Ann Oper Res (2012) 199:225–247
DOI 10.1007/s10479-011-0970-8

Determination of storage tanks location for optimal short-term scheduling in multipurpose/multiproduct batch-continuous plants under uncertainties

Guillermo A. Durand · Fernando D. Mele · J. Alberto Bandoni

Published online: 22 September 2011

Abstract A multipurpose/multiproduct plant has to deal with many combinations of tasks sequences and operation rates that lead to accumulation problems. These problems can be handled using storage tanks, but usually their location within the flowsheet is predetermined and not subject to optimization, missing the opportunity to better satisfy the customers. In this work we will determinate the optimal location of storage tanks for the short-term scheduling under uncertainty. A hybrid simulation-based optimization (SBO) strategy was developed and implemented to solve the problems combining stochastic and deterministic solution algorithms.

Keywords Decision-making · Uncertainty · Simulation-based optimization · Batch plant · Semi-continuous plant · Scheduling · Multipurpose/multiproduct plants

1 Introduction

The manufacturing process industry operates in a rapidly changing and uncertain environment. Uncertainties appear, for example, in customer demands for products, raw materials and products prices, lead times for the supply of raw materials, production and distribution of the end products, and in process and quality failures.

This situation involves making decisions at different levels in order to satisfy multiple goals. For example, companies must decide the tuning of the inventory control policies several times during a relative long-term period (e.g., once a month), and, more frequently (e.g.,

An outline of this work was first presented in the ALIO-INFORMS Joint International Conference 2010, Buenos Aires.

G.A. Durand (✉) · J.A. Bandoni
PLAPIQUI (CONICET-UNS), Camino La Carrindanga km 7, 8000, Bahía Blanca, Argentina
e-mail: gdurand@plapiqui.edu.ar

F.D. Mele
Dep. Process Engineering, FACET, Universidad Nacional de Tucumán, Av Independencia 1800, T4002BLR, San Miguel de Tucumán, Argentina

once a week), they may have to schedule production to satisfy customer demands. This type of problem—which is referred to as sequential decision-making under uncertainty—arises in many decision-making contexts and in a variety of situations in which one makes decisions at different stages in response to uncertainty that is resolved over time. This class of problems has attracted increasing attention from the research community, and various theoretical and computational works have been devoted to model and solve it. The resulting complexity of these problems, from theoretical and computational perspectives, is a challenge for both practitioners and researchers.

Particularly, a problem of interest is the short-term scheduling of industrial processes, which involves finding the optimal series of actions to carry out when operating the process equipment in order to meet a determinate objective (maximization of profit, minimization of makespan, etc.). In many cases this is done with deterministic optimization models with fixed values for parameters. However, this often leads to sub-optimal or infeasible solutions for the real-world cases, because many of the assumed parameter values are, in fact, uncertain. Although this is acknowledged since long time, the increase in complexity when these uncertainties are considered in scheduling problems has hindered their study to very recent years.

A common way of operation, applied in practice to try to mitigate the adverse effects of uncertainty, is the use of storage tanks. In general, the existence of storage vessels decouples processing steps and enables better utilization of resources. For example, storage before and after a bottleneck stage leads to higher production because it allows uninterrupted utilization of the most scarce resources. Even Storage for raw materials and final products can potentially lead to more flexible operation. Even in some cases, storage itself is often a necessary processing step; e.g., when products have to be held for at least some time before the next processing step begins (Sundaramoorthy and Maravelias 2008).

Nevertheless, finding the location of these storage tanks is not a trivial problem because it can influence the scheduling results. Therefore, it constitutes a bi-level problem: a design problem (tanks location) coupled with an operational one (optimal scheduling). Hence the problem can be seen as a sequential decision problem under uncertainty.

The objective of this work is to develop a technique based on a hybrid simulation-based optimization (SBO) approach, combining stochastic and deterministic algorithms, to determinate the optimal storage tanks location during the short-term scheduling in multipurpose/multiproduct batch and/or continuous plants.

The paper is organized as follows. First, the literature review on the two key points involved -decision-making under uncertainty, and SBO- is given. Next, a motivating example is used to outline the problem and present a suitable approach for solving it. The proposed methodology is then explained in the following section. The strategy is applied to two case studies and their results are discussed. Finally, some concluding remarks are provided.

2 Background

According to the concepts and tools presented in this paper, the review is divided into two interrelated sections: decision-making processes under uncertainty and simulation-based optimization.

2.1 Decision processes under uncertainty

The concepts involved in decision-making under uncertainty are closely linked to those of optimization under uncertainty. Literature on optimization under uncertainty very often divides the approaches into two categories: “wait and see” and “here and now”. In the “wait and see" approaches, one has to wait until an observation is made on the random elements, and then solve the deterministic problem. Conversely, a "here and now" problem involves optimization over some probabilistic measure of the system performance—usually the expected value. It should be noted that many realistic problems have both "here and now", and "wait and see" problems embedded in them. The trick to overcome this complication situation is to divide the decisions into these two categories and use a coupled approach (Diwekar 2002).

In this regard, many advances have been observed in the supporting theory, including algorithmic developments and computational capabilities for solving this class of problems, most of which fall into one of these two approaches: multistage stochastic programming and stochastic optimal control.

In multistage stochastic programming decisions are based on past observations and decisions before the future events occur (Birge and Louveaux 1997). A finite set of scenarios is often generated to represent the space, therefore, the stochastic program becomes a deterministic equivalent program, the size of which can easily grow out of hand for a large number of scenarios, making the direct solution approaches numerically intractable, thus requiring methods of decomposition or aggregation (Balasubramanian and Grossmann 2004).

Stochastic optimal control describes a sequential decision problem in which the decision-maker chooses an action in the state involved at any decision stage according to a decision rule or policy. Dynamic programming provides the framework for designing algorithms to compute an optimal control policy. However, for large problems, dynamic programming also suffers numerically from dimensionality. Both approaches—stochastic programming and optimal control—are essentially equivalent, but they exhibit differences in formulation and solution, with the consequent advantages and disadvantages for specific problems (Cheng et al. 2003, 2004a; Duenas and Petrovic 2000; Kuster et al. 2010).

Efficient numerical solution proposals can be achieved by combining several techniques that belong to each approach. The resulting strategy needs to be adapted to solve the specific problem, defining some approximations or heuristic-based methods. The works by Cheng et al. (2004b) and Jung et al. (2004) are relevant examples in this regard. Among these approximate approaches, an attractive alternative to address large-scale problems is SBO.

2.2 Simulation-based optimization

SBO is an attractive combined strategy that addresses the situation in which the analyst wishes to find which of many possible sets of input parameters lead to the optimal performance of a system. Thus, a simulation model can be understood as a function whose explicit form is unknown and which converts input parameters to performance measures (Law and Kelton 2000).

Typically, the optimization search process is slow and inefficient due to the presence of some noise coming from the simulation output as well as the fact that each simulation run takes significant execution time. However, SBO provides an alternative for problems in which analytical methods are inefficient. Moreover, SBO is an active area of research in the field of stochastic optimization (Gosavi 2003). In the literature on Process Systems Engineering (PSE), SBO approaches have received some attention and are currently awaiting further study. The works by the group of Reklaitis (Subramanian et al. 2001; Jung et al. 2004; and Wan et al. 2005), and Mele et al. (2006b) are well-regarded contributions to this field.

There are a variety of ways of optimizing processes represented by means of a simulation model. Among those widely studied SBO algorithms, response surface methodologies

(Nedderneijer et al. 2000; Silva and Salcedo 2009; Kim et al. 2009), stochastic approximation (Kleinman et al. 1999; Gürkan 2000; George and Farid 2008) and metamodelling or surrogate model building (Barton 1994; Wan et al. 2005; Mele et al. 2006a; Sundar Raj and Lakshminarayanan 2008) play important roles.

The response surface SBO scheme proposed in this work uses a genetic algorithm (GA) to perform the search process. The success of metaheuristics in carrying out this task is perhaps that they are designed to seek global optimality and their properties are apparently robust in practice, even though they do not yet have a solid theoretical basis (Gosavi 2003). Evolutionary algorithms, in general and GA, in particular, forms an important subset of metaheuristic methods, which have been used to optimize multimodal, discontinuous, and non-differentiable functions, 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.

