THE ROLE OF PROCESS VARIABLES IN THE DESIGN OF MULTIPRODUCT BATCH PROTEIN PRODUCTION PLANTS

Jorge M. Montagna [(1)], Aldo R. Vecchietti [(1)], Oscar A. Iribarren [(1)(*)]
José M. Pinto [(2)] and Juan A. Asenjo [(3)]
(1) Institute for Development and Design Ingar, Conicet, Santa Fe, Argentina
(2) Department of Chemical Engineering, Univ. of Sao Paulo, Sao Paulo, Brazil
(3) *Center for Biochem Eng and Biotechnology, Univ. of Chile, Santiago, Chile*

Keywords: Multiproduct batch plant design, process performance models, design variables.

Abstract

This work reports findings about the role of process variables in the design of multiproduct batch plants. Unlike continuous processes, batch processes are subject to size and time constraints which depend on the structure of the plant: the number of units at each stage and the provision of intermediate storage. We used simple process performance models (yet involving all the process variables with significant economic impact) to get explicit expressions for these size and time factors. The traditional approach uses fixed size and time factors. So the addition of those expressions to the original fixed factors model, permitted to simultaneously optimize the plant structure and process variables, and study the role of the latter in the design.

We found that if the plant structure constraints are disregarded (with a Free Unlimited Storage operating policy), process variables behave just alike in continuous processes. They trade off cost components with the Total Annual Cost being quite insensitive to them in the neighborhood of the optimal solution. As setting the process variables sets the size and time factors, this means that near the optimal set of process variables, cycle times and size factors can be accommodated to the plant structure, with little effect on the cost of equipment.

1. Introduction

While the constant time and size factors model (Biegler *et al*, 1997) is the most widespread used to model multiproduct batch processes, the process performance models (Salomone *et al*., 1992) describe these time and size factors as functions of the process variables selected as optimization variables. Thus the same mathematical model for the plant is used in both approaches, with the process performance models as additional constraints in the second case.

The process performance models are obtained from the mass balances and kinetic expressions that describe each unit operation. They are kept as simple as possible still retaining the influence of the process variables selected to optimize the plant.

In the first part of this paper, we briefly describe the processes and the performance models used. Then we study the influence of the process variables on the design and operation of this process.

Multiproduct batch processes are size and time constrained, with these constraints depending on the structure of the plant. This makes the optimization of process variables be nested. Setting a value for the process variables sets the mass balances, which in turn sets the recipe of the process, and these recipes, in the form of fixed size and time factors were usually taken as input to optimize the plant structure.

Using a hierarchical approach (Douglas, 1988), we first optimize process variables disregarding the plant structure using Single Product – Free Intermediate Storage scenarios. We show that disregarding the plant structure constraints, process variables behave as in continuous processes trading off cost components with a smooth dependence of the Total Annual Cost on the process variables.

2. Process description

Figure 1 shows the flowsheet of a multiproduct batch plant for the production of Human Insulin, Vaccine for Hepatitis B, Chimosin and Cryophilic Protease, produced by saccharomices cerevisiae.

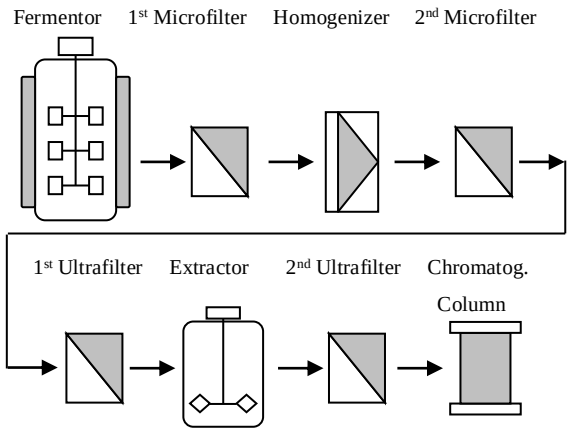


Fig. 1. Flowsheet for proteins production plant

All metabolites are produced as the cells grow in the Fermentor. Due to the Vaccine and Protease being intracellular, the first Microfilter concentrate the cell suspension, which is afterwards sent to the Homogenizer for cell wall disruption to liberate the metabolites. The second Microfilter remove the cell debris from the solution of proteins.

