Theory and Applications of Mathematics & Computer Science

(ISSN 2067-2764)
http://www.uav.ro/applications/se/journal/index.php/tamcs

Theory and Applications of Mathematics & Computer Science 1 (2) (2011) 71–88

Design and Scheduling of Chemical Batch Processes: Generalizing a Deterministic to a Stochastic Model

João Luís de Miranda[a,*], Miguel Casquilho[b]

[a]*Escola Superior de Tecnologia e Gestão, Instituto Politécnico de Portalegre: Lugar da Abadessa, Apt. 148; 7301-999 Portalegre; Portugal.*
[b]*Instituto Superior Técnico, Universidade Técnica de Lisboa (IST/UTL) Ave. Rovisco Pais, IST; 1049-001 Lisboa; Portugal, Centre for Chemical Processes at IST/UTL.*

Abstract

A stochastic optimization model for the design and scheduling of batch chemical processes is developed in a Two-Stage Stochastic Programming framework, with the uncertainty formulated through a number of discrete scenarios. The sparse model presents binary variables in the first stage and systematically generalizes a deterministic model chosen from the literature, in an approach based on computational complexity. The combination of single product campaign (SPC) with multiple machines was found to be the most promising from a computational standpoint, and it is here generalized toward a stochastic environment within the relaxation of the soft demand constraints. Numerical examples are presented, and the results point to a significant reduction of 8-20 % of the investment costs in comparison to the SPC non-relaxed case, without real losses if the multiple product campaign (MPC) policy is adopted.

Keywords: robust optimization, two-stage stochastic programming, design, scheduling, batch processes.
2000 MSC: 90B36.

1. Introduction

The number of industrial cases published on the design and scheduling of batch processes is very small, and the industrial works published in the open literature are also illustrative of the difficulties to conjugate different time ranges with efficiency and detail, as referred in the reviews of (Floudas & Lin, 2004) and (Barbosa-Póvoa, 2007).

This paper addresses the design of batch chemical processes and simultaneously considers the scheduling of operations. A generalization of literature models (Voudouris & Grossmann, 1992) is proposed, from a deterministic Mixed Integer Linear Programming (MILP) model to a Two-Stage Stochastic Programming (2SSP) one. The generalization to a stochastic, multiperiod, and robust model is based on computational complexity studies (Miranda, 2011*a*). When reducing the multiperiod model into one single time period, the optimality study shows that significant reduction of investment costs is possible.

*Corresponding author
Email addresses: jlmiranda@estgp.pt (João Luís de Miranda), mcasquilho@ist.utl.pt (Miguel Casquilho)

The generalized model treats different time ranges, namely, the investment and scheduling horizons. Furthermore, the 2SSP framework allows robustness promotion: i) in the solution, by penalizing the deviations; and ii) in the model, with relaxation of the integrality constraints in the second phase variables.

In the design of batch processes for multiproduct units (flowshop batch plant), the implementation of difficult MILP/MINLP models is common: it is required to design and enumerate the equipment, simple or distributed in parallel, or to consider single product (SPC) or multiple products (MPC) campaigns. This usually corresponds to the first phase of the 2SSP framework. (The word "phase" is used here instead of "stage" because of the traditional use of "stage" in chemical engineering scheduling problems).

In order to better select the equipments, the optimal production policy must also be found since it directly affects the equipment sizing. However, it involves the detailed solution of scheduling subproblems where decomposition schemes are pertinent. These subproblems are focused in the second phase of 2SSP, where the control variables (recourse) occur. The integer and binary variables related to the scheduling and precedence constraints are disregarded as control variables, as they would make very hard the treatment of the recourse problem. Consequently, the second phase variables are assumed continuous (for example, the number of batches) and binary variables occur only in the 2SSP first phase.

The generalization approach is not like the one proposed by (Moreno *et al.*, 2007), who treat a similar problem of design and scheduling of batch processes. This approach coincides with Ahmed and Sahinidis (Ahmed & Sahinidis, 2000), (Ahmed & Sahinidis, 2003) and (Liu & Sahinidis, 1997) in the sense that analytical studies of computational complexity can yield significant improvements in terms of algorithms and problem structures, with good and realistic solutions.

In this study, the generalization to a 2SSP framework within a multiperiod and robustness environment points to a significant reduction (8%-20 %) in the investment costs. And the results motivate to pursue this generalization approach, with foreseen developments.

The paper sequentially addresses: in Section 2, the design and scheduling of batch processes, and the deterministic model that is enlarged here; in Section 3, the presentation of the generalized model on a stochastic and robust 2SSP framework; in Section 4, the optimality study, which is directed to promote robustness; finally, the main conclusions are presented in Section 5.

2. The Design and Scheduling of Batch Processes

In this section, the issue of scheduling of batch processes is integrated with process design (sizing), combining the short term decisions with the long term investment planning. The models that were systematically studied in (Miranda, 2007) are from the open literature and the state-of-art at that time is described.

The study of existing models in the literature induces the enlargement of models and related applications (Miranda, 2007), and this generalization of models simultaneously causes increasing complexity and difficulties. A design and scheduling, deterministic, and single time period model (Voudouris & Grossmann, 1992) that seems to have no improvements for more than a decade is addressed in this paper.

The models studied were those in (Voudouris & Grossmann, 1992), and appear to belong to a research line initiated in (Birewar & Grossmann, 1989) , featuring a quick resolution, and finished in (Voudouris & Grossmann, 1993), realizing the approach impracticability. The approach option was changed to the jobshop framework in (Voudouris & Grossmann, 1996), but flowshop continues to be widely used in chemical industry.

The options set and the successive generalizations are the main criteria to the models selection: i) single machine vs. multiple parallel machines in each stage; ii) single product campaigns (SPC) or multiple products (MPC); and iii) assumption of some storage policy or zero wait (ZW) operations. The complexity of the problems leads to the adoption of alternative methodologies, such as the evolutionary procedures of Xia and Macchietto in (Xia & Macchietto, 1997) and Tan and Mah in (Tan & Mah, 1998). Pekny and Miller in (Pekny & Miller, 1991) recommended the utilization of heuristics whenever a satisfactory method is not available for the problem at hand. The latter authors treated the scheduling flowshop problem with a ZW policy, both exactly, through a branch-and-bound algorithm, and heuristically, through a permutations procedure and simulated annealing. Jayaraman et al. in (Jayaraman *et al.*, 2000) addressed the design and scheduling of batch processes using a heuristic method (ant colony) and, they obtained exact results for short sized instances. Cavin et al. in (Cavin *et al.*, 2004) used tabu search to address the combinatorial and multiobjective optimization in the design and scheduling of a multipurpose batch plant problem.

A different path was adopted by Liu and Sahinidis in (Liu & Sahinidis, 1997) and Ahmed and Sahinidis in (Ahmed & Sahinidis, 2000) and (Ahmed & Sahinidis, 2003) when assessing the possibility to develop exact and efficient algorithms. They developed analytical investigations to verify that their problems are NP-hard, both the static version and the dynamic version of their planning process models. Using computational complexity techniques, they verified that it is not possible to develop exact polynomial algorithms and pointed the need to systematically build good procedures.

