

Design and planning of the bioethanol supply chain via simulation-based optimization: The case of Argentina

Guillermo A. Durand¹, Fernando D. Mele², Gonzalo Guillén-Gosálbez³, Alberto Bandoni^{1,*}

1 PLAPIQUI, Camino La Carrindanga km7, Bahía Blanca, Argentina
{gdurand, abandoni}@plapiqui.edu.ar

2 Universidad Nacional de Tucumán, SM de Tucumán, Argentina
fmele@herrera.unt.edu.ar

3 Universitat Rovira i Virgili, Tarragona, Spain
gonzalo.guillen@urv.cat

Abstract. This paper addresses the strategic planning of bioethanol supply chains (SC) under uncertainty in the demand. The design problem is setting as a reformulation of a multi-scenario mixed-integer linear programming (MILP) problem. It consists in the adoption of a simulation-based optimization scheme to accommodate the variables belonging to different management levels. The strategy supports decisions on the capacity expansions of the production and storage facilities of the network over time along with the associated planning decisions (i.e., production rates, sales, etc.). The capability of the approach is demonstrated through a case study based on the Argentinean sugarcane industry. Results include the investment strategy for the optimal SC configuration.

Keywords: Supply Chain Management, uncertainty, sugar cane industry

1 Introduction

Following a global trend, Argentina published law 26 093 in 2006, which provides the framework for investment, production, and marketing of biofuels. This law establishes a minimum content of biofuel in gasoline and diesel, with the purpose of diversifying the supply of energy and promoting the development of rural areas. Most of the ethanol in Argentina is currently produced by 15 sugar mills located in the north-west of the country, using sugar molasses as main feedstock.

To meet the official requirements, Argentina needs to expand its sugar cane industry. The country has abundant natural resources, a very efficient agricultural production sector, and good processing and export infrastructures [1]. However, there are some drawbacks such as land competition with food, environmental impact of transportation, and the generation of large amounts of wastewater. Moreover, the rapid expansion of ethanol consumption has affected the international market of sugar, the coproduct of ethanol. In particular, one of the key issues that still remains open is how to develop a more approaches to design networks and facilities capable of meeting the growing demand of sugar and ethanol, such as the one proposed here.

Decisions involved in the design, planning, production, and delivery of products to final customers are the focus of supply chain management (SCM). In the last few decades, the process systems engineering community (PSE) has developed tools to facilitate decision-making in this area. Among the available methods, those based on mathematical programming such as linear programming (LP) and mixed-integer linear programming (MILP) are the prevalent approaches.

Particularly, few works have focused on the optimization of bioethanol/sugar SCs. Kawamura *et al.* [2] introduced LP models to find the optimal SC network by minimizing transportation and storage costs in Brazil. López Milán *et al.* [3] and Ioannou [4] did a similar work for the Cuban and Greek cases. Some other references can be found in the work by Mele *et al.* [5].

The studies mentioned above assume that all model parameters are perfectly known. But, in practice, some of them show certain degree of variability. Various approaches have been proposed to tackle optimization models with uncertainty [6], being the two-stage stochastic programming the common approach [7]. It involves two types of decisions: first-stage decisions that must be made before the realization of the uncertain parameters, and second-stage decisions that are taken once the uncertainty is unveiled. The goal is to choose the first-stage variables in a way that the expected value of the objective function is optimized for all the scenarios. Alternative approaches are robust optimization [8], and fuzzy programming [9].

Regarding optimization of biofuel infrastructures, there are just a few works in the literature that have accounted for uncertainty. As relevant examples, it can be mentioned the works in [10-12], all of them using variations of the two-stage approach.

This paper presents an approach that is an alternative to treat the uncertainty in SCM. The proposed approach relies on a simulation-based optimization (SbO) strategy that uses an MILP model of the SC. This model is coupled with a metaheuristic optimization algorithm designed to improve the network configuration.