The ultrafiltration prior to the Extractor is used for concentrating the solutions in order to minimize the Extractor volume. In the liquid-liquid extraction, salt concentration is manipulated to first drive the product to a Polyethylene Glycol Phase (PEG) and back again

* Author to whom correspondence should be addressed.

into an aqueous phosphate solution. In this process, many of the proteins other than products are removed. Ultrafiltration is used again for concentrating the solution, and finally the last stage is a chromatography where selective binding is used to further separate the product of interest from other proteins.

Insulin and Chimosin are extracelular. They are in the permeate that crosses the filtration membrane of the first Microfilter. In order to reduce the amount of valuable product lost in the retentate, extra water is added to the cell suspension. The filtration operation with make up of water is also called diafiltration and dilutes the solution of proteins. They skip the Homogenizer and Microfilter for cell debris removal, but then the Ultrafilter is necessary to concentrate the dilute solution prior to extraction. The final steps of extraction, ultrafiltration and chromatography are common with the intracellular products.

Insulin and Vaccine are commercial products. The plant would produce the technical grade products with further purification steps. Otherwise, Chimosin and the Protease are newer products that could be made with part of the plant shown in Figure 1. While there is enough information about the Chimosin, the Cryophilic Protease is still in its development stage and most of the process information is estimated.

3. Fixed Factors Model

The general batch process literature as in Biegler *et al* (1997) describes batch plants through size and time equations. For batch stages we use:

$$V_j \geq S_{ij} \quad B_i \tag{1}$$

$$T_{ij} = T_{ij}^0 + T_{ij}^1 \quad B_i \tag{2}$$

where V_j is the size of stage j [m^3], B_i is the batch size for product i,[kg of product exiting the last stage] and S_{ij} is the size factor at stage j to produce 1 kg of final product i. T_{ij} is the time required at stage j to process a batch of product i. T_{ij}^0 is a time factor that accounts for fixed amounts of time while T_{ij}^1 permits to account for time demands proportional to the batch size to be processed. For semicontinuous units we use:

$$R_j \geq D_{ij} \quad \frac{B_i}{\theta_j} \tag{3}$$

where R_j is the size of the semicontinuous item j, usually a processing rate as in the case of the homogenizer capacity [m^3/h], but in the case of the filters the size is the filtration area A [m^2]. D_{ij} is the duty factor. In the case of composite stages with a semicontinuous item that processes the material hold in a batch item (as in the case of the Homogenizer) we follow the modeling approach in Salomone *et al* (1994). The stage is described with equation (1) for the batch item size, but the batch processing time T_{ij} includes the operating time θ_j of the semicontinuous item, so replacing θ_j from equation (3) into equation(2) gives :

$$T_{ij} = T_{ij}^0 + D_{ij} \quad \frac{B_i}{R_j} \tag{4}$$

If the size and time factors S_{ij}, D_{ij}, T_{ij}^0 and T_{ij}^1 in equations (1) to (4) are constants, this gives rise to a geometric model for the process (Asenjo *et al*, 1999). To get these constant factors it is necessary to guess or estimate a value for every process variable so as to cover the degrees of freedom of the process mass balances.

4. Process Performance Models

In the approach of this paper we use process performance models as simple as possible, still retaining the influence of the process variables that we a priori expect to have the largest impact on the economics of the process, in line with Douglas (1988).

Once these variables have been selected, we write the mass balances and kinetic equations that describe each stage guessing or estimating values for every non-selected process variable, but not for the chosen optimization variables. As a result, we get analytical expressions for the size and time factors that will be functions of these process variables.

The mathematical optimization model for the design of the multiproduct batch plant will be exactly the same as the one with constant size and time factors, plus the additional constraints that describe these factors as functions of the process variables.

The process variables that have been selected as optimization variables are: the biomass concentration at the Fermentor (X_{fer}) and Microfilter 1 (X_{mf1}) for all products, the volumetric ratio of diafiltration water to suspension feed at Microfilter 1 (W_{mf1}) for extracelular Insulin and Chimosin and at Microfilter 2 (W_{mf2}) for intracellular Vaccine and Protease after cell disruption, the number of passes through the Homogenizer (N_p) for intracellular Vaccine and Protease, and the volumetric ratio (R) of PEG to Phosphate phases at the Extractor for all products.

Following, is a brief description of the process performance models for the Fermentor as a typical batch stage, and the Homogenizer as a typical combined batch - semicontinuous stage. Most of the information needed to develop them was taken from Asenjo (1990) and Belter *et al* (1988). A more detailed description can be found in Asenjo *et al* (1999).