Moreno et al. in (Moreno *et al.*, 2007) addressed multiproduct batch plant and considered SPC policy. They presented a multiperiod model aiming at the sizing and planning in batch multiproduct plants, considering: return values, and operation and investment costs; assignment of discrete dimensions to the processes, batch or semicontinuous; implementation of intermediate storage; and variations on demands and component prices, due both to seasonal and structural effects. The approach presented here is different from the one of (Moreno *et al.*, 2007), but the kind of problem focused is similar.

This work uses analytical results and applies computational complexity techniques at the deterministic model *MS* (Miranda, 2011*a*), which feature multiple machines per stage and SPC (Multiple machine, SPC). This model, was selected because (Miranda, 2011*b*):

- for industrial applications with realistic number and quantities of products, the option of the multiple processes in parallel at each stage should be considered: otherwise unfeasibility will certainly occur;
- the option for the SPC mode arises from the current difficulties to apply MPC in a *multiple machine* environment, just due to insufficiency of the related model;
- the investment cost assuming SPC is estimated to exceed in near 5 % the cost of the more efficient MPC policy; this surplus results from a selection of the next discrete dimension on nearly half of the stages (values derived from comparable instances); that is, the SPC sizing is *a priori* overdesigned, and this will permit to introduce new products, or even to accommodate un-forecast growth on product demands.

3. Development of a Generalized Robust Model

In this section, a robust model is presented for the design and scheduling of batch processes, generalizing the deterministic model *MS* to a stochastic context in two phases (2SSP), with promotion of robustness. This permits the treatment of the risk associated to medium and long term investments.

When a high investment is needed, usually, a long return term is associated. The generalized model includes the optimization of long term investment and also considers the short term scheduling of batch processes. Deterministic models do not conveniently address the risk of a wider planning horizon, and scheduling models often deal with certain data in a single time horizon. Thus, difficulty increases when the combinatorial scheduling problem is integrated with the uncertainty of the design problem.

The objective of the 2SSP model is the maximization of a robust measure of the Net Present Value (NPV), by selecting the discrete dimensions for the batch processes and the number of processes operating in parallel (out-of-phase operation). In addition, multiple time periods are supposed in the NPV evaluation. Given the uncertainty of the quantities and unit returns of each product, then returns are evaluated in a probabilistic way.

The stochastic model aims to maximize the robust NPV, Φ , considering the expected return minus the investment costs, with the latter occurring only in the first period. Robustness is promoted by penalizing the expected values or estimators of: *i)* the variability of the discrete scenarios solutions, *dvtn*; *ii)* the non-satisfied product demands, *Qns*; and *iii)* the capacity slacks, *slk*. That is:

$$\begin{aligned}[\max]\,\Phi = &\sum_{r=1}^{NR} prob_r\,\xi_r - \lambda dsv \sum_{r=1}^{NR} prob_r\, dvtn_r - \lambda qns \sum_{r=1}^{NR} \frac{prob_r}{NC.NT}\left(\sum_{j=1}^{NC}\sum_{t=1}^{NT} Qns_{jtr}\right)\\ &-\lambda slk \sum_{r=1}^{NR} \frac{prob_r}{M.NC.NT}\left(\sum_{i=1}^{M}\sum_{j=1}^{NC}\sum_{t=1}^{NT} slk_{ijtr}\right).\end{aligned} \tag{3.1}$$

The objective function uses technical estimators (Appendix A) that are built to assess the quality of the generalized model. Each probabilistic component, ξ_r, corresponds to the NPV obtained at each discrete scenario *r*, and this component is obtained from: *i)* the present amount of sales return, obtained in the second phase of the 2SSP (probabilistic net values, related to materials purchases and operations costs); minus *ii)* the investment costs, defined in the first phase of the 2SSP (deterministic and discrete costs, accordingly with the discrete dimensions of equipments at each stage). This means:

$$\xi_r = \sum_{j=1}^{NC}\sum_{t=1}^{NT} ret_{jtr} W_{jtr} - \sum_{i=1}^{M}\sum_{s=1}^{NS(i)}\sum_{p=1}^{NP(i)} c_{isp} y_{isp}\ ,\quad \forall r. \tag{3.2}$$

A linear measure is used for penalizing variability. The negative deviation of NPV, $dvtn_r$, is evaluated at each discrete scenario to measure only the deviations below the NPV's expected value, and provided that each probabilistic component, ξ_r, is computed from relations 3.2,

$$dvtn_r \geq \sum_{r'=1}^{NR} (prob_{r'}\xi_{r'}) - \xi_r \ \geq 0\ ,\quad \forall r. \tag{3.3}$$

Assuming ZW policy and the non-existence of intermediate storage, the non-satisfied demand of each product *j*, Qns_{jtr}, is defined at each time period, *t*, and for each scenario, *r*, by the related definition constraint slack. That is, by the difference between the quantities of product demands, Q_{jtr}, and quantities produced, W_{jtr}. The following constraints sets are employed:

$$W_{jtr} + Qns_{jtr} = Q_{jtr}\ ,\quad \forall j,t,r. \tag{3.4}$$

$$Qns_{jtr} \geq 0,\quad \forall j,t,r. \tag{3.5}$$

The global quantities produced for each product, W_{jtr}, at each discrete scenario and each time period are related to the aggregated batches number, nc_{ijsptr},

$$S_{ij} W_{jtr} \leq \sum_{s=1}^{NS(j)}\sum_{p=1}^{NP(j)} dv_{is} nc_{ijsptr},\quad \forall i,j,t,r. \tag{3.6}$$

Thus, the global excess on the implemented production capacities (slk_{ijtr}) results directly from

$$S_{ij}W_{jtr} + slk_{ijtr} = \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} dv_{is} nc_{ijsptr}, \quad \forall i, j, t, r. \tag{3.7}$$

The disaggregated number of batches, nc_{ijsptr} (further details in Appendix A), corresponds to the product-aggregation of variables ($n_{jtr} \, . \, y_{isp}$). Three logical sets of constraints are required: upper bounds; only one value is selected; and the definition of the selected value. Respectively:

$$nc_{ijsptr} - nc_{ijsp}^{Upp} y_{isp} \leq 0\,, \;\forall\, i, j, s, p, t, r, \tag{3.8}$$

$$\sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} y_{isp} = 1, \quad \forall i, \tag{3.9}$$

$$\sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} nc_{ijsptr} - n_{jtr} = 0, \quad \forall i, j, t, r. \tag{3.10}$$