In SbO, the simulation model can be understood as a function whose explicit form is unknown and which converts input parameters to performance measures [13]. The choice of metaheuristics to optimize a system represented by a simulation model is because they are designed to seek global optimality in a robust way, even though they do not have a sound theoretical basis. Evolutionary algorithms, among them genetic algorithms (GA), are a particularly important subset of metaheuristic methods, whose main advantage is that they are capable of exploring a larger area of the solution space with a smaller number of objective function evaluations. Because, in the context of SbO, evaluating the objective function entails running the simulation model, being able to find high-quality solutions early on in the search is of critical importance. The authors have already presented contributions within the field of SbO [14-15].

2 Problem statement

The SC considered in this work integrates facilities for the combined production of ethanol and sugar, in which final products are stored in warehouses before being delivered to the final markets (Fig. 1).

The problem addressed in this article can be stated as follows. Given are a set of potential locations for the SC facilities, the capacity limitations associated with these technologies, the demand and prices of final products and raw materials and the investment and operating cost of the network. The goal of the study is to determine the configuration of the SC along with the associated planning decisions that maximize its economic performance. The demand is assumed to be uncertain, and it is described through a set of scenarios with a given probability of occurrence. Failure to properly account for product demand fluctuations may result in either unsatisfied customer demand or excess of products. The first scenario leads to a loss of potential revenues and market share, whereas the second one generates large inventory costs. Modelling has been done according to the features of the Argentinean sugarcane industry.

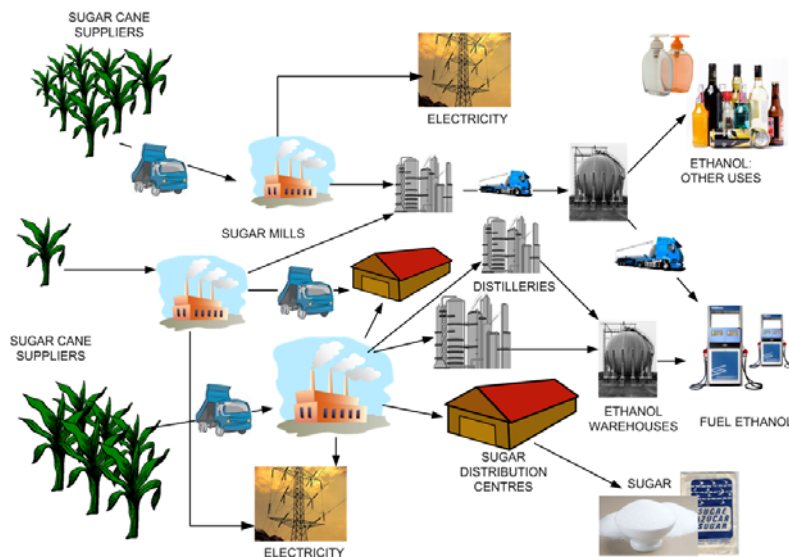


Fig. 1. Supply chain network.

Production. Cane juice is extracted from sugarcane by milling to produce white and raw sugar. There are two available technologies to carry out this process. One of them generates molasses (T1) as a by-product, whereas the other one produces a secondary honey (T2) as a by-product. Anhydrous ethanol is produced from different raw materials (via fermentation and dehydration): molasses (T3), honey (T4), and cane juice (T5). Thus, the model considers five different technologies, two for sugar production and three for ethanol production (distilleries). Nine materials are considered among raw materials and products: sugarcane, ethanol, molasses, honey, white and raw sugar, and vinasses type 1, 2 and 3. Each plant incurs fixed capital and operating cost, and can be expanded in capacity over time in order to follow some demand pattern.

Storage. Two different types of storage facilities are included: warehouses for liquid products (S1), and for solid materials (S2). For each one, fixed capital and unit storage costs have been considered, along with lower and upper limits on capacity expansions. Similarly, as with the plants, the storage capacity can be expanded.