4.1 Fermentor

A Monod like kinetics constrained by a maximum biomass concentration is assumed for cell grow:

$$\frac{dX}{dt} = \phi\, X \left(1 - \frac{X}{X_{max}}\right) \tag{5}$$

We estimated the same kinetic constant ϕ and maximum biomass concentration X_{max} for all products. The batch size at the Fermentor is related to the Fermentor size through the biomass concentration:

$$B_i^{fer} = 0.8\, V_{fer}\, X_{fer,i}\, k_i \tag{6}$$

that assumes that the batch volume occupies 80% of the vessel, and k_i is a stoichiometric ratio [kg of product i / kg of biomass] which is 0.08 for Protease, 0.06 for Chimosin, 0.02 for Insulin, and 0.04 for

Vaccine. Then, it must be taken into account that the batch size at any stage j is related to the batch size exiting the plant through the yields of all stages between this particular stage and the exit from the plant:

$$B_i = B_i^j \prod_n \eta_n \qquad n = j, j+1,...M \tag{7}$$

where M is the number of stages of the plant.
Integrating equation (5) between an initial biomass concentration 0.05 X_{max} (inoculum seeded amounts to a 5% of the fermentor capacity) and X_{fer} and adding an estimated downtime of 4 hr, gives:

$$T_i[hr] = 4 + 3.8\, \ell n\left[\frac{0.35\, X_{fer,i}}{\left(1 - \frac{X_{fer,i}}{55.}\right)}\right] \tag{8}$$

which is the same for all four products.

Replacing (7) into (6), gives the S_i expressions for the Fermentor. For Insulin and Chimosin:

$$S_i\left[m^3 / kg\right] = \frac{1.25}{k_i X_{fer,i} \eta_{m1,i} \eta_{ext,i} \eta_{chr,i}} \tag{9}$$

that contains the yields different from one.

4.2 Homogenizer

The holding vessel capacity corresponds to the final volume in the retentate vessel of Microfilter 1. So the S_i expression for intracellular products is:

$$S_i\left[m^3 / kg\right] = \frac{1.25}{k_i X_{mf1,i} \eta_{hom,i} \eta_{mf2,i} \eta_{ext,i} \eta_{chr,i}} \tag{10}$$

The denominator contains the yields different from one, for intracellular products. The time is proportional to the volume fed to the homogenizer V_{hom} [m^3] and inversely proportional to the Homogenizer capacity Cap [m^3/h] plus a downtime:

$$T_{hom,i} = T^0_{hom,i} + \frac{V_{hom}}{Cap} \tag{11}$$

The volume fed to the Homogenizer is the batch volume times the N_p, and estimating a 1.25 h downtime gives the time for the homogenizer:

$$T_i = 1.25 + \left[\frac{1.\quad N_{p,i}}{k_i X_{mf1,i} \eta_{hom,i} \eta_{m2,i} \eta_{xt,i} \eta_{chr,i}}\right] \frac{B_i}{Cap} \tag{12}$$

Successive passes through the Homogenizer drive the fraction of cells disrupted asinthotically to 1. This is also the fraction of proteins released F_r:

$$F_{r,i} = 1 - exp(-k_1 N_{p,i}) \tag{13}$$

but the same law works to estimate the fraction of released denatured proteins F_d:

$$F_{d,i} = 1 - exp(-k_2 N_{p,i}) \tag{14}$$

with k_1 larger than k_2 because larger particles are more easily disrupted (k_1=1.5 and k_2=0.03). The yield of the homogenizer is the fraction released times the fraction not denatured:

$$\eta_{hom,i} = F_{r,i}\,(1 - F_{d,i}) \tag{15}$$

Replacing (13) and (14) into (15) gives the yield:

$$\eta_{hom,i} = \left[1 - exp\left(-1.5\, N_{p,i}\right)\right] exp\left(-0.03\, N_{p,i}\right) \tag{16}$$

5. The Single Product – Free Intermediate Storage Scenario (SP-FISS)

The model for the multiproduct batch plant consists of the traditional geometric program where the size and time factors are fixed, plus the process performance models that describe each of these factors as functions of the process variables. This is a Mixed Integer Non Linear Program, which lacks a definite structure, just because the nonlinear process performance models depend on the particular unit operations involved in the process at hand. The global optimization of this problem, which does have multiple local optima, is still an open problem (Floudas and Pardalos, 1999).