The campaign times must be determined, $tcamp_{jtr}$, either in relation to the disaggregated number of batches, nc_{ijsptr}, and by satisfying the time horizon, H:

$$\sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} \left(\frac{\tau_{ij}}{p(i)} nc_{ijsptr} \right) - tcamp_{jtr} \; \leq 0, \forall i, j, t, r, \tag{3.11}$$

$$\sum_{j=1}^{NC} tcamp_{jtr} \leq H, \quad \forall t, r. \tag{3.12}$$

Conjugating the described relations, the stochastic model *spbatch_ms* for the design and scheduling of batch chemical processes aims at the maximization of NPV, promotes robustness in solution and model, and it assumes *flowshop* configuration, several processes in parallel at each stage (*multiple machine*), single product campaigns (SPC), and ZW policy:

Model *spbatch_ms*:

$$\begin{aligned}[][\max]\, \Phi = \sum_{r=1}^{NR} prob_r \xi_r - \lambda dsv \sum_{r=1}^{NR} prob_r . dvtn_r - \lambda qns \sum_{r=1}^{NR} \frac{prob_r}{NC.NT} \left(\sum_{j=1}^{NC} \sum_{t=1}^{NT} Qns_{jtr} \right) \\ -\lambda slk \sum_{r=1}^{NR} \frac{prob_r}{M.NC.NT} \left(\sum_{i=1}^{M} \sum_{j=1}^{NC} \sum_{t=1}^{NT} slk_{ijtr} \right), \end{aligned} \tag{3.13}$$

subject to,

Relations 3.2 to 3.4
Relations 3.7 to 3.12

$$\xi_r, dvtn_r, \; slk_{ijtr}, nc_{ijsptr}, \; n_{jtr}\,, \; Qns_{jtr}, \; tcamp_{jtr}, W_{jtr} \geq 0, \quad \forall i, j, s, p, t, r \tag{3.14}$$

$$y_{isp} \in \{0;1\}, \quad \forall i, s, p. \tag{3.15}$$

The application of this generalized model for the design and scheduling of batch processes is illustrated through numerical examples: the design of batch processes is satisfying uncertain demands on a unique time period (”static”), and the minimization of investment costs considers a stochastic and robust formulation. Appendix B contains further details of the development of the generalized 2SSP model, while Appendix C describes how the 2SSP model relates to the deterministic model *MS*.

4. Illustrative Examples

The problem of sizing batch chemical processes is now discussed to illustrate the characteristics of the stochastic model *spbatch_ms*. The former deterministic model, *MS* (theoretically focused in Appendix C), is targeted as reference model. Through the usual reasoning of polynomial reduction of problem instances, the following is assumed: ***i)*** only one time period; and ***ii)*** zero value of return in products.

Given the single time period considered, the unsuitability of NPV maximization must be noticed: NPV is usually addressed in a multiperiod horizon (dynamic problem) due to high investment costs that do not allow payback on the first time period. Then, assuming $ret_{jtr} = 0$, means not to account for cash flows returning, and the NPV variables ξ are representing only the investment costs.

$$\xi = \xi_r = -\sum_{i=1}^{M}\sum_{s=1}^{NS(i)}\sum_{p=1}^{NP(i)} c_{isp} y_{isp}, \quad \forall r. \tag{4.1}$$

The NPV variables ξ are scenario independent, so no variability is presented, and a null deviation, $dvtn_r = 0$, will be observed in all scenarios. One mainly wishes to satisfy the uncertain product demands, and the penalization of capacity slack will not be considered ($slk = 0$). The objective function in equation 3.1 is thus reduced to the robust minimization of investment costs, assuming only the penalization of non-satisfied demand:

$$\begin{aligned}[\max]\,\Phi &= \xi \, . \sum_{r=1}^{NR} prob_r - \lambda qns \sum_{r=1}^{NR} \frac{prob_r}{NC.NT} \left(\sum_{j=1}^{NC}\sum_{t=1}^{NT} Qns_{jtr} \right) \\ &= -\sum_{i=1}^{M}\sum_{s=1}^{NS(i)}\sum_{p=1}^{NP(i)} c_{isp} y_{isp} - \lambda qns \sum_{r=1}^{NR} \frac{prob_r}{NC.NT} \left(\sum_{j=1}^{NC}\sum_{t=1}^{NT} Qns_{jtr} \right),\end{aligned} \tag{4.2}$$

or,

$$[\min]\,\Psi = \sum_{j=1}^{M}\sum_{s=1}^{NS(i)}\sum_{p=1}^{NP(i)} c_{isp} y_{isp} + \lambda qns \sum_{r=1}^{NR} \frac{prob_r}{NC.NT} \left(\sum_{j=1}^{NC}\sum_{t=1}^{NT} Qns_{jtr} \right). \tag{4.3}$$

The model defined in the relations set 3.13 is restricted, and the following can be neglected: the time period index, *t*, because only one time period is considered; the constraint sets concerning the definition of the probabilistic variables, ξ , which will have a constant and scenario-independent value; and, the same for the deviation definitions, *dvtn*, which consequently will be null and useless.

The main characteristics and the average execution times for the instances treated in the various examples (EX1 to EX6) are described in Table 1. A laptop ASUS-F3JC (Intel Core2 T5500, 1.55GHz and 2GB of RAM) and GAMS/OSL are used (data generation in Appendix D).

Table 1: Numbers of parameters, variables and constraints corresponding to examples solved assuming: $NC = 4; M = 3; NS = 5; NP = 3; NT = 1$.

Numerical examples	Parameters NR	Binary variables	Continuous variables	Constraints	Execution times(s)
EX1	1	45	209	226	0.33
EX2	3	45	603	672	1.68
EX3	7	45	1391	1564	7.48
EX4	15	45	2967	3348	21.46
EX5	30	45	5922	6693	91.60
EX6	100	45	19712	22303	940.31

4.1. Robustness and Number of Scenarios

The effect related to the utilization of distinct numbers of discrete scenarios is analyzed in the generalized and stochastic model, which conceptually reduces to the deterministic one when considering a single scenario. Although the number of binary variables is kept constant for the various examples, the number of continuous variables and the number of constraints vary linearly with the number of scenarios, *NR*. Instead, the evolution of the average execution times is approximately quadratic.

Graphical representations are shown for the variation of the different estimators (robust cost, ψ; expected value of the non-satisfied demand, *Ensd*; expected value of the capacity slacks, *Eslk*; and non-robust cost, *Ecsi*), with the increasing scenarios number, NR, and penalization for non-satisfied demand, λqns. The subsequent analysis is presented.

Due to the near coincidence of the different lines represented, from NR = 1 (EX1) to NR = 100 (EX6), six lines may not be presented in some of the graphs. A general and descriptive approach is intended for the various examples, because they are targeted: the estimators' sensitivity to the variation of penalization parameter: and to support model and solution robustness.