Transportation. Transportation units deliver the final products to the customers, supply the production plants, and dispose wastes. In the model materials can be transported by three types of trucks: heavy trucks for sugar cane (TR1), medium trucks for sugar (TR2), and tank trucks for liquid products (TR3). Each transportation mode has fixed capital and unit transportation costs, and lower and upper capacity boundaries.

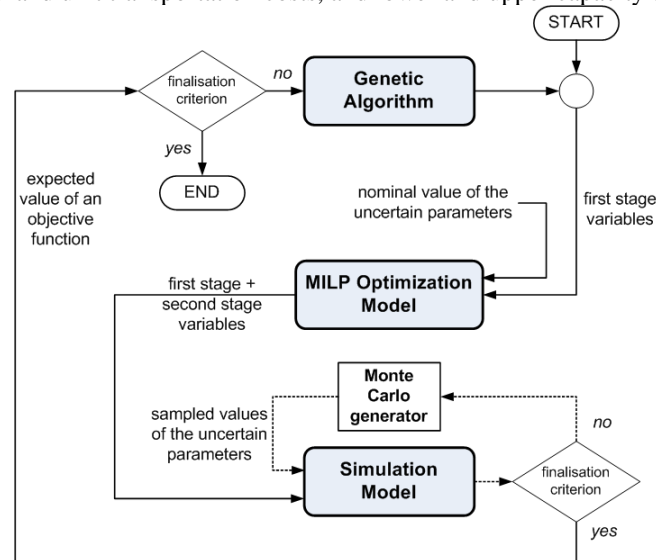


Fig. 2. Simulation-based optimization strategy.

3 Simulation-based optimization strategy

The proposed algorithm can be followed from Fig. 2. It involves an outer loop which is a GA and an inner loop which is a Monte Carlo simulation over a MILP deterministic SC model. Variables have been divided into first- and second-stage variables.

First-stage variables stands for SC design decisions, namely the number of production, storage and transportation units, and their initial capacities and capacity expansions over the time horizon. The reason for this is that we assume that they are taken at the beginning of the time horizon, before the demand is known. Second-stage variables include the amount of products to be produced and stored, the flows of materials transported among the SC entities and the product sales. These last variables include a subscript e that denotes the particular scenario realization for which they are defined.

At the beginning of the algorithm the values for the stage variables are given and the uncertain parameters, demand, are fixed in the nominal value. A MILP optimization model is executed to evaluate the second variables. The objective function for this model is the net present value (NPV). Then, the calculated variables are considered to be fixed and confronted to variations in the demand; this is done by running n times the same model with different values of the uncertain parameters sampled by a

Monte Carlo generator. Then, the expected value of a new objective function is calculated and this value is fed to a GA implementation to propose new values for the first-stage variables. This new objective function is the customer satisfaction *CSat*. The outer loop, the GA, keeps on working until a finalisation criterion is satisfied, in this case, when a given maximum number of generations is reached. The equations comprising the *inner* MILP model can be roughly classified into three main blocks: mass balance equations, capacity constraints and objective function equations.

3.1 Mass balance constraints

The overall mass balance for each region g is enforced via Eq. (1). For every material i and scenario e , the initial inventory ($ST_{is,g,t-1,e}$) plus the amount produced (PT_{igte}), the amount of raw materials purchased (PU_{igte}) and the input flow from other facilities ($Q_{ilg'gte}$) must equal the final inventory (ST_{isgte}) plus the amount delivered to the customers (DTS_{igte}) plus the output flow to other facilities ($Q_{ilgg'te}$) and the wastes (W_{igte}).

$$\begin{aligned} & \sum_{s \in IS(i,s)} ST_{isg,t-1,e} + PT_{igte} + PU_{igte} + \sum_{l \in IL(i,l)} \sum_{g' \neq g} Q_{ilg'gte} = \\ & = \sum_{s \in IS(i,s)} ST_{isgte} + DTS_{igte} + \sum_{l \in IL(i,l)} \sum_{g' \neq g} Q_{ilgg'te} + W_{igte} \quad \forall i, g, t, e \end{aligned} \quad (1)$$