3 Motivating problem

The motivating problem for this article is to determine the optimal storage tanks location into a multipurpose/multiproduct batch-continuous plant. Current approaches to this problem transform it into different optimization problems whose structures are normally complex (nonlinear, non-convex problems with mixed integer and continuous variables). Specifically, the consideration of the uncertainty and the modelling of the heuristic rules that drive the plant operations usually results in large-scale mixed-integer nonlinear programming (MINLP) problems. Solving these problems by means of mathematical programming techniques requires excessive computing time, because of their size and complexity when real scenarios are considered.

In this work, a SBO approach is proposed to address the sequential decision-making problem under uncertainty in the area of batch-continuous plant scheduling, which has the objective of overcoming the computational limitations of current solution methods. This approach considers two important components: a simulation model used to represent the real system, and a metaheuristic optimization algorithm that is designed to improve the operation of this system.

Consider the following situation to illustrate the effect of intermediate tanks in the operation of batch-continuous plants under uncertainty. In the situation of no uncertainties in the duration of tasks -tasks last as expected- intermediate materials transfers between units happen as determined by the deterministic short-term scheduling (as shown in Fig. 1a, between units j, j' and j''). However, when applying the deterministic schedule in a realistic situation, uncertainties in the length of tasks can arise. These differences with the expected lengths lead to cases were the mass in, e.g., unit j' cannot be transferred to the following unit j'', thus not allowing the release of unit j' for the processing of other tasks. Figure 1b shows that unit j cannot transfer its mass to unit j' because the last one is still waiting for unit j'' to finish, unit j will have to wait until unit j' releases its content, as shown if Fig. 1c. However, if a storage tank is located between units j' and j'' for intermediate material s', then unit j' will be idle to continue with the scheduled tasks (Fig. 1d).

Deviations produced by uncertainties will almost surely lead to worse-than-expected performances for a given objective function. As an example, consider Fig. 2. The histogram on the left (Fig. 2a) shows the results of simulation runs for the multipurpose/multiproduct

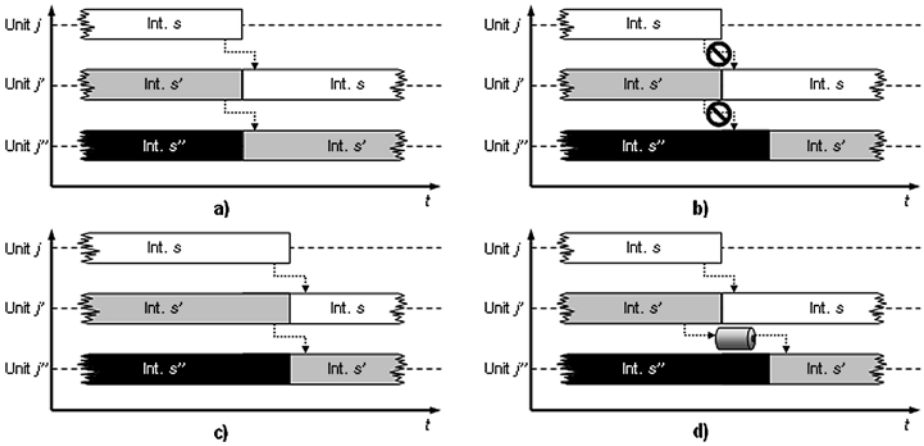

Fig. 1 Schematic representation of the intermediate material transfer in a multipurpose/multiproduct plant

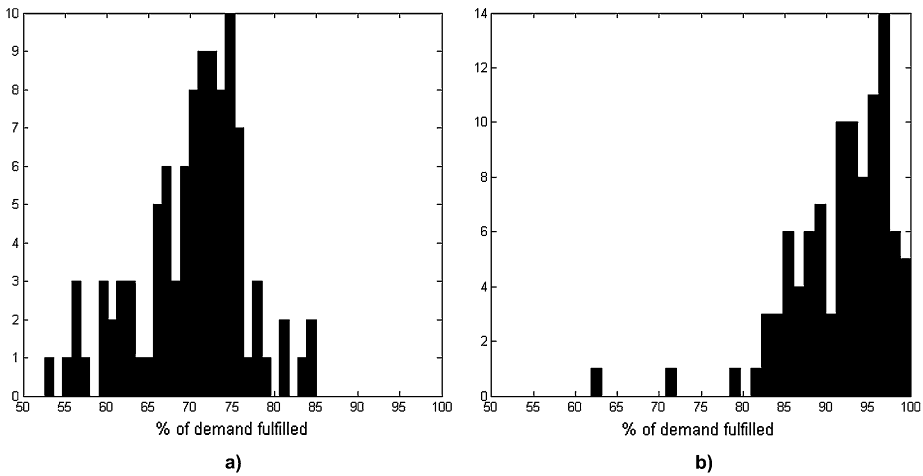

Fig. 2 Histograms from 100 simulations for a given batch plant. (**a**) No storage tanks (mean 70.25%). (**b**) Three storage tanks (mean 91.72%)

batch plant of Case Study I, in terms of total fulfilled product demands. The histogram was obtained by first doing a deterministic short-term scheduling that complies with product demands (100% of fulfilled demand), and then implementing it 100 times in a simulation model with given uncertainties in the task duration and no storage tanks. The mean value of Fig. 2a histogram is 70.25%. Figure 2b shows the same situation but having three storage tanks available in the plant flowsheet (one tank for intermediate material of product A produced by task 1 and consumed by task 2, one for intermediate B between tasks 2 and 3, and one amid tasks 3 and 4 of product C). The incidence of the storage tanks increases the mean value to 91.72% of fulfilled demand.

The goal of our contribution can be thought as a discrete-time decision process over a given time horizon, in which managers periodically make their decisions based on the available information. A partition of the set of decisions must be performed according to the frequency with which they are updated. The system evolves over time and generates values for some performance indexes. The problem then becomes one of finding the decision strategies that can optimize these performance indexes.

Particularly, if the problem of storage tanks location in batch-continuous plants is considered, the plant uses different resources to produce products s in a time horizon of T periods. The company has the option to install intermediate storage tanks at the beginning of each period $\theta \in \{1, \dots, T\}$. Within each period θ, the plant schedule is reviewed and optimized, and decisions are made at the beginning of each subperiod $\tau \in \{1, \dots, L\}$, given that the size of τ is less than the size of θ.

The uncertain parameters, such as product demand or task duration, are modelled as a discrete-time stochastic process that gives a value of the uncertain variable at some instants in the time horizon. Hereafter, one realization of the uncertainty is identified with ω.

Decisions are made in stages as information arrives and uncertainty is unveiled over time. The following sequence summarizes the events occurrence during each period:

(a) At the beginning of each period $\theta = 1, \dots, T$, having observed the information available, new information about storage tanks location η are chosen.
(b) At the beginning of each subperiod θ, and based on the current information, the simulator mimics the plant behaviour executing a given scheduling plan.
(c) The procedure is repeated from step (a) (rolling horizon) until $\theta = T + 1$.

It is important to notice that this problem can be considered as embedded within a wider decision-making problem. There might be long-lasting decisions, such as strategic decisions, to update at the beginning of each period $h \in \{1, \dots, H\}$, with the size of h being greater than the size of θ (for instance, plants or warehouses capacities), such that the scope of these decisions would include the decisions previously described.

At each stage of the decision process, as a consequence of the selected decision, the company receives a profit or incurs a cost. After choosing and implementing the decision strategy, the decision-maker receives the outcomes for the system performance. The manager's motivation is to choose a decision strategy such that these outcomes are as good as possible, that is to say, optimal. This necessitates the definition of performance measures to compare alternative decisions. In this work, the performance criterion adopted is the expected demand satisfaction, which under a given decision strategy η, can be expressed as follows:

$$F(\eta) = E[(f(\eta, \omega)] \tag{1}$$

where the expectation $E[\bullet]$ is taken with respect to all random variables in ω.

Therefore, the optimization problem can be stated mathematically as:

$$\begin{aligned}
&\max F(\eta) = E[(f(\eta, z, \omega)] \\
&\text{s.t.} \quad f(\eta, z, \omega) = \frac{1}{N} \sum_i \frac{\textit{produced quantity}}{\textit{demanded quantity}} \\
&\qquad h_s(\eta, z, \omega) = 0 \quad g_s(\eta, z, \omega) \leq 0 \\
&\qquad h_{io}(\eta, z, \omega) = 0 \\
&\qquad g_{io}(\eta, z, \omega) \leq 0 \\
&\qquad \eta^L \leq \eta \leq \eta^U, \quad z^L \leq z \leq z^U, \quad \omega^L \leq \omega \leq \omega^U
\end{aligned} \tag{2}$$

where: η accounts for the decision variables in a higher level, z, for decision variables (continuous and binary) at a lower level, and ω, for the uncertain parameters. h_s, g_s are rules-based relationships corresponding to usually non-explicit equation simulation models, and h_{io} and g_{io} correspond to variable relationships embedded in the model.