From Figure 1, the values of the robust cost (NPV, ψ) vary significantly in the range of λqns from 1 to 4, which reveals the most sensitive zone and to which attention will be drawn. This evolution pattern is similar in the following figures that successively present *Ensd*, *Eslk*, and *Ecsi* evolutions, and a general range of stability for the estimator values is also observed for λqns greater than 5. Although not represented in *Figure 1*, when λqns increases in the range of 20 to 40, the robust costs ($NR > 1$) tend to the deterministic cost: the expected values of the non-satisfied demand, *Ensd*, are already small in that range, so the penalized component shall present a significant value only after a strong increment of λqns.

In relation to Figure 2, the expected value in *Ensd* for $\lambda qns = 0$ (not represented) is the result of the minimum configuration: a single process in each stage and with minimum dimension is selected, to which corresponds a virtual value of high non-satisfied demand (about $83 \times 10^3 kg$). For the first value represented, $\lambda qns = 1$, *Ensd* varies between 30×10^3 and $40 \times 10^3 kg$, representing about 23 % of the average demand ($150 \times 10^3 kg$) or a reduction to less than half in relation to the null penalization parameter value.

The non-satisfied demand concentrates in quite specific situations, worthy of a directed analysis: in one or several products, for one or several instances. From the observation of the optimum values for the variables Qns_{jtr}, it becomes clear that these variables are null in almost all the examples and for almost all the products. It will then be important to verify if the non-satisfaction of the demand of a specific product will have other consequences (accomplishment of trade agreements, customer service, etc.). In this case, penalization parameters associating the product in analysis may be introduced.

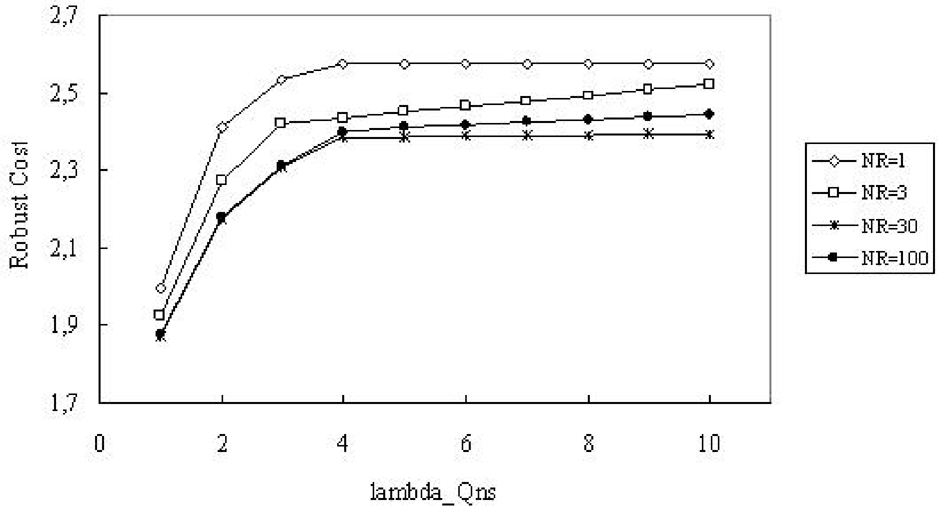

Figure 1: Variation of robust costs (10^5 €) versus the penalty values on non-satisfied demand, λqns, for various numerical examples.

In Figure 3, $Eslk$ tends to stabilize on an upper limit (about 1400 L) and, supposing the maximum configuration of discrete volumes, the plateau represents a percent underutilization (aggregating the discrete capacities of all the stages, 10929 L are obtained) of up to 13 %. Anyway, $Eslk$ represents a value of relative importance, as it depends on the instance at hand.

If, however, attention is paid to the percentage of Eslk as a function of the total capacity, stabilized values in the range 11-13 % are observed (as in Table 3, for $NR = 7$) in all the examples. In fact, for $\lambda qns = 0$, the minimum configuration corresponds to about 3451 L, in which a percent underutilization is verified (aggregated for all the stages) of 6.6 %. For the first values represented ($\lambda qns = 1$), *Eslk* is about 712 L, or 6.5 % of underutilization of the maximum capacity (10929), but corresponding to 11.7 % of the optimum capacity (6087 L) for this example (Table 2 and Table 3). Therefore, when λqns increases, *Eslk* also increases rapidly until the upper limit mentioned (about 1400 L), but the percentages of equipment underutilization in each example stay around between 11 and 13 %.

Figure 4 also yields a strong initial growth of non-robust cost, *Ecsi*, which then tends to stabilize in an upper value. This upper limit of *Ecsi* corresponds to the complete satisfaction of all the instances of the uncertain demand, and it also corresponds to the deterministic cost of investment of about 2.6×10^5 €. This value is significantly greater (about 8 %) than the remaining ones that are around 2.4×10^5€. Nevertheless, *Ecsi* presents a "permanent" value (for $\lambda qns = 0$, the minimum configuration) of about 1.25×10^5€, which represents a fraction of 48 %, relatively to the maximum configuration. For the first values ($\lambda qns = 1$ or 2) presented, *Ecsi* is about 1.6×10^5€, that is, 62 % of the cost of the maximum configuration. Then such configuration would decrease the complementary expected cost by about 38 %.

In Figure 5, the difference between the robust cost (ψ, or $NPVrob$ in the graph legend) and the non-robust cost (*Ecsi*) significantly comes from the consideration of the penalization on the non-satisfied demands, λqns: the variation of robust cost is linearly related to the evolution of this penalization parameter, when λqns is greater than 5 and non-satisfied demand is stabilized on a low value.

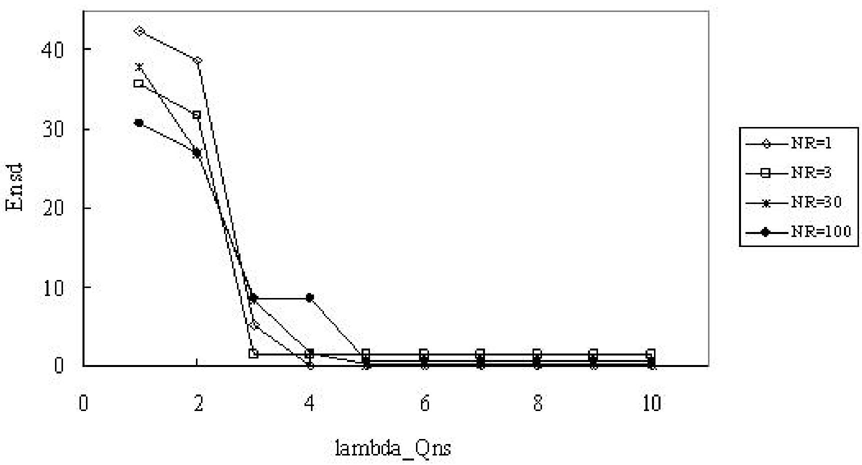

Figure 2: Variation of the expected value for non-satisfied demand, $Ensd(10^3 kg)$, with the evolution of λqns, for various numerical examples.