In this paper, we decompose the problem into hierarchical levels. First, we assign a value to process variables, i.e. we construct the recipe for the processes. Next, we use the resulting size and time factors as an initial point to optimize a structure for the plant. This assignment should be as unbiased as possible with respect to the plant structure, which has been left as a second level decision.

The optimization of the process variables in a SP-FISS fulfills this objective. We do an arbitrary partition of the annual operating time among the products that we expect to produce in the same plant and define a production rate Pr_i for each product:

$$Pr_i\left[kg/hr\right] = Q_i\left[kg\right] / H_i\left[hr\right] \tag{17}$$

where Q_i are the annual target production and H_i the time horizons assigned to each product i

The most expensive stage is fermentation, so a reasonable partition of the total horizon time should consider a similar Fermentor size requirement through the stoichiometric ratios k_i (Flatz, 1981). As in our case the biomass production step demands the same amount of time regardless from the protein being produced:

$$H_i = H \frac{Q_i / k_i}{\sum_i Q_i / k_i} \tag{18}$$

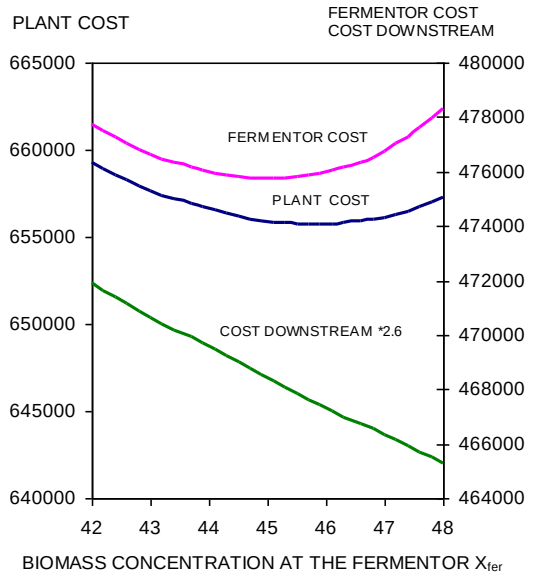
While the single product assumption permits to ignore the size constraints, the free intermediate storage assumption permits to ignore the time constraints so that each stage works with its own cycle time, starting a new cycle right after finishing one, satisfying the production rate assigned to the product.

We note that the level of storage that would be needed for this operation is not unlimited, but the decoupling level (Modi and Karimi, 1989) which would permit to simultaneously hold both the batch sizes entering and leaving the tank.

Following, we describe the role of the process variables of the protein production plant, in SP-FISS. We took Cryophilic Protease as example because, being intracellular, it goes through all the processing stages. We plotted the cost items versus each variable, with the rest of the variables at their optimal value.

6. Role of the Selected Variables

6.1 Biomass Concentration at the Fermentor X_{fer} (Fig. 2)



The X_{fer} increases monotonically with time, at a decreasing pace as it asymptotically approaches a maximum. Thus, the rate of production of biomass at the fermentor has a maximum at an intermediate value of 45.1 kg / m^3 in coincidence with the minimum cost for the fermentor. However, the cost of the downstream process decreases monotonically when increasing the concentration. As a result, the downstream shifts the overall optimum to a larger value of 45.8 kg / m^3

6.2 Biomass Concentration at Microfilter I X_{mf1} (Fig. 3)

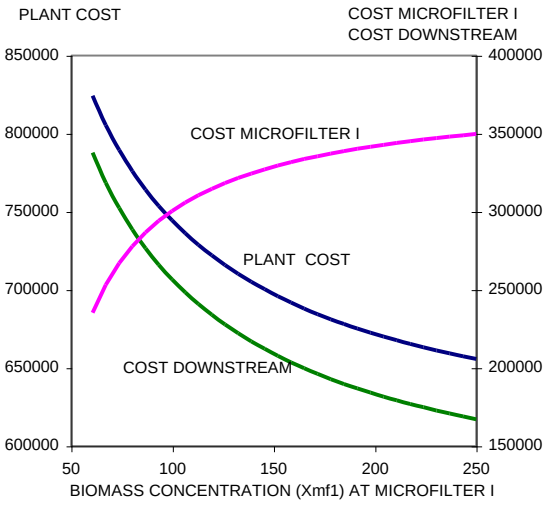


Fig.3. Tradeoff in the selection of biomass concentration at microfilter I

Larger concentrations require larger volumes of liquid to be permeated through the membrane. At a constant permeation rate, this requires more area and thus, an increased filter cost. On the other hand, both the Homogenizer and Microfilter II sizes are inversely proportional to this concentration, so their costs decrease monotonically. The optimal biomass concentration is at its upper bound of 250 kg / m^3