4 Proposed methodology

The problem under study is complicated enough that there is no possibility to express it explicitly or analytically; thus, an approach that considers this problem by means of an SBO approach is very well suited to address these systems. SBO proposes a hybrid framework for optimization under uncertainty, in which a "here and now" solution strategy is assumed, but incorporating "wait and see" features. The SBO approach has three valuable advantages:

- Using simulation provides a framework with the flexibility to accommodate arbitrary stochastic elements.
- It is a very flexible solution strategy, a feature that has leaded to applications with quite different purposes, e.g. portfolio selection for R&D pharmaceutical projects (Subramanian et al. 2001, 2003), industrial supply chain optimization (Jung et al. 2004; Mele et al. 2006b) or life-support system design for manned missions to Mars (Aydogan et al. 2005).
- Using simulation facilitate problem resolution of models with larger amount of details than in math programming.

These three strengths have to be balanced with two main shortcomings:

- It can take much time and computational resources to solve a given problem. Each access to the objective function requires many simulations and deterministic optimizations in order to obtain a representative value.
- Its usefulness is strongly dependent of the selection and distribution of the decision variables in the two loops.

The SBO requires decision variables to be divided into two levels: for each uncertain parameters instance and for the stochastic optimization. In the particular case of interest in our work, the location of storage tanks in a multipurpose batch-continuous plant, the classification in levels of the variables has been done by taking advantage of the decision hierarchy in manufacturing processes proposed by the International Society for Measurement and Control (1995, 2000) (ANSI/ISA-S88, ANSI/ISA-S95) (see Fig. 3). The higher level (outer loop) is related to design whereas the lower level (inner loop) is related to short-term scheduling at the operational level.

This approach proposes a hybrid strategy, combining simulation with deterministic optimization and stochastic optimization in two levels. In the inner level a deterministic optimization algorithm ignores all the random elements present in the problem and obtains a deterministic optimal solution for the problem, then the simulator implements it handling uncertainties while respecting the guidelines given by the optimization algorithm. This level is embedded in an outer level loop where a stochastic optimization technique utilizes the information from the simulation to search the decision space systematically, trying to improve the performance of the solution. The particular choice of the problem variables that will be controlled by each level, as well as the optimization and simulation algorithms, depends on the case study.

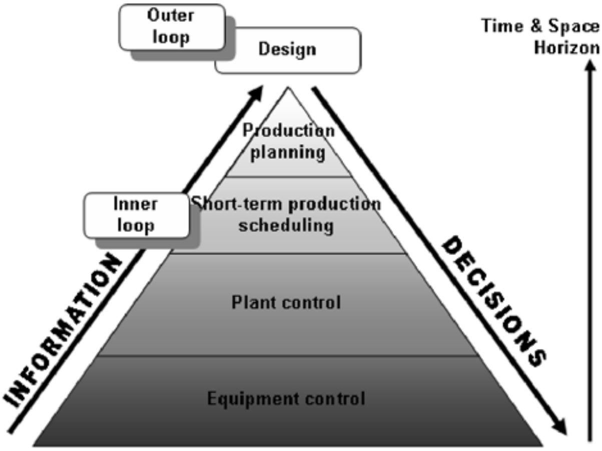

Fig. 3 Hierarchy of decision-making and information flow in manufacturing processes

In this work, a simulation model of a multipurpose batch-continuous plant is coupled with a GA to optimize certain parameters related to the operation of the plant. The logical (or heuristic) rules, which are implemented in the simulation model to respond to the arriving events, are parameterized, and the metaheuristic algorithm is then invoked to act over these parameters searching for an improvement of the objective function. The motivation for using GAs is to take advantage of the features of this technique, applying it to complex problems such as that previously described (see Sect. 3). Particularly, in the presented examples (Sects. 5 and 6), the framework permits to obtain the parameters values associated with the storage tank location, in such a way that the customers' demand satisfaction is optimized over a given time horizon. Nevertheless, the proposed approach is general enough to enable the choice of any other parameter for optimization (equipment capacity, inventory policy selection, backlogged orders management, etc.), as well as any other objective functions.

Storage tanks location is usually determined at the design stage of a project, using an expected demand as the objective to fulfil, and the proposed approach can be applied to solve not only this problem, but also tank capacities.

Although, as demands' characteristics will surely change over time, the selected locations will be no longer optimal and re-design stages will be needed. Since in multipurpose/multiproduct plants storage tanks location can be changed simply by modifying tubing and connections, without actually physically moving the vessels, the frequency of this re-designing corresponds to long- and/or medium-term planning. The time durations of these planning stages depend on the industry, but, because of this, for our technique to be applicable, it must have good computational performance, i.e., to solve the problem in a reasonable time despite the large number of uncertain parameters involved.

The proposed approach applied to the problem presented previously is summarized in Fig. 4. First, and to determine the value of the performance index, $f(\eta, \omega)$, a set of scenarios of the uncertain variables extended over the entire planning horizon ω are generated in a Monte Carlo manner. A simulation run of the batch plant is then executed, considering a specific distribution of storage tanks, for the full horizon, and for a given scenario. The result of a set of such simulations serves to provide an estimate of the objective function, $E[f(\eta, \omega)]$. The parameters for setting the storage tank location at the plant are then adjusted by means of the outer optimization loop. The outer optimizer uses the parameters of the tanks location as decision variables, whereas the inner problem involves a simulation with a series of embedded scheduling sub-problems.

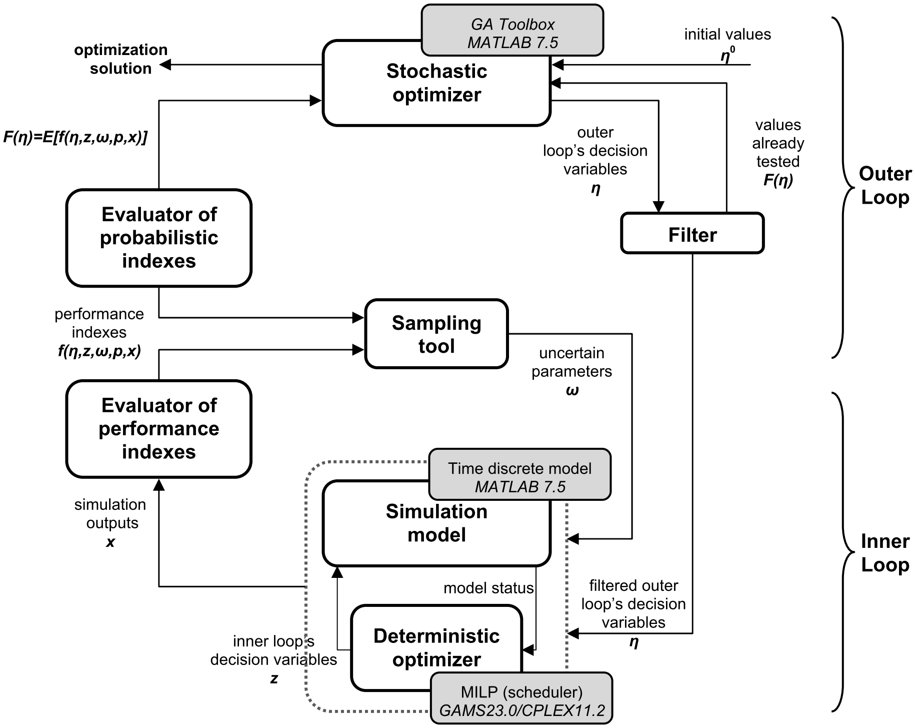

Fig. 4 Scheme of the proposed solution strategy

Next, the outer optimization sub-problem and the inner planning sub-problem, are defined, discussing in more detail the various computational details that are needed to link these sub-problems and drive the computations.

4.1 Outer loop optimization

The objective function of the outer optimization involves a stochastic optimization in which the customers' demand satisfaction is maximized by choosing the appropriate tank location η, through the following equation:

$$\max_{\eta} F(\eta) \tag{3}$$

Equation (3) is evaluated using the inner loop's results of multiple Monte Carlo samplings with embedded simulation, which, in turn, involves *if-then* rules, and scheduling optimizations.