The total production rate of a material is determined from the production rates associated with each technology p installed in that region (PE_{ipgte}):

$$PT_{igte} = \sum_p PE_{ipgte} \quad \forall i, g, t, e \quad (2)$$

The production of by-products and consumption of raw materials are obtained from the material balance coefficient ρ_{pi} , and the production rate of the main product:

$$PE_{ippgte} = \rho_{pi} PE_{ipgte} \quad \forall i, p, g, t, e \quad \forall i' \in IM(i, p) \quad (3)$$

The amount of product i sent to the final markets located in region g in scenario e should be less than or equal to the corresponding demand in that region (SD_{igte}):

$$DTS_{igte} \leq SD_{igte} \quad \forall i, g, t, e \quad (4)$$

3.2 Capacity constraints

For each scenario e and time interval t , the purchases of sugarcane are limited by the capacity of the existing sugarcane plantation in region g :

$$PU_{igte} \leq CapCrop_{gt} \quad i = \text{sugarcane}, \quad \forall g, t, e \quad (5)$$

The total inventory (ST_{isgt}) of product i is forced by the storage capacity ($SCap_{sgt}$):

$$\sum_{i \in IS(i,s)} ST_{isgte} \leq SCap_{sgr} \quad \forall s, g, t, e \quad (6)$$

The production rate of each technology p is bounded according to eq. (7), where τ is the minimum desired percentage of the available technology:

$$\tau PCap_{pgt} \leq PE_{ipgt} \leq PCap_{pgt} \quad \forall i, p, g, t \quad (7)$$

The capacity of technology p equals the sum of the existing capacity at the end of the previous period, and the capacity expansion carried out in this period ($PCapE_{pgt}$):

$$PCap_{pgt} = PCap_{pgt-1} + PCapE_{pgt} \quad \forall p, g, t \quad (8)$$

Eq. (9) limits the capacity expansion between upper and lower bounds that are calculated from the number of plants installed in that region (NP_{pgt}) and the minimum and maximum capacities related to each technology p (\underline{PCap}_p and \overline{PCap}_p).

$$\underline{PCap}_p NP_{pgt} \leq PCapE_{pgt} \leq \overline{PCap}_p NP_{pgt} \quad \forall p, g, t \quad (9)$$

Similar constraints are defined for the storage and transportation capacity.

3.3 Objective function

The model exhibits different performance in each scenario, which is measured through the net present value (NPV_e) and the customer satisfaction ($CSat_e$). NPV_e is the objective function of the inner loop and the expected value of the $CSat_e$ is the objective function of the outer loop.

$$E[CSat] = \sum_e pr_e CSat_e \quad (10)$$

where pr_e is the probability of scenario e . NPV_e is determined from the cash flows (CF_{te}) generated in each time interval t and the interest rate ir :

$$NPV = \sum_t \frac{CF_t}{(1+ir)^{t-1}} \quad (11)$$

Details on how to calculate these terms, which involve fixed and operating costs calculation, can be found in [12].

4 Results and discussion

The SbO strategy has been tested on the sugarcane industry of Argentina. The problem considers 12 regions each one with an associated demand of sugar and ethanol. The entire set of data for this cases study can be found in [12]. It includes the mean

demand values associated to each region; the prices for white sugar, raw sugar and ethanol; distances between regions; sugarcane crop capacity; minimum and maximum capacities; fixed and variable investment cost coefficients, and unit cost. The length of the planning horizon is equal to 3 years.