6.3 Number of Passes through the Homogenizer Np (Fig. 4)

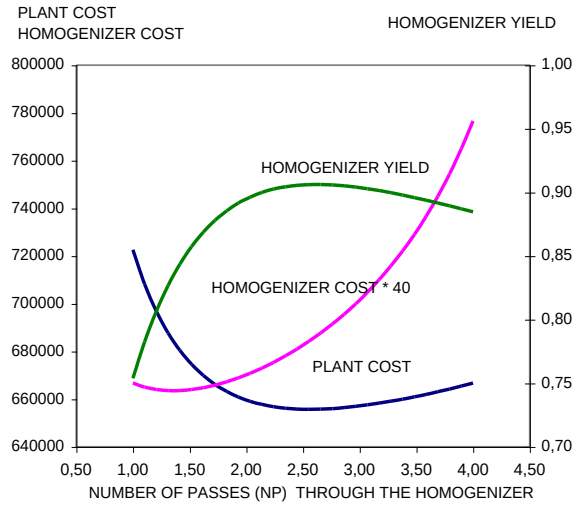


Fig.4. Tradeoff in the selection of number of passes (NP) through the Homogeneizer

Increasing Np increases the number of cells disrupted (protein released) but also increases the amount of released protein that is being denatured. As a result we have a maximum yield of product (released but not denatured, with respect to the total amount inside the cells before processing) at a number of passes of Np=2.65

The size of the Homogenizer is proportional to Np and inversely proportional to the yield. It has a minimum at Np = 1.35. However, the yield affects the whole plant (specially increasing the size required from the units upstream of the Homogenizer) and the optimal value for the plant is Np = 2.55 which is very close to the maximum yield.

6.4 Washing Water at Microfilter II Wmf2 (Fig. 5)

At Microfilter II the already released protein is recovered by diafiltration with distilled water. An increase in the amount of water increases the size of both Microfilter II itself and of Ultrafilter I whose job is to re-concentrate the diluted protein solution.

So the cost of Microfilter II and Ultrafilter I increase monotonically, as well as the yield of product at Microfilter II. The increase in yield decreases the size required from the upstream units: Fermentor, Microfilter I and Homogenizer. As a result, we have an overall optimum for the plant at a ratio of washing water to feed $W_{mf2} = 1.15$

6.5 Volumetric Ratio of PEG to Phosphate Phases R (Fig. 6)

Increasing R increases the yield of the first extraction from the Phosphate into the PEG phase, but

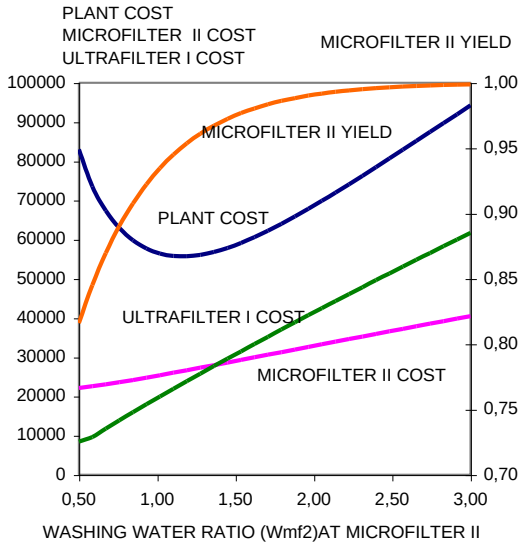


Fig.5. Tradeoff in the selection of washing water ratio at Microfilter II

decreases the yield of the back extraction into the new salt – free Phosphate phase. This is so because of the smaller amount of the new phase, and because of a poorer dilution of the ClNa remaining in the PEG, which jeopardizes the partition constant for this back extraction. As a result, the overall extraction yield has a maximum at R=0.63.

The cost of the Extractor grows linearly with R (R it is the volume of PEG phase to be added, per volume of the batch entering this stage) and inversely proportional to the extraction yield. As a consequence, the cost of the Extractor shows a minimum at a phase ratio of R=0.2 However the extraction yield also affects all the upstream stages size, so the overall process optimum is at R=0.61 close to the maximum yield.