The outer optimization algorithm consists of a GA-based strategy, to improve the values of the variable parameters. The unavailability of explicit equations to model the plant behaviour has encouraged the GA-based strategy as a suitable tool for performing the stochastic optimization in the outer loop. The general resolution strategy is as follows:

- **Step I.** Create an initial population of N randomly generated individuals (a set of parameters for tanks location), η_{gen}^{k}, $gen = 0$, for all products s. The term gen is the counter for the GA generations, i.e., the outer loop iterations, and the parameter k accounts for each individual in the population.

- **Step II.** Run a sufficient number (n) of Monte Carlo samplings from a pool of values of the uncertain variable (e.g. task duration and/or product demand) to generate different sample paths ω.
- **Step III.** Execute the simulator $card(\omega)$-times with its embedded optimizations and rules to obtain a reliable estimate of the expected customers' demand satisfaction for each individual k, $F(\eta_{gen}^k)$. Step II and III are carried out by the inner loop of the SBO strategy.
- **Step IV.** Apply genetic operators over the current population of η_{gen}^k values to generate a new one.
- **Step V.** If $gen \geq MaxGen$, with $MaxGen$ being the maximum number of generations, stop. Otherwise, set $gen = gen + 1$ and go to step II. Until this moment, the best solution found is η_{gen}^*: $F(\eta_{gen}^*) = \max_k\{F(\eta_{gen}^k), k = 1, \ldots, N\}$.

4.2 Inner loop

The computations inside the inner loop optimization involve the repeated simulation of the short-term scheduling of the batch-continuous plant operation over the planning horizon, each with a given Monte Carlo sample of the uncertain parameters. Within each of these simulations, a series of decisions are periodically made to execute a given schedule. The simulator has the outstanding advantage of being capable of representing the real multipurpose/multiproduct plant operation as fitted as desirable. One of the actions taken by the simulator is the call for the scheduler. Each complete simulation run travels across a so-called timeline. The procedure for executing a timeline is as follows:

- **Step I.** At the beginning of each access to the inner loop, take a sample of all the product demands ω_D, generate a deterministic short-term scheduling to comply with these demands.
- **Step II.** Run a simulation for the scheduling horizon H with expected tasks' durations affected by uncertainties, ω_C. Problems arising from deviation from expected values (as explained in Sect. 3) unleashing a set of events, mostly through if-then rules, that drive the simulation imitating the plant operation in the real world.
- **Step III.** Calculate and record performance measuring function, $f(\eta, \omega)$.

By repeating the aforementioned procedure for a number n of Monte Carlo samples, the performance results necessary to compute the expected value, $F(\eta) = E[f(\eta, \omega)]$, can be collected. Note that the accuracy of the expected value will be governed by the number of analyzed timelines and their representativeness.

It is worth noting that solutions given by the simulation-based approach generally cannot guarantee optimality. Nevertheless, the proposed approach is capable of providing reasonable solutions, in time and effort, using building blocks that are readily implementable in practice.

4.3 Implementation of the proposed strategy

The computational scheme for implementing the proposed strategy of Fig. 4 has been programmed using MATLAB 7.5 (2007). The simulator, also encoded in MATLAB, generates the input data for the scheduling model of the plant process, which is used periodically to make production decisions. The scheduling model is formulated as a mixed-integer linear programming problem (MILP), which is solved using GAMS 23.0/CPLEX 11.2 (Brooke et al. 2008). GAMS is interfaced with MATLAB, using the library developed by Ferris et al. (2005).

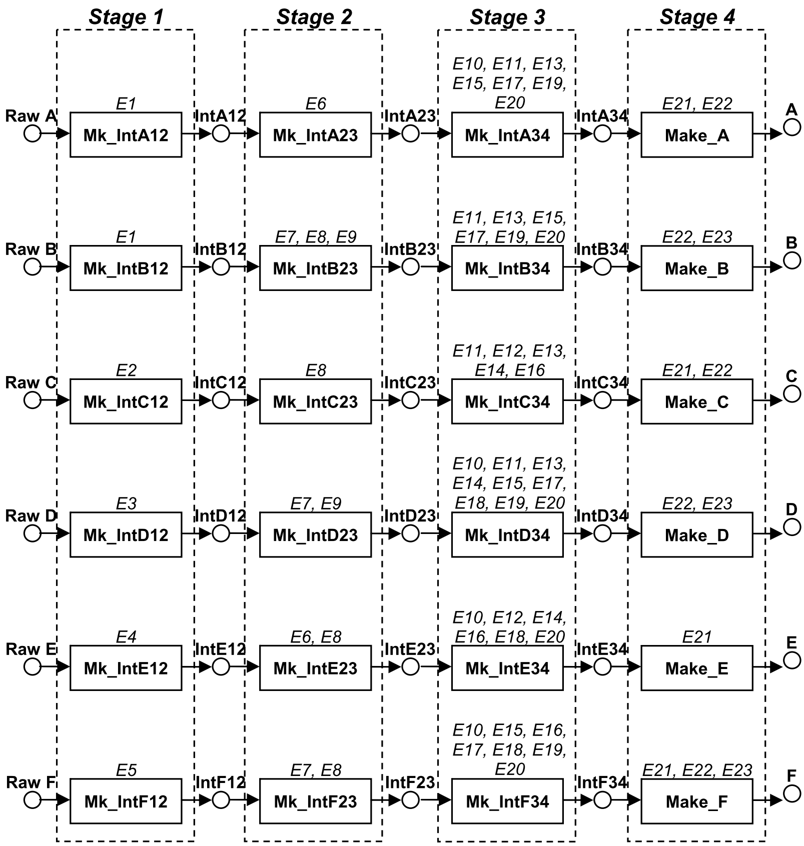

Fig. 5 State-Task Network representation for Case Study I

The objective function calculation module encloses necessary data to calculate the simulation performance (objective function) from the simulation outputs and is encoded in MATLAB. The outer loop optimization is handled by the GA toolbox of the MATLAB 7.5 package, with default option values. This toolbox is able to handle non-linear constraints (needed to avoid repeating tanks' locations) and integer variables.

5 Case Study I: batch paint plant

5.1 Problem description

This case study involves a multipurpose/multiproduct batch plant, where 6 products (A, B, C, D, E and F) are manufactured in 4 stages that are carried out in 23 equipment units (E1-E23). This test example was extracted from Adonyi et al. (2008). Its State-Task Network representation is shown in Fig. 5. It takes into account equations related to mass balances and task duration, and constraints related to allocation, capacity, sequence, tightening, storage and demand. The recipes of the products are given in Table 1. The changeover time is 70 min for equipment units E6 to E9, and 100 min for equipment units E1-E5 and E10 to E20. All

Table 1 Recipes for products A, B, C, D, E, and F

Task	Product A		Product B		Product C		Product D		Product E		Product F	
	Eq.	Time (min)	Eq.	Time (min)	Eq.	Time (min)	Eq.	Time (min)	Eq.	Time (min)	Eq.	Time (min)
1	E1	60	E1	60	E2	60	E3	60	E4	40	E5	40
2	E6	310	E7	240	E8	120	E7	240	E6	300	E7	240
			E8	120			E9	240	E8	120	E8	120
			E9	240								
3	E10	60	E11	120	E11	120	E10	60	E10	60	E10	60
	E11	120	E13	60	E12	70	E11	120	E12	90	E15	120
	E13	60	E15	120	E13	70	E13	60	E14	90	E16	90
	E15	120	E17	60	E14	60	E14	90	E16	90	E17	60
	E17	60	E19	120	E16	50	E15	120	E18	90	E18	90
	E19	120	E20	60			E17	60	E20	60	E19	120
	E20	60					E18	90			E20	60
							E19	120				
							E20	60				
4	E21	720	E22	540	E21	720	E22	540	E21	720	E21	720
	E22	540	E23	720	E22	540	E23	720			E22	540
											E23	720

Table 2 Mean values of demands for each product

Product	A	B	C	D	E	F
Number of batches	3	5	1	3	9	3

other changeover times are supposed to be zero. This case study's plant is able to produce a maximum total of 27 batches of product in a week of operation time (8000 minutes). The number of batches to be produced at the end of the horizon time is given in Table 2.