Table 1. Production capacity installation (y. 1) and expansion (y. 2 and y. 3). Optimal solution*

	Córdoba	Mesopotamia	Bs. Aires	Cuyo	North	NW
Year 1		0.100 (T3)				1.634 (T1)
						0.199 (T3)
						0.100 (T4)
						2.793 (T5)
Year 2		0.300 (T2)				1.934 (T1)
		0.100 (T3)				0.199 (T3)
						0.100 (T4)
						2.793 (T5)
Year 3		0.300 (T2)				1.934 (T1)
		0.100 (T3)				0.199 (T3)
						0.100 (T4)
						2.793 (T5)
	Tucumán	Santa Fe	La Pampa	Santiago	West	Patagonia
Year 1	3.500 (T2)	0.360 (T2)	0.300 (T1)	3.500 (T2)		
	0.784 (T4)	0.100 (T3)	0.300 (T2)	0.784 (T4)		
		0.100 (T4)	0.100 (T3)			
		0.100 (T5)				
Year 2	3.500 (T2)	0.360 (T2)	0.300 (T1)	3.500 (T2)		
	0.784 (T4)	0.200 (T3)	0.300 (T2)	0.784 (T4)		
		0.100 (T4)	0.100 (T3)			
		0.100 (T5)	0.100 (T4)			
Year 3	3.500 (T2)	0.360 (T2)	0.300 (T1)	3.500 (T2)		
	0.784 (T4)	0.200 (T3)	0.600 (T2)	0.784 (T4)		
		0.100 (T4)	0.100 (T3)			
		0.100 (T5)	0.100 (T4)			

* Quantities in 10^5 t/year. Type of technology installed or expanded in parentheses

The MILP deterministic model was written in GAMS® and solved with CPLEX 11.2 on an Intel Pentium D 945 desktop PC with 1 GB of RAM by using the “rolling horizon” strategy introduced in a previous work [16]. The outer loop GA was implemented in Matlab® R2007b. The stochastic simulations of the inner loop considered a factor $N = 1.30\%$ for the estimated sugar demand, and $N = 1.5\%$ for the ethanol one. The Monte Carlo generator was set to provide $n = 100$ samples.

Table 1 shows the evolution of the SC network along the horizon in terms of production. It can be seen that while the optimal solution covers all the regions with delivery/storage centres, it tries to concentrate production plants near the plantation zones. The final geographical distribution of the SC nodes can be seen in Fig. 3.

Table 2. SbO statistics and results

Population	100
Generations	40
Objective function (best individual)	82.17%
NPV (best individual)	\$227.7x10 ⁶
Best individual found at generation	38
Tried combinations	100 x 41 = 4100
Unique combinations	3975 (1627 not valid)
Average CPU time per MILP solving	0.846 seconds
Total CPU time	61 min 1.47 seconds

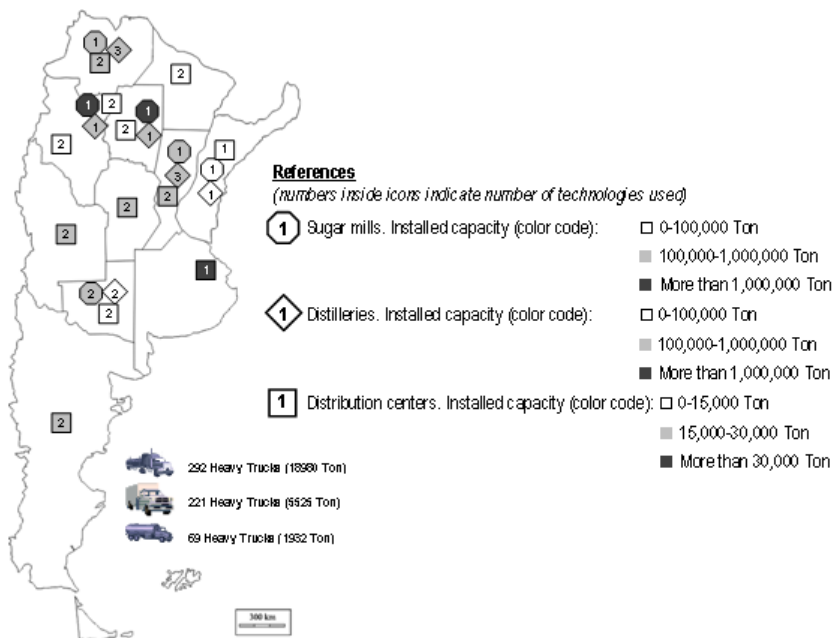


Fig. 3. SC configuration of the minimum GWP100 solution.