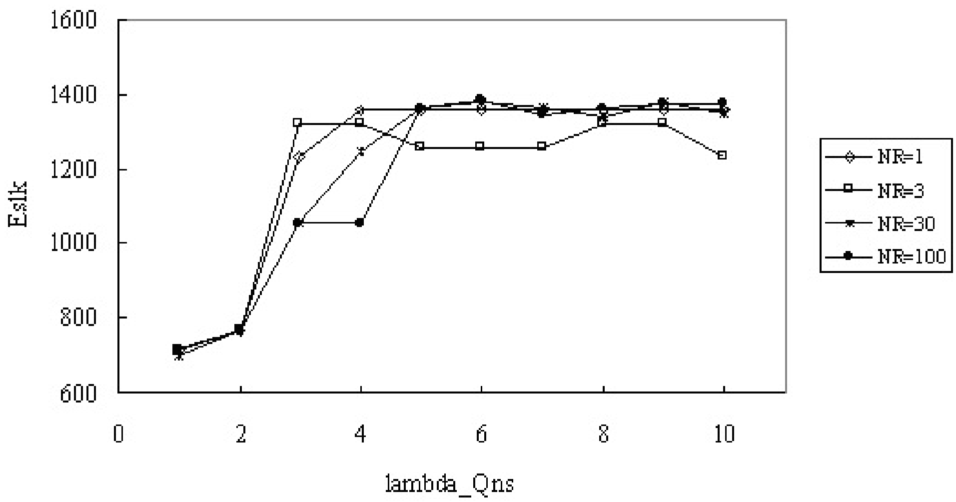

Figure 3: Variation of the expected value for capacity slacks, $Eslk(L)$, with the evolution of penalty values, λqns, for various numerical examples.

Accordingly with the figures 1 to 5, two key subjects are noticed:

i) the similarity of the behavior of the numerical instances when different number of scenarios is considered, *NR* from 1 to 100; and

ii) the model robustness, with adequate sensitivity of technical estimators to the evolution of the non-satisfied demand penalization parameter, λqns.

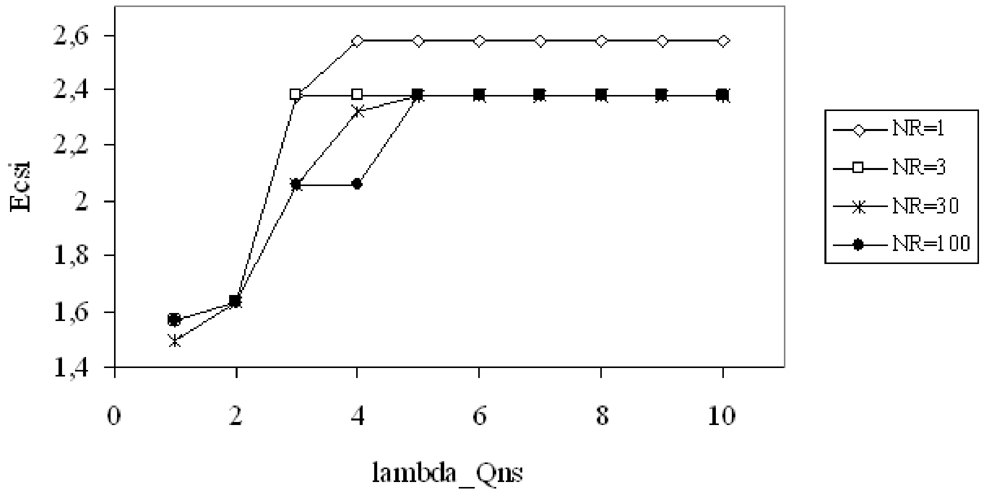

Figure 4: Variation of the expected value for investment costs, *Ecsi* (10^5 €), with the evolution of penalty values, λqns, for various numerical examples.

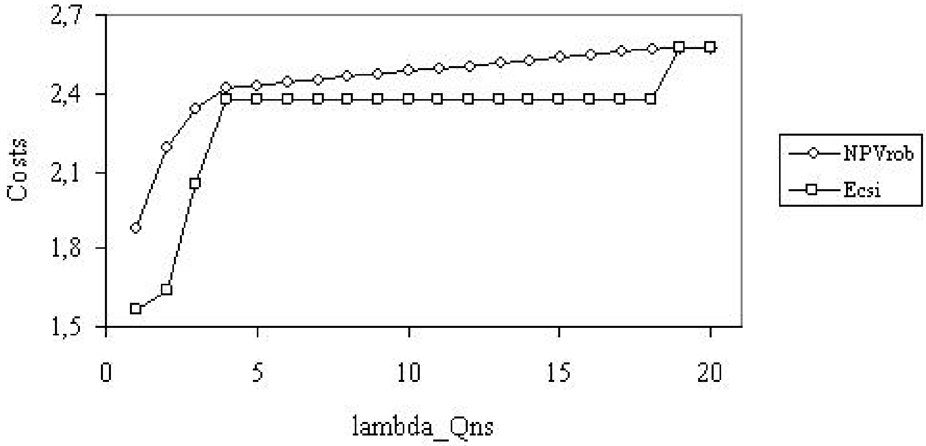

Figure 5: Variation of robust ($NPVrob$) and non-robust (*Ecsi*) costs (x10^5), with the evolution of penalty values, λqns (for $NR = 15$).

4.2. Focusing Robustness in Capacity Slack

In the following tables, Table 2 and Table 3, significant values are shown for instances of EX3, in which instances are considered with seven scenarios, but whose variation with the growth of the penalization on the non-satisfied demand, λqns, is similar to all the other examples, with different numbers of scenarios. These significant values are the values that are associated to alterations in the optimum configuration.

In Table 2 the expected values (robust and non-robust costs, and *Edvt*, *Ensd*, *Eslk*) of interest are shown, proving numerically the type of the evolutions observed in the previous graphs. Similarly to the expected value for the variability deviations, *Edvt*, the theoretical prediction is verified numerically: null return values, $ret_{jtr} = 0$, imply that the NPV to be maximized corresponds to the minimization of investment costs, which depends only on the binary variables.

Such order of results does not prove the robustness in the solution, but invariability is verified in this case: the equipment dimensions of the processes being defined in the first phase, the quantities produced will always be as great as possible, as the non-satisfied demand is penalized, which is equivalent to promoting production.

The prompt variation in the expected values of the non-satisfied demand, *Ensd*, with the increase of the respective penalty parameter, even making null this estimator, permits to state the model robustness.

Table 2: Significant values of the stochastic optimization, considering the evolution of penalty values and distinct instances of numerical example EX3 (for $NR = 7$).