The inner loop optimization involves a deterministic short-term scheduling problem, and the MILP model was developed with the formulation presented in Shaik and Floudas (2009). Two objective functions are generally used in short-term scheduling: (a) profit maximization and (b) makespan minimization. For this case study makespan minimization is chosen in order to free the process units at earlier times and decrease the impact of parametric uncertainty on the objective function of the outer loop (mentioned below).

The simulation model is a linear discrete-time model capable of meeting the timing constraints, mass balances at production units and storage tanks, and the bounds on variables.

The flowchart shown in Fig. 6 represents the system of rules utilized in the simulation model to avoid incurring in violation of units' limits and storage vessels' capacities.

The objective for the outer optimization loop is to maximize the likelihood (expected value) of complying with consumers demands within a week of production (8000 minutes), by locating N_t storage tanks with a given maximum capacity in the flowsheet.

The sources of uncertainties are two: product demands and duration of batch cycle times. Regarding the demand for each product, the mean value has been perturbed with a factor which variation follows a normal distribution *Norm*(1,1). Demand values are then corrected to integer values, and checked to be positive. Moreover, total number of demanded batches

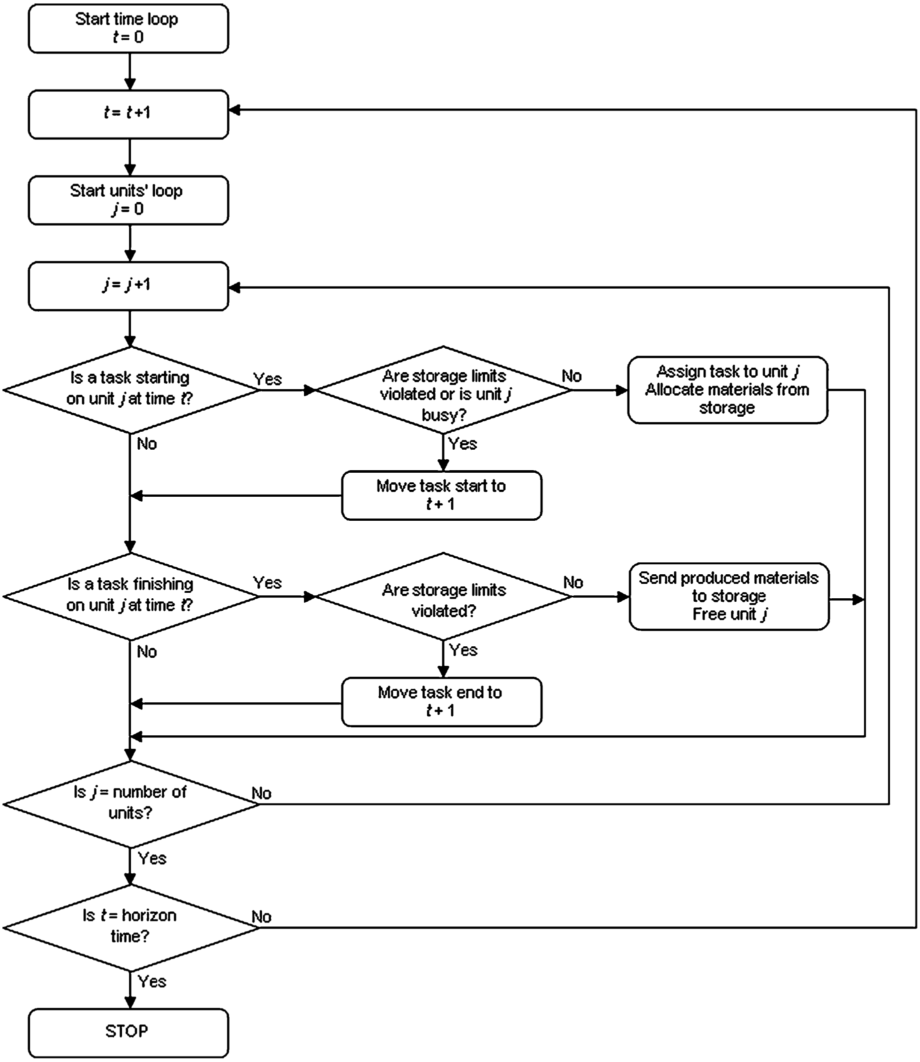

Fig. 6 Flowchart of simulation model for Case Study I

cannot be greater than 27 (maximum plant capacity). Each time the objective function is accessed; three samplings are done over the product demands, each one with their corresponding deterministic scheduling. For each one of the samplings five simulations are carried out. It has been found that with this number of samplings and simulations is enough to get a representative value of the objective function.

For the case of batch cycle times, the mean value has been affected by a factor which also has a normal distribution $Norm(1,\sigma)$. Two values for the standard deviation have been considered: $\sigma = 0.010$ (Case Study Ia, or CSI-a) and $\sigma = 0.025$ (Case Study Ib, or CSI-b). The resulting configuration utilized in this example involves around 50 uncertain parameters; the exact number depends on the time all the tasks are executed.

Table 3 Results of Case Study I

	CSI-a	CSI-b
Population	32	16
Generations	8	10
Objective function (best individual/s)	97.92%	97.69%
Found at generation	0	5
Objective function value for the no-tanks configuration: 73.11%		
Valid solutions	$(24-7+1)^3 = 5{,}832$	$(24-7+1)^3 = 5{,}832$
Solutions found	$32 \times 9 = 288$	$16 \times 11 = 176$
Unique solutions	137	61
Average CPU time per scheduling[a]	15.41 sec	15.79 sec
Total CPU time[a]	6 h 59 min 7 sec	3 h 40 min 21 sec

[a]Computing time using an AMD Athlon 64 X2 Dual Core with 1 GB of RAM

5.2 Results

Case Study I has been divided into two sub-cases. In CSI-a, the GA-based outer loop has been set with a population of 32 individuals and a maximum number of generations of 8. In CSI-b, the population size has been 16 individuals and the maximum number of generations equal to 10. For both CSI-a and CSI-b, the number of storage tanks to locate is three, each with a maximum storage capacity of 3 batches.

The location of all storage vessels has been restricted to the storage of intermediate materials, and tanks locations are allowed to superpose. This last situation is considered as if the storage capacity of such intermediate has duplicated or triplicated.

Results for both CSI-a and CSI-b are reported in Table 3. The population size and number of generations for the GA are shown, as are the objective function value of the best solution found in each case and the generation where they have been first located. For comparison, the value of the objective function when there are no tanks is also presented. This value has been obtained by accessing the inner loop (realizing 3 demands' sampling with deterministic scheduling, each with 5 simulations) with no locations selected for the tanks. "Valid solutions" indicates how many possible combinations of tanks' locations are valid, while "Solutions found" shows how many combinations have been tried in the sub-cases. Some of the combinations tried in each sub-cases have been repeated, as explained in Sect. 4, those solutions are not cycled again through the inner loop, but their values are repeated by a filter. Therefore they do not add computational load to the sub-case solution process. The numbers in the "Unique solutions" row indicate how many non-repeated solution have been tried. "Average CPU time per scheduling" and "Total CPU time" stand for the elapsed computing time to solve the average scheduling problem of the inner loop and the stochastic optimization of the outer loop, respectively.

As a general rule in the proposed framework, the values of the decision values of the outer loop do not have a direct impact on the inner loop objective function's value. This characteristic stems from the fact that outer loop's decision variables move, if they are correctly chosen, the expected value of an objective function that is affected by the present parametric uncertainty, a condition that is not "seen" by the objective function of the inner loop's deterministic optimization. In Case Study I this translates to the tanks' location having little or no impact over the minimization of makespan, as is shown in Table 4, where the difference on the average values is around 1%.

Table 4 Inner loop results of Case Study I

Average makespan	No Tanks (min)	CSI-a (min)	CSI-b (min)
Deterministic scheduling[a]	6511.1	6453.3	6589.4
Simulation[b]	12347.0	8048.7	8509.0

[a]Average over 100 runs

[b]Average over 3 demands' samplings, each one with 5 simulations

Table 5 Solutions of Case Study Ia

Solution	Individuals[a]	Tank 1	Tank 2	Tank 3
1	5	Int.C for tasks 2-3	Int.E for tasks 2-3	Int.D for tasks 1-2
2	2	Int.C for tasks 2-3	Int.D for tasks 3-4	Int.D for tasks 1-2
3	1	Int.C for tasks 2-3	Int.E for tasks 2-3	Int.D for tasks 3-4

[a]Number of individuals with the same solution at the final GA generation

However, the impact of the outer loop's decision variables on the value of the makespan length calculated from the Monte Carlo simulations is expected to be large, since the simulations do consider parametric uncertainty. As it can be seen in Table 4, the average makespan is reduced in 34.8% and 31.1% from the no-tanks configurations to the better ones found in CSI-a and CSI-b. This fact is coherent, since a smaller average makespan means more batches produced in less time, thus increasing the likelihood of complying with product demand.