Table 2 presents some statistics of a SbO run. It took just over an hour to try 4100 combinations of outer loop variables, and each instance of the inner loop MILP was solved in about 0.85 seconds. “Unique combinations” refers to the quantity of combinations that were not a duplicate of another tried combination. If duplication happened, the combination was not sent to the inner loop but given the same objective value of its duplicate, thus improving the CPU time. Of 3975 unique combinations, 1627 were deemed invalid because the inner loop did not find a positive NPV value. Those combinations were given an outer loop objective value of 0%, in order to make the GA to discard them. The best individual was found in the latter generations (38th), and further SbO runs with increasing number of generations showed that the objective

function value did not improve significantly. The objective function value of the best individual was over 80%. As comparison, when the problem was solved as a complete deterministic MILP (the MILP handling all the decision variables), the resulting SC only reached a 65.98% of demand satisfaction on the Monte Carlo generator.

Fig. 4.a shows the objective function values of the best and worst individual and the average for each generation. Fig. 4.b shows the value of the *NPV* of the best individual of each generation. The number of not valid combination over the generations is given in Fig. 4.c. Although it was expected that better demand satisfaction comes with more installed production plants, storage facilities and transport units, thus increasing cost and decreasing the *NPV*, the SbO framework manages to improve both the demand satisfaction and the *NPV* along the GA generations.

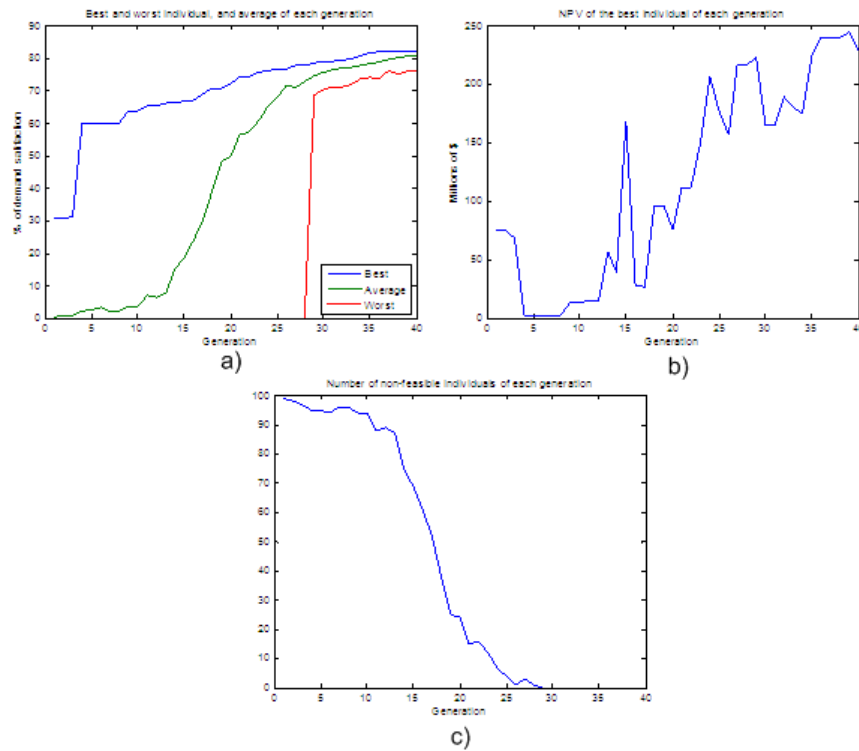


Fig. 4. Evolution of the GA outer loop.