λqns	**Costs (rob)**	**Costs**	***Edvt***	***Ensd***	***Eslk***
0.	124596.08	124596.08	0.0	153052.69	725.64
1.	188885.36	156737.60	0.0	32147.76	712.72
2.	220447.78	163661.91	0.0	28392.93	767.74
3.	234553.01	205467.00	0.0	9695.34	1042.93
4.	241546.56	237689.36	0.0	964.30	1404.58
5.	242510.86	237689.36	0.0	964.30	1233.89
	(...)		(...)		(...)
10.	247332.35	237689.36	0.0	964.30	1359.57
	(...)		(...)		
20.	256975.34	237689.36	0.0	964.30	1305.95
21.	257332.14	257332.14	0.0	0.0	1448.17
	(...)		(...)		(...)
40.	257332.14	257332.14	0.0	0.0	1400.87

Similarly to the capacity slacks, it should be noted that the *Eslk* estimator contains a character of permanence, as the equipment underutilization is underlying this type of problem: the underutilization or a slack would not exist, in the context of the ZW policy, if and only if all the products would present equal values for the technical parameters (namely, the dimension factors, S_{ij}, in L/kg) in the different stages. This numerical information on the expected values is complemented with the values in Table 3, relating the penalty parameters λqns and the (non-robust) costs, showing:

- the order of the discrete dimension (size, *s*) selected in each stage, *Ord(s)*; for example, "1/ 3/ 4", indicates that the first dimension in the first stage was chosen, the third dimension in the second stage, and the fourth dimension in the third stage; these values come directly from the binary solution;
- the sum of the discrete dimensions or equipment volumes selected, *Sum(dv)*, being opportune for the analysis of the slacks, given the interest of estimating *Eslk* in relative terms for each example;
- the percentage of the expected value of the non-satisfied demand, *%Ensd*, while 100 is the expected value of the uncertain demand in each example;
- the percentage of the expected value of the capacity slacks, *%Eslk*, the sum of volumes being the basis of the calculation of this estimator, *Sum(dv)*, in each example.

5. Conclusions

The model *spbatch_ms* is generalized from the deterministic model *MS*, featuring a two-stage stochastic framework with promotion of robustness. The generalized model simultaneously treats the scheduling of the production cycle embedded in the problem of design of batch processes, considering multiproduct environment (flowshop), multiple processes in each stage, SPC production and ZW storage policies.

The combination of SPC with multiple machines was found to be the most promising from a computational standpoint (Miranda, 2011*b*), and the model *MS* was here generalized toward a stochastic environment within the relaxation of the soft constraints regarding uncertain demand. Even for a reduced number of scenarios, results point to a significant reduction (8-20 %) on the investment costs in comparison to the deterministic non-relaxed case. If the MPC policy is adopted or if a slight relaxation is made to the impositions on the uncertain demands (respectively, of 1 % to 6 %), the demand relaxation does not cause real losses.

Beyond this significant decrease in terms of investment costs, other conclusions are obtained. Analyzing the variation of the number of scenarios, a similar behavior was observed for the various estimators. This similarity for the different numbers of scenarios (from 1 to 100) is to be remarked due to their impact in the execution times.

The robustness of the model *spbatch_ms* was also observed, with estimators responding adequately to the variation of the penalty parameter for non-satisfied demand. Also, the invariability of the configurations is mainly due to the realistic presupposition of discrete volumes. These results motivate to further continue this approach, which is integrating other developments under way.

Acknowledgements

This work was supported in part by CPQ ("Centro de Processos Químicos", Centre for Chemical Processes), IST, Technical University of Lisbon, Portugal. We thank the College of Technology and Management at the Polytechnic Institute of Portalegre and CIIST (IST Computing Centre).

Nomenclature

Index and sets

M	– number of stages *i*;
NC	– number of components or products *j*;
NP(*i*)	– (cardinal) number of processes *p*(*i*) per stage;
NR	– number of discrete scenarios *r*;
NS(*i*)	– (cardinal) number of discrete dimensions *s*(*i*) in the process of stage *i*;
NT	– number of time periods *t*;

Parameters

τ	– processing times (h), for each product j in stage i;
λdvt	– Negative deviation on NPV penalization parameter;
λqns	– non-satisfied demand penalization parameter;
λslk	– capacity slack penalization parameter;
c	– equipment cost related to process $p(i)$ and size $s(i)$ selected in stage i;
dv	– discrete equipment volume in each stage;
H	– time horizon;
$ncUpp$	– upper limit for disaggregated number of batches;
$prob$	– probability of scenario r;
$p(i)$	– (ordinal) number of processes in stage i;
Q	– demand quantities (uncertain) for each product i;
ret	– unit (uncertain) values of return (net values) of the products j, in period t and scenario r;
$s(i)$	– (ordinal) number of process discrete dimensions in stage i;
S	– dimension factor (L/kg), for each product j in stage i;
V	– equipment volume (continuous value) in each stage;

Variables

ξ	– NPV value in scenario r;
$dvtn$	– negative deviation on the value of NPV in scenario r;
n	– number of batches of product j, in period t and scenario r;
nc	– number of batches of product j, in period t and scenario r, disaggregated by process $p(i)$ and size $s(i)$ in each stage i;
Qns	– non-satisfied demand quantities of product j, in period t and scenario r;
slk	– capacity slacks in each stage i, concerning totality of the *batches* of each product j, in period t and scenario r;
$tcamp$	– campaign times (SPC) relative to each product j;
W	– global quantities produced of product j, in period t and scenario r;
y	– binary decision related to process $p(i)$ and size $s(i)$ selected in stage i;

Glossary

MPC	– multiple product campaign;
MS	– deterministic model focusing multiple processes and SPC;
spbatch_ms	– stochastic MILP model focusing multiple processes and SPC;
SPC	– single product campaign;
ZW	– zero wait storage policy;

6. Appendix A: Technical estimators

Non-robust NPV expected value:

$$Ecsi = \sum_{r=1}^{NR} prob_r \xi_r \tag{6.1}$$

Negative deviation expected value:

$$Edvt = \sum_{r=1}^{NR} prob_r.dvtn_r \tag{6.2}$$

Non-satisfied demand expected value:

$$Ensd = \sum_{r=1}^{NR} \frac{prob_r}{NC.NT} \left(\sum_{j=1}^{NC} \sum_{t=1}^{NT} Qns_{jtr} \right) \tag{6.3}$$

Capacity slack expected value:

$$Eslk = \sum_{r=1}^{NR} \frac{prob_r}{M.NC.NT} \sum_{j=1}^{NC} \sum_{t=1}^{NT} \left\{ \sum_{i=1}^{M} \sum_{s=1}^{NS} \sum_{p=1}^{NP} p(i).y_{isp}.\left(dv_{js} - S_{ij}.\frac{W_{jtr}}{n_{jtr}} \right) \right\} \tag{6.4}$$

Percent non-satisfied demand expected value:

$$\%Ensd = \frac{Ensd}{Qmed}.100, \text{ with } \ Qmed = \sum_{r=1}^{NR} \frac{prob_r}{NC.NT} \left(\sum_{j=1}^{NC} \sum_{t=1}^{NT} Q_{jtr} \right) \tag{6.5}$$