5.2.1 Results Case Study Ia

Three solutions were found with the best objective value of Case Study Ia, with more than one individual with the same combination in two of them:

Analysing the results on Table 5, it can be seen that solution 1 has been found at $gen = 0$, which would seem that a population size of 32 is too large for this problem. In all three solutions a tank is placed for the recipe of product C (between stages 2–3), although it is the product with lowest mean demand. Finally, all solutions locate a storage tank in the lines for product D or products D and E.

5.2.2 Results Case Study Ib

As in Case Study Ia, three solutions were found with the best objective function value. They are shown in Table 6.

In CSI-b, solution 1 was found at $gen = 5$. On increasing the standard deviation, σ, in the distribution of perturbation for batch cycle times (0.025 for Case Study Ib), the stochastic solver allocates the tanks for those products with greater mean demands (products B, E and F). This result indicates great sensibility of the objective function to the standard deviation.

Table 6 Solutions of Case Study Ib

Solution	Individuals[a]	Tank 1	Tank 2	Tank 3
1	7	Int.E for tasks 1-2	Int.B for tasks 3-4	Int.F for tasks 2-3
2	2	Int.E for tasks 1-2	Int.B for tasks 2-3	Int.F for tasks 2-3
3	1	Int.E for tasks 2-3	Int.B for tasks 2-3	Int.F for tasks 3-4

[a]Number of individuals with the same solution at the final GA generation

6 Case Study II: continuous paint plant

6.1 Problem description

The test example used in this Case Study is a modified version of the problem presented in Sect. 3.1.2 of the work of Ierapetritou and Floudas (1998). This is a well known benchmark case, which has been used in several works presenting short-term scheduling formulation techniques.

Figure 7 shows the State-Task Network representation of the studied paint plant. It obtains 15 different products from 3 different bases. The 3 bases are processed to make 7 different intermediates using 3 mixers. The intermediates are stored in storage tanks and then sent to 5 packing lines that produce the 15 final products. Unlike the previous case study, all the equipment units operate in continuous mode and there are necessary clean-up times for certain combinations of operations. The operation rates and clean-up requirements of each unit are given in Table 7 and the minimum required mass of each final product at the end of the week (120 hours) are given in Table 8.

In the original version of the problem, three tanks of 60 tons each are used for the storage of intermediates materials Int1 to Int7. The tanks are available for all intermediates provided that they only store one intermediate at any time. The modified problem presents the variation of having dedicated storage tanks for each intermediate. The aim of our strategy is to find their optimal location.

The short-term scheduling in this case study is carried out with a model obtained with the technique of Castro et al. (2004). Again, the simulation loop is done with a linear discrete-time model, complying with timing constraints, mass balances at production units and storage tanks, and variables' limits. The systems of rules utilized to assure limits' complaining is shown through the flowchart of Fig. 8.

The goal of the outer optimization loop is to maximize the expected value of complying with consumers demands within a week of production (8000 minutes). Case Study II has been divided into three sub-cases. In the first one (Case Study IIa, or CSII-a), the objective is to find the optimal location of three tanks with the same capacity (60 tons). The goal of the Case Study IIb, (CSII-b) is to get the optimal location of three different tanks, with capacities of 30, 60 and 120 tons respectively. Finally, the third one (Case Study IIc, CSII-c) involves 7 different tanks, each one with capacities between 0 (no tank) to 120 tons, with the optimal capacity value decided by our algorithm. Case Study IIc, is the fully expanded expression of the problem of optimal tank location, with no limits imposed over the number or capacities of storage tanks. It is included to validate the results obtained in CSII-a and CSII-b.

The sources of uncertainties are again the products demands and the durations of processing tasks. The uncertainty in product demands was modeled as a normal distribution, with mean values equal to the rates indicated in Table 7, and a standard deviation of 10%.

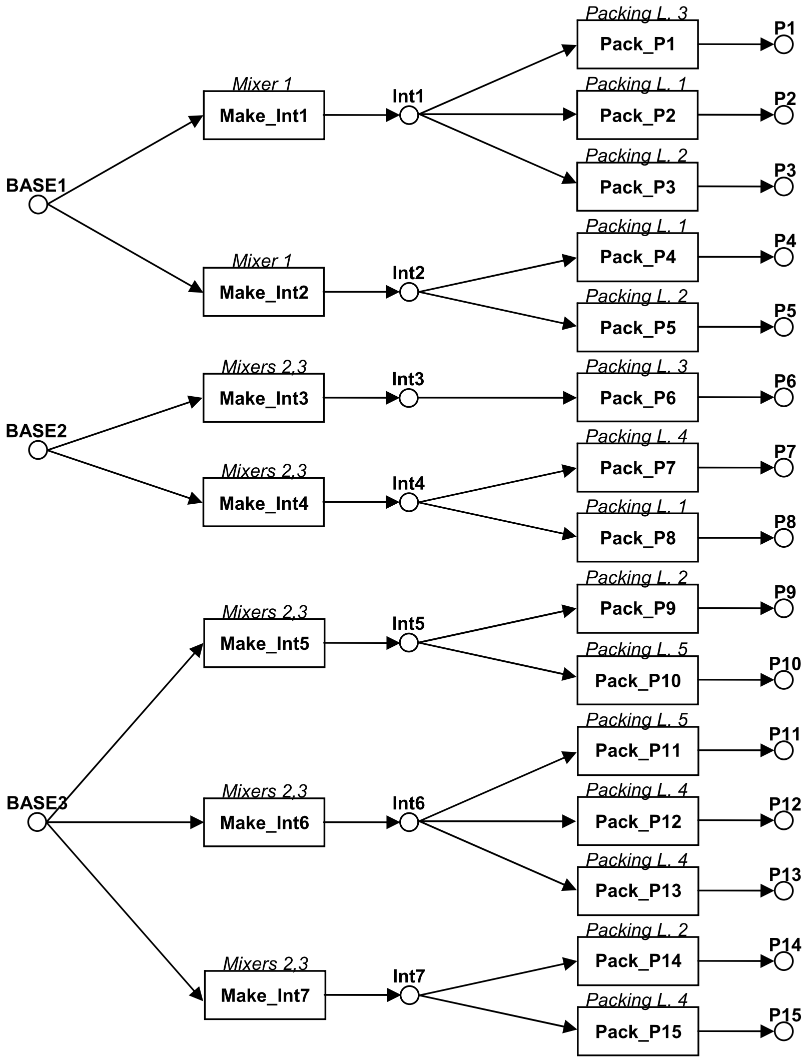

Fig. 7 State-Task Network representation of Case Study II

Processing tasks' durations are obtained by multiplying the value predicted by the deterministic scheduling by a stochastic factor with a normal distribution of $Norm(1, 2.5\%)$. Since in each simulation the total quantity of tasks done is around 100, so is the number of uncertain parameters.

To assure a representative value for each objective function, five products' demands samplings, each with ten simulations were performed.

6.2 Results

Table 9 shows the results for cases CSII-a, CSII-b and CSII-c. CSII-a and CSII-b have been solved using 10 generations of 8 individuals, thus generating 88 accesses to the GA objective function, out of 343 valid combinations. For both cases, combinations with more

Table 7 Equipment's data for Case Study II

Equipment	Rate (ton/h)	Clean time (h)
Mixer 1	17.00	–
Mixer 2	12.24	–
Mixer 3	12.24	–
Pack Line 1	5.833	1
Pack Line 2	2.708	4
Pack Line 3	5.571	1
Pack Line 4	3.333	2
Pack Line 5	5.357	–

Table 8 Product demand for Case Study II

Product	Demand (ton)	Product	Demand (ton)
P1	220	P9	42.5
P2	251	P10	114.5
P3	116	P11	53
P4	15	P12	2.5
P5	7	P13	16.5
P6	47	P14	13.5
P7	8.5	P15	17.5
P8	144		

Table 9 Results of Case Study II

	CSII-a	CSII-b	CSII-c
Population	8	8	32
Generations	10	10	10
Objective function (best individual/s)	96.32%	98.11%	99.45%
Found at generation	0	2	3
Objective function value for the no-tanks configuration: 66.27%			
Valid solutions	$7^3 = 343$	$7^3 = 343$	$25^7 = 6{,}103{,}515{,}625$
Solutions found	$11 \times 8 = 88$	$11 \times 8 = 88$	$11 \times 32 = 352$
Unique solutions	22 (8 not valid)	20	220
Average CPU time per scheduling[a]	11.71 sec	18.44 sec	13.06 sec
Total CPU time[a]	1 h 16 min 23 sec	1 h 48 min 49 sec	19 h 54 min 44 sec

[a]Computing time using an AMD Athlon 64 X2 Dual Core with 1 GB of RAM

than one storage tank per intermediate were considered not valid; therefore, those combinations were assigned a low customers' demand satisfaction value (30%). In CSII-c there are 6,103,515,625 valid combinations (since 7 tanks can have any of the capacity values between 0 and 120 tons with 5 tons intervals), thus the population size was increased to 32.