5 Conclusions

A SbO strategy has been implemented to solve the problem of optimal design of the sugar/ethanol SC in Argentina under parametric uncertainty. A model taken from literature was used to implement a two level optimization framework that combines MILP solving with Monte Carlo simulation and GA. The proposed framework han-

dles around 70 uncertain parameters (the products demands for each region and time period) and geographically distributes production/ storage/ distribution nodes in Argentina, considering different technologies. The model decides the production/storage capacities of the nodes, the quantity of transport units and the period in which each installation and/or expansion should be made. The proposed SbO strategy combines two objective functions, *CSat* and *NPV*, thus allowing increasing both although it was expected them to be opposite.

Acknowledgements. This work was partially supported by “Consejo Nacional de Investigaciones Científicas y Técnicas” (CONICET), “Universidad Nacional del Sur” and “Universidad Nacional de Tucumán”, Argentina.

References

1. Joseph, K. GAIN Report. Argentina. Biofuels Annual. USDA For. Agric. Service (2009)
2. Kawamura, M.S., Ronconi, D.P.Y., Yoshizaki, H.T.Y. Optimizing transportation and storage of final products in the sugar and ethanol industry: A case study. *Int. Trans. Operat. Res.* 13 (5), 425–439 (2006)
3. López Milán, E., Miquel Fernández, S., Miquel Pla Aragonés, L. Sugar cane transportation in Cuba, a case study. *Eur. J. Operat. Res.* 174 (1), 374–386 (2006)
4. Ioannou, G. Streamlining the SC of the Hellenic sugar industry. *J. Food Eng.* 70 (3), 323–332 (2005)
5. Mele, F.D., Kostin, A.M., Guillén, G., Jiménez, L.: Multiobjective model for more sustainable fuel SCs. *Ind. Eng. Chem. Res.* 50, 4939–4958 (2011)
6. Sahinidis, N. Optimization under uncertainty: state-of-the-art and opportunities. *Comp. Chem. Eng.* 28 (6–7), 971–983 (2004)
7. Liu, M., Sahinidis, N. Optimization in process planning under uncertainty. *Ind. Eng. Chem. Res.* 35 (11), 4154–4165 (1996)
8. Li, P., Arellano-García, H., Wozny, G. Chance constrained programming approach to process optimization under uncertainty. *Comp. Chem. Eng.* 32 (1–2), 25–45 (2008)
9. Zimmermann, H.-J. Fuzzy set theory and its application, Kluwer AP: Boston (1991)
10. Dal-Mas, M., Giarola, S., Zamboni, A., Bezzo, F. Strategic design and investment capacity plan. of the ethanol SC under price uncertainty. *Biom. Bioen.* 35 (5), 2059–2071 (2011)
11. Kim, J., Realf, M., Lee, J. Optimal design and global sensitivity analysis of biomass SC networks for biofuels under uncertainty. *Comp. Chem. Eng.* 35, 1738–1751 (2011)
12. Kostin, A.M., Guillén-Gosálbez, G., Mele, F.D., Bagajewicz, M.J., Jiménez, L. Design and planning of infrastructures for bioethanol and sugar production under demand uncertainty. *Chem. Eng. Res. Des.* 90, 359–376 (2012)
13. Law, A.M., Kelton, W.D. *Simulation Modeling & Analysis*, McGraw-Hill: NY (2000)
14. Mele, F.D., Guillén, G., Espuña, A., Puigjaner, L.: A simulation-based optim. framework for parameter optim. of SC networks. *Ind. Eng. Chem. Res.* 45, 3133–3148 (2006)
15. Durand, G.A., Mele, F.D., Bandoni, J.A. Determination of storage tanks location for optimal short-term scheduling in multipurpose/multiproduct batch-continuous plants under uncertainties. *Ann. Oper. Res.* DOI: 10.1007/s10479-011-0970-8 (2011)
16. Kostin, A., Guillén-Gosálbez, G., Mele, F., Bagajewicz, M., Jimenez, L. A novel rolling horizon strategy for the strategic planning of SCs. *Comp. Chem. Eng.* 35, 2540–2563 (2011)