Percent capacity slack expected value:

$$\%Eslk = \frac{Eslk}{Vtotal}.100, \text{ with } Vtotal = \sum_{i=1}^{M} \sum_{s=1}^{NS} \sum_{p=1}^{NP} \left(y_{isp}.dv_{is} \right) \tag{6.6}$$

7. Appendix B: Aspects of the generalized model

The generalization approach at hand foresees a dynamic treatment of the multiperiod horizon, within the uncertainty considered in the NPV maximization. The uncertainty in product demands, Q_{jtr}, is modeled through discrete scenarios. If full satisfaction of demand is required, the equipment sizing may be directed for scenarios of low probability but requiring large dimensions. Thus, the production flows, W_{jtr}, are defined through soft constraints, simultaneously with the definition of non-satisfied demand, Qns_{jtr}, which is penalized in the robust objective function:

$$W_{jtr} \le Q_{jtr} \Rightarrow W_{jtr} + Qns_{jtr} = Q_{jtr}, \quad \forall j, t, r. \tag{7.1}$$

When the production flows variable, W_{jtr}, is used instead of the parameter on uncertain demand, Q_{jtr}, non-linearities occur in the constraints defining the number of batches, n_{jtr},

$$W_{jtr}.y_{isp} = Waux_{ijsptr}, \quad \forall i, j, s, p, t, r, \tag{7.2}$$

would lead to the simultaneous consideration of the following three sets of constraints:

$$Waux_{ijsptr} \le Waux_{ijsp}^{Upp} y_{isp}, \ \forall\, i, j, s, p, t, r, \tag{7.3}$$

$$\sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} y_{isp} = 1, \quad \forall i, \tag{7.4}$$

$$\sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} Waux_{ijsptr} = W_{jtr}, \quad \forall i, j, t, r. \tag{7.5}$$

Thus, the number of variables *(M.N.NS.NP.NT.NR)* and the number of constraints *(M.N.NS.NP.NT.NR + M.N.NT.NR)* would be highly increased, and a coincidence would also occur with the usual aggregated variable:

$$n_{jtr}.y_{isp} = nc_{ijsptr}, \quad \forall i, j, s, p, t, r. \tag{7.6}$$

Now, using the original formulation (Kocis & Grossmann, 1988), which had batch size, B_{jtr}, associated with the equipment volume dimension, V_i, and if production flows, W_{jtr}, are further formulated in association with discrete volumes, dv_{is}, it follows:

$$\begin{aligned} &S_{ij}B_{jtr} \le V_i, \quad \forall i, j, t, r \quad \Rightarrow S_{ij}B_{jtr} \le \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} dv_{is}y_{isp} \Rightarrow \\ &S_{ij}\frac{W_{jtr}}{n_{jtr}} \le \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} dv_{is}y_{isp} \Rightarrow S_{ij}W_{jtr} \le \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} dv_{is}y_{isp}n_{jtr} \Rightarrow \\ &S_{ij}W_{jtr} \le \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} dv_{is}nc_{ijsptr}, \quad \forall i, j, t, r. \end{aligned} \tag{7.7}$$

Consequently, the semicontinuous variables, nc_{ijsptr}, are required, and the auxiliary variables, $Waux_{ijsptr}$, and the inherent constraints are avoided.

Furthermore, the capacity slacks must be penalized in the robust model. The constraint slacks, slk_{ijtr}, occur in each stage and they are related to the global quantity produced, W_{jtr}, in all time periods and discrete scenarios,

$$S_{ij}W_{jtr} + slk_{ijtr} = \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} dv_{is}nc_{ijsptr}, \quad \forall i, j, t, r. \tag{7.8}$$

These constraint slacks, slk_{ijtr}, correspond to the aggregated value of the slacks that are occurring in all of the production cycles. Such variables can be related to the effective slack that occurs in each one of the started batches, Zp_{ijtr},

$$\begin{aligned} &S_{ij}\frac{W_{jtr}}{n_{jtr}} + \frac{slk_{ijtr}}{n_{jtr}} = \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} dv_{is}y_{isp} \Rightarrow \\ &S_{ij}B_{jtr} + Zp_{ijtr} = \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} dv_{is}y_{isp}. \end{aligned} \tag{7.9}$$

The effective slack verified in the selected equipment, Zp_{ijtr}, can be obtained through the disaggregation of the global slack, slk_{ijtr}, in association with the number of started batches, n_{jtr}:

$$Zp_{ijtr} = \frac{slk_{ijtr}}{n_{jtr}}, \quad \forall i, j, t, r. \tag{7.10}$$

This non-linear relation between the two slack variables forbids the consideration of the effective capacity slack, Zp_{ijtr}, in the robust objective function. Nevertheless, this variable can be used to assess the slacks verified in the selected configuration of equipment volumes.

8. Appendix C: Reduction to Static *MS* Problem

The computational complexity of the model *spbatch_ms* is studied, as follows:

- if one supposes $NT = NR = 1$, with $prob(1) = 1$, then the time t and random r index can be neglected;
- provided that $NR = 1$, the linear negative deviation is null,

$$dvtn_r = \sum_{r'=1}^{1} (prob_{r'}\xi_{r'}) - \xi_r = 0; \tag{8.1}$$

- assuming $\lambda qns = bigM$, the optimization resolution will lead to

$$Qns_i = 0. \Rightarrow W_i = Q_i \ , \quad \forall i; \tag{8.2}$$

- assuming $\lambda slk = 0.$ (i.e., do not penalize capacity slack) and $ret_{jtr} = 0.$ (do not account return or cash flows), the objective function presents the following pattern of minimization of the investment costs:

$$\max \ \Phi = \max \ \xi_r = \max \left\{ -\sum_{j=1}^{M} \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} c_{isp} y_{isp} \right\} = \min \left\{ \sum_{j=1}^{M} \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} c_{isp} y_{isp} \right\}; \tag{8.3}$$

- the remaining constraint sets are equivalent to the constraint sets of *MS*; it follows, by instance reduction, that if it is proved that *MS* problem belongs to NP-Hard class, it also drives the computational complexity of *spbatch_ms* problem.