As in Case Study I, the objective function value of the deterministic scheduling is not expected to be greatly affected by the outer loop's decision variables. This is expressed in Table 10, where the results of inner loop's scheduling show little variation of the plant

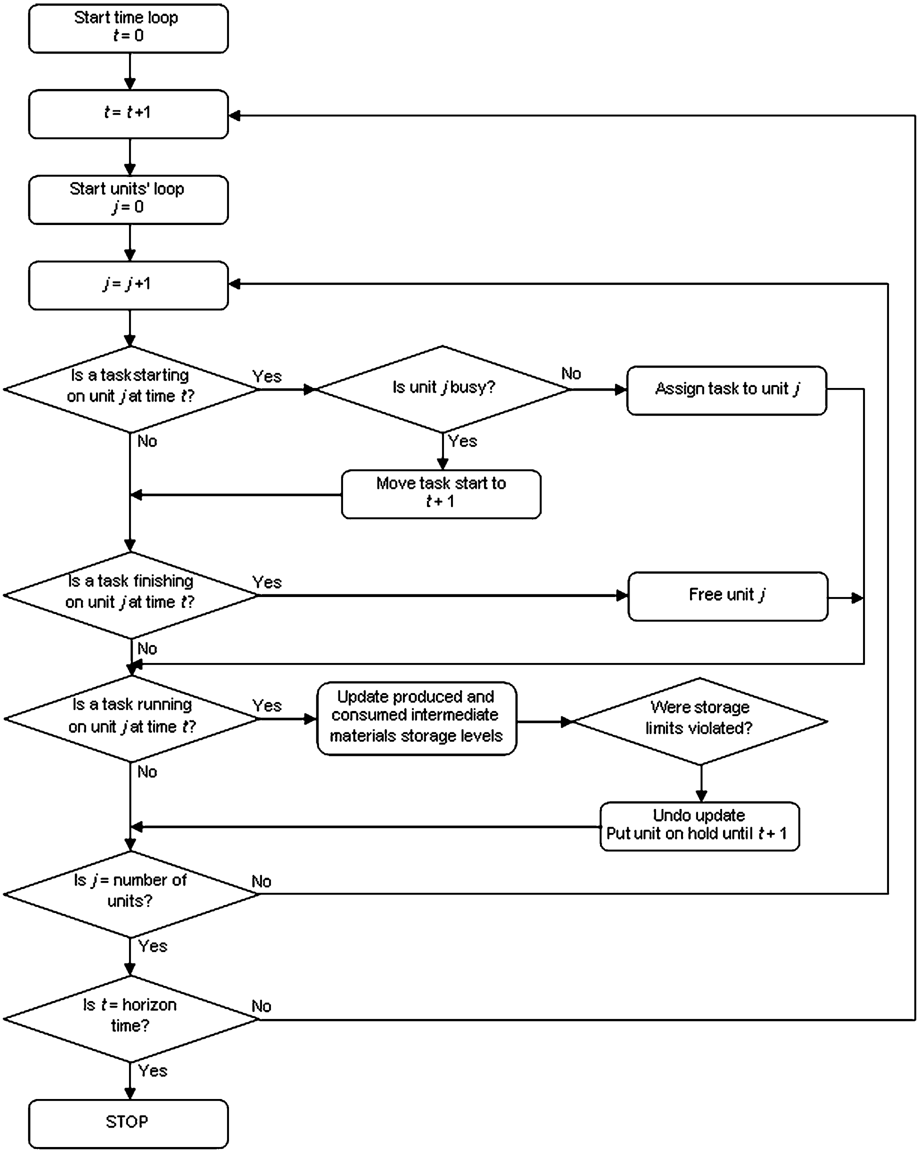

Fig. 8 Flowchart of simulation model for Case Study II

average makespan between the no-tanks configuration and the best solution of each subcase.

But, again as in Case Study I, there is great impact of the outer loop's decision variables over the simulations' average makespans, with averages 15.5% to 19.8% smaller for the best solutions found.

Table 10 Inner loop results of Case Study II

Average makespan	No Tanks (h)	CSII-a (h)	CSII-b (h)	CSII-c (h)
Deterministic scheduling[a]	93.53	93.12	93.04	92.86
Simulation[b]	147.08	124.21	125.88	117.94

[a] Average over 100 runs

[b] Average over 5 demands' samplings, each one with 10 simulations

Table 11 Solution of Case Study IIa

Individuals[a]	Tank 1 (60 tons)	Tank 2 (60 tons)	Tank 3 (60 tons)
5	Intermediate Int2	Intermediate Int5	Intermediate Int4

[a]Number of individuals with the same solution at the final GA generation

Table 12 Solution of Case Study IIb

Individuals[a]	Tank 1 (30 tons)	Tank 2 (60 tons)	Tank 3 (120 tons)
6	Intermediate Int2	Intermediate Int5	Intermediate Int4

[a]Number of individuals with the same solution at the final GA generation

6.2.1 Results Case Study IIa

In the final generation the best solution is present with 5 individuals, each with the combination shown in Table 11, which gives a customers' demand satisfaction value of 96.32%.

The best solution was found at generation 0, meaning that the population size is too large for this problem. The best solution found locates a tank for intermediate Int2, which is transformed into products P4 and P5, both with some of the lowest demands; thus, any mass that fails to comply with their consumers' demand has a greater impact in the objective function. A tank is also located for intermediate Int5, which is transformed into products P9 and P10. The packing of P9, with a lower than average demand, is done by Packing Line 2, one of the busiest and slowest equipment units, thus necessitating a tank as a buffer. The tank for intermediate Int4 is needed because a combination of one of the lowest demands and slow and busy equipment (Packing Line 4) hinders the manufacture of product P4.

6.2.2 Results Case Study IIb

As presented in Table 12, the best solution for CSII-b is found by 6 individuals, with an objective function value of 98.11%.

The best solution, which was first found at generation 2, is very similar to CSII-a's one, with the three tanks used for intermediates Int2, Int4 and Int5, although with different capacities. The outer loop locates the biggest tank for intermediate Int4, which has one of its derived products, P4, with its production being hindered by slow and busy equipment and

Table 13 Solution of Case Study IIc

Individuals[a]	Tank capacity for intermediate (tons)						
	Int1	Int2	Int3	Int4	Int5	Int6	Int7
12	0	20	5	115	85	115	30

[a]Number of individuals with the same solution at the final GA generation

with great impact on the objective function. The improvement of having 60 more tons of available storage over CSII-a solution increases the likelihood of obtaining a better consumers' demand fulfillment value.

6.2.3 Results Case Study IIc

Table 13 shows the best tanks' capacity combination for Case Study IIc. In the final generation the best solution is present with 12 individuals, each with the following combination which gives a customers' demand satisfaction value of 99.45%.

Found at generation 3, the best solution for CSII-c almost repeats the available storage values for intermediates Int2, Int4 and Int5 of the best solution of CSII-b. By also allocating storage availability for the other intermediates excepts Int1, the objective function value improves to almost 100%, but the present problem does not considers the extra capital cost incurred by having more storage vessels.

7 Conclusions

A simulation-based optimization (SBO) strategy has been implemented to solve the optimal location of storages tanks, in multipurpose/multiproduct batch and/or continuous plants under parametric uncertainty.

The strategy has shown to be effective to handle problems with large number of uncertain parameters. It also has enough flexibility to allow the combination of different solution algorithms that better fit the characteristics of each application.

