_{i=1}^{m} u_{ij}\,, \quad j = 1, \ldots, n \tag{4}$$

$$\sum_{l=1}^{q} x_{jls} \leq t_s(\phi) \sum_{i=1}^{m} u_{ij}\,, \quad j = 1, \ldots, n\,,\ s = 1, 2 \tag{5}$$

$$\sum_{l=1}^{q} w_{kl} \leq \sum_{j=1}^{n} v_{jk}\,, \quad k = 1, \ldots, p \tag{6}$$

$$(1 - t_3(\phi)) \sum_{i=1}^{m} u_{ij} \leq H_j y_j\,, \quad j = 1, \ldots, n \tag{7}$$

$$\sum_{j=1}^{n} v_{jk} \leq R_k z_k\,, \quad k = 1, \ldots, p \tag{8}$$

$$y_j, z_k \in \{0,1\}\,, \quad j = 1, \ldots, n\,,\ k = 1, \ldots, p \tag{9}$$

$$\begin{aligned} u_{ij}, v_{jk}, w_{kl}, x_{jls} \geq 0, \quad & i = 1, \ldots, m\,,\ j = 1, \ldots, n\,, \\ & k = 1, \ldots, p\,,\ l = 1, \ldots, q\,,\ s = 1, 2 \end{aligned} \tag{10}$$

So this formulation also indicates which parameters are determined to change in the scenarios. In order to strengthen the formulation the following valid inequalities are included:

$$u_{ij} \leq \min \left\{ o_i(\phi)\,,\, \frac{H_j}{1 - t_3(\phi)} \right\} y_j\,, \quad i = 1, ..., m\,, \quad j = 1, ..., n \tag{11}$$

$$v_{jk} \leq \min \left\{ t_3(\phi) \sum_{i=1}^{m} o_i(\phi)\,,\, R_k \right\} z_k\,, \quad j = 1, ..., n\,,\; k = 1, ..., p \tag{12}$$

$$w_{kl} \leq \min \{ d_{l1}(\phi)\,,\, R_k \}\, z_k\,, \quad k = 1, ..., p\,,\; l = 1, ..., q \tag{13}$$

$$x_{jls} \leq \min \left\{ d_{ls}(\phi)\,,\, \frac{t_s(\phi)}{1 - t_3(\phi)} H_j \right\} y_j\,, \quad j = 1, ..., n\,,\; l = 1, ..., q\,,\; s = 1, 2 \tag{14}$$