Consequently, the study of the deterministic and static *MS* problem (a single time period) presented characteristics that remain valid in all of the discrete scenarios and time periods considered in the stochastic and multiperiod *spbatch_ms* problem:

- from constraint set 3.10, in all stages i, one and only one of discrete dimension *s'(i)* and number of machines *p'(i)* can be selected,

$$\begin{cases} y_{isp} = 0, & \forall \quad s \neq s'(i), \quad p \neq p'(i) \\ y_{i,s'(i),p'(i)} = 1, & \exists^1 \ (s'(i), p'(i)) \end{cases} \tag{8.4}$$

- by conjugation of constraint sets 3.8 and 3.9, the nc_{ijsp} variables present zero value, with the exception of the ones associated to the selected discrete dimension *s'(i)* and number of machines *p'(i)*, thus the value presented by n_j is equal in all stages; further, the trivial satisfaction of the upper limit in logic constraint 3.7 ($nc^{Upp} = bigM$) is supposed and it follows that

$$\begin{cases} nc_{ijsp} = 0, & \forall \quad s \neq s'(i), \quad p \neq p'(i) \\ nc_{ij,s'(i),p'(i)} = n_j, & \exists^1 \ (s'(i), p'(i)) \end{cases} \tag{8.5}$$

- from constraint set 3.7 it is observed that, for each product j, the number of batch n_j definition must be satisfied in all stages i; thus, the n_j value obtained corresponds to the maximum of the values assessed through the i index,

$$S_{ij} W_j \leq p'(i) dv_{i,s'(i)} n_j, \quad \forall i, j \Rightarrow n_j = \max_i \left\{ \frac{S_{ij} W_j}{p'(i) dv_{i,s'(i)}} \right\}, \ \forall i \tag{8.6}$$

- in constraint set 3.11, the time campaign $tcamp_j$ for each product j must be satisfied in each stage i; it thus corresponds to the maximum value of the sum formulated in this constraint set; in the sum, the number of production cycles nc_{ijsp} is limited by the nj value in all stages i, as expressed by constraint sets 3.8 and 3.10,

$$tcamp_j = \max_i \left\{ \frac{\tau_{ij}}{p'(i)} n_j \right\} = \max_i \left\{ \frac{\tau_{ij} S_{ij} W_j}{p'(i) dv_{i,s'(i)}} \right\}, \quad \forall j \tag{8.7}$$

- in constraint set 12, the time campaign $tcamp_j$ sum must satisfy the time horizon H. Then it follows that

$$\sum_{j=1}^{NC} \max_i \left\{ \frac{\tau_{ij} S_{ij} W_j}{p'(i) dv_{i,s'(i)}} \right\} \leq H \tag{8.8}$$

In conclusion, the robust objective of maximizing NPV must be observed. Thus the equipment configuration (s', p') must be selected satisfying the constraint sets in all stages i and for all products j, given all discrete scenarios in all time periods. This way, the production flow W_{jtr} is associated to the return (cash) flows. The decrease of solution variability is foreseen through the penalization of the negative deviation on NPV, and model robustness is promoted through the penalization of capacity slacks.

9. Appendix D: Data generation for numerical examples

Comprehensively and in GAMS environment, the generation of random data for numerical examples is specified through the following lines of code.

```
OPTION SEED = 08012007
H = UNIFORM (6000, 8000);
alpha(j)= UNIFORM (300, 700); /*cost function coefficient */
beta(j) = UNIFORM (0.5, 0.7); /*cost function exponent */
LOOP (j,
    dv(j,"1") = UNIFORM(800,1300);
    dv(j,"2") = UNIFORM(1500,2000);
    dv(j,"3") = UNIFORM(2200,2700);
    dv(j,"4") = UNIFORM(2800,3300);
    dv(j,"5") = UNIFORM(3500,4000);
);
c(j,s) = alpha(j)*dv(j,s)**beta(j);
S(i,j) = UNIFORM (1,5);
τ(i,j) = UNIFORM (2,9);
Q_med = 150000; Q_dsv = Q_med / 6.;
Q(i,t,r) = NORMAL (Q_med, Q_dsv);
```

References

Ahmed, S. and N. V. Sahinidis (2000). Analytical investigations of the process planning problem. *Computers and Chemical Engineering* **23**, 1605– 1625.

Ahmed, S. and N. V. Sahinidis (2003). An approximation scheme for stochastic integer programs arising in capacity expansion. *Operations Research* **53**(3), 461–471.

Barbosa-Póvoa, A.P. (2007). A critical review on the design and retrofit of batch plants. *Computers and Chemical Engineering* (31), 833–855.

Birewar, D.B. and I. E. Grossmann (1989). Efficient optimization algorithms for zero-wait scheduling of multiproduct batch plants. *Ind. Eng. Chem. Res.* (28), 1333–1345.

Cavin, L., U. Fischer, F. Glover and K. Hungerbhler (2004). Multi-objective process design in multi-purpose batch plants using a tabu search optimization algorithm. *Computers and Chemical Engineering* (28), 459–478.

Floudas, C.A. and X. Lin (2004). Continuous-time versus discrete-time approaches for scheduling of chemical processes: a review. *Computers and Chemical Engineering* (28), 2109–2129.

Jayaraman, V. K., B. D. Kulkarni, S. Karale and P. Shelokar (2000). Ant colony framework for optimal design and scheduling of batch plants. *Computers and Chemical Engineering* (24), 1901–1912.

Kocis, G.R. and I.E. Grossmann (1988). Global optimization of nonconvex mixed-integer nonlinear programming (MINLP) problems in process synthesis. *Ind. Eng. Chem. Res.* (27), 1407–1421.

Liu, M.L. and N.V. Sahinidis (1997). Bridging the gap between heuristics and optimization: capacity expansion case. *AIChE J.* (43), 2289–2299.

Miranda, J. L. (2007). Optimização em sistemas de processos químicos: Generalização de modelos com planeamento e sequenciamento. PhD dissertation, Instituto Superior Técnico, Technical University of Lisbon, Lisboa.

Miranda, J. L. (2011*a*). Computational complexity studies in the design and scheduling of batch processes.. In: *Book of Abstracts of IO2011 - 15th Congress of APDIO*. University of Coimbra, Coimbra, Portugal.

Miranda, J.L. (2011*b*). The design and scheduling of chemical batch processes: building heuristics and probabilistic analysis. *Theory and Applications of Mathematics & Computer Science* **1**(1), 45–62.

Moreno, M.S., J.M. Montagna and O.A. Iribarren (2007). Multiperiod optimization for the design and planning of multiproduct batch plants. *Computers and Chemical Engineering* (31), 1159–1173.

Pekny, J. F. and D. L. Miller (1991). Exact solution of the no-wait flowshop scheduling problem with a comparison to heuristic methods. *Computers and Chemical Engineering* **15**(11), 741–748.

Tan, S. and R. S. H. Mah (1998). Evolutionary design of noncontinuous plants. *Computers and Chemical Engineering* **22**(1/2), 69–85.

Voudouris, V. T. and I. E. Grossmann (1992). Mixed-integer linear programming reformulations for batch processes design with discrete equipment sizes. *Industrial and Engineering Chemistry Research* (31), 1315–1325.

Voudouris, V. T. and I. E. Grossmann (1993). Optimal synthesis of multiproduct batch plants with cyclic scheduling and inventory considerations. *Industrial and Engineering Chemistry Research* (32), 1962–1980.

Voudouris, V. T. and I. E. Grossmann (1996). Milp model for scheduling and design of a special class of multipurpose batch plants. *Computers and Chemical Engineering* (20), 1335–1360.

Xia, Q. and S. Macchietto (1997). Design and synthesis of batch plants - minlp solution based on a stochastic method. *Computers and Chemical Engineering* **Suppl.**(21), S697–S702.