Two case studies corresponding to realistic problems have been solved. The Case study I represents a multipurpose/multiproduct batch plant problem in the paint industry, with a large number of possible pathways for product manufacturing. The SBO was able to find the optimal tank locations reaching a likelihood of customers' demand satisfaction to almost 98% as compared with only 70% for the case without intermediate tanks in a reasonable computing time for a design problem. Our strategy was able to solve the case study in almost 7 hours of computing time in CSI-a, but a comparable objective function value was obtained in CSI-b with only 3 hours and 40 minutes of running time.

Case study II corresponds also to a multipurpose/multiproduct paint plant but with continuous operation mode, with approximately 100 uncertain parameters. The SBO strategy was able to solve in less than 2 hours of computing time the design problem of optimally locating 3 tanks (cases CSII-a and CSII-b), increasing the likelihood to 96 and 98% respectively. These results were validated running a fully general case, where no limit was imposed in the number and capacity of the tanks; verifying that the optimal locations found in the cases with three tanks were kept in the same positions. The fully general case required 19 h of computing time and increased the likelihood of customers' satisfaction to 99.5%.

The computational performance showed in both case studies proves that the proposed approach can be applied not only for the designing stage of the plant, but also for re-designing of long- and/or medium-term planning.

Acknowledgements This work was partially supported by "Consejo Nacional de Investigaciones Científicas y Técnicas" (CONICET) and "Universidad Nacional del Sur" of Argentina.

References

Adonyi, R., Biros, G., Holczinger, T., & Friedler, F. (2008). Effective scheduling of a large-scale paint production system. *Journal of Cleaner Production*, *16*, 225–232.

Aydogan, S., Orçun, S., Blau, G., Pekny, J. F., & Reklaitis, G. V. (2005). A simulation-based optimization approach to the evolution of an advanced life support system for Mars base. In *Proceedings of 2005 AIChE annual meeting*, Cincinnati (EEUU), 10/30/05–11/04/05, oral presentation code 104f.

Balasubramanian, J., & Grossmann, I. E. (2004). Approximation to multistage stochastic optimization in multiperiod batch plant scheduling under demand uncertainty. *Industrial & Engineering Chemistry Research*, *43*, 3695–3713.

Barton, R. R. (1994). Metamodeling: a state of the art review. In J. D. Tew, S. Manivannan, D. A. Sadowski, & A. F. Seila (Eds.), *Proceedings of the 1994 winter simulation conference* (p. 237).

Birge, Z., & Louveaux, S. (1997). *Principles on stochastic programming*. New York: Springer.

Brooke, A., Kendrick, D., Meeraus, A., & Raman, R. (2008). *GAMS: a user guide*. Washington: GAMS Development Corporation.

Castro, P. M., Barbosa-Póvoa, A. P., Matos, H. A., & Novais, A. Q. (2004). Simple continuous-time formulation for short-term scheduling of batch and continuous processes. *Industrial & Engineering Chemistry Research*, *43*, 105–118.

Cheng, L., Subrahmanian, E., & Westerberg, A. W. (2003). Design and planning under uncertainty: issues on problem formulations and solutions. *Computers & Chemical Engineering*, *27*, 781–801.

Cheng, L., Subrahmanian, E., & Westerberg, A. W. (2004a). A comparison of optimal control and stochastic programming from a formulation and computation perspective. *Computers & Chemical Engineering*, *29*, 149–164.

Cheng, L., Subrahmanian, E., & Westerberg, A. W. (2004b). Multiobjective decisions on capacity planning and inventory control. *Industrial & Engineering Chemistry Research*, *43*, 2192–2208.

Diwekar, U. (2002). Optimization under uncertainty: an overview. *SIAG/Optimization Views-and-News*, *13*(1), 1–8.

Duenas, A., & Petrovic, D. (2000). An approach to predictive-reactive scheduling of parallel machines subject to disruptions. *Annals of Operations Research*, *159*(1), 65–82.

Ferris, M. C., Dirkse, S., & Jain, R. (2005). MATLAB and GAMS: interfacing optimization and visualization software. Computer Sciences Department, University of Wisconsin at Madison, Madison. http://www.cs.wisc.edu/math-prog/matlab.html. Accessed October 7, 2010.

George, E. D., & Farid, S. S. (2008). Stochastic combinatorial optimization approach to biopharmaceutical portfolio management. *Industrial & Engineering Chemistry Research*, *47*, 8762–8774.

Gosavi, A. (2003). *Simulation-based optimization: parametric optimization techniques and reinforcement learning*. New York: Springer. ISBN-10: 1402074549, ISBN-13: 978-1402074547.

Gürkan, G. (2000). Simulation optimization of buffer allocations in production lines with unreliable machines. *Annals of Operations Research*, *93*, 177–216.

Ierapetritou, M. G., & Floudas, C. A. (1998). Effective continuous-time formulation for short-term scheduling: II. Continuous and semi-continuous processes. *Industrial & Engineering Chemistry Research*, *37*, 4360–4374.

International Society for Measurement and Control (1995). ANSI/ISA-S88.01-1995, Batch Control, Part 1: Models and Terminology, 23/10/1995.

International Society for Measurement and Control (2000). ISA-S95.01-1999, Enterprise Control System Integration, Part 1: Models and Terminology. Draft 13.

Jung, J. Y., Blau, G., Pekny, J. F., Reklaitis, G. V., & Eversdyk, D. (2004). A simulation based optimization approach to supply chain management under demand uncertainty. *Computers & Chemical Engineering*, *28*, 2087–2106.

Kim, M., Rao, A. S., & Yoo, C. (2009). Dual optimization strategy for N and P removal in a biological wastewater treatment plant. *Industrial & Engineering Chemistry Research*, *48*, 6363–6371.

Kleinman, N. L., Spall, J. C., & Naiman, D. Q. (1999). Simulation-based optimization with stochastic approximation using common random numbers. *Management Science*, *45*, 1570.

Kuster, J., Jannach, D., & Friedrich, G. (2010). Applying local rescheduling in response to schedule disruptions. *Annals of Operations Research*, *180*(1), 265–282.

Law, A. M., & Kelton, W. D. (2000). *Simulation modeling & analysis* (3rd edn.). New York: McGraw-Hill.

MATLAB 7.5 (2007). *Simulink 6.0, Stateflow 6.0. User's manual*. Natick: The MathWorks, Inc.

Mele, F. D., Espuña, A., & Puigjaner, L. (2006a). Supply chain management through dynamic model parameters optimization. *Industrial & Engineering Chemistry Research*, *45*, 1708–1721.

Mele, F. D., Guillén, G., Espuña, A., & Puigjaner, L. (2006b). A simulation-based optimization framework for parameter optimization of supply-chain networks. *Industrial & Engineering Chemistry Research*, *45*, 3133–3148.

Nedderneijer, H. G., van Oortmarssen, G. J., Piersma, N., & Dekker, R. (2000). A framework for response surface methodology for simulation optimization. In *Proceedings of the 1999 winter simulation conference*.

Shaik, M. A., & Floudas, C. A. (2009). Novel unified modeling approach for short-term scheduling. *Industrial & Engineering Chemistry Research*, *48*, 2947–2964.

Silva, H. G., & Salcedo, R. L. R. (2009). A coupled strategy for the solution of NLP and MINLP optimization problems—benefits and pitfalls. *Industrial & Engineering Chemistry Research*, *48*, 9611–9621.

Subramanian, D., Pekny, J. F., & Reklaitis, G. V. (2001). A simulation-optimization framework for research and development pipeline management. *AIChE Journal*, *47*, 2226–2242.

Subramanian, D., Pekny, J. F., Reklaitis, G. V., & Blau, G. E. (2003). A simulation-optimization framework for stochastic optimization of R&D pipeline management. *AIChE Journal*, *49*, 96–112.

Sundaramoorthy, A., & Maravelias, C. T. (2008). Modeling of storage in batching and scheduling of multistage processes. *Industrial & Engineering Chemistry Research*, *47*, 6648–6660.

Sundar Raj, T., & Lakshminarayanan, S. (2008). Multiobjective optimization in multiechelon decentralized supply chains. *Industrial & Engineering Chemistry Research*, *47*, 6661–6671.

Wan, X., Pekny, J. F., & Reklaitis, G. V. (2005). Simulation-based optimization with surrogate models—application to supply chain management. *Computers & Chemical Engineering*, *29*, 1317–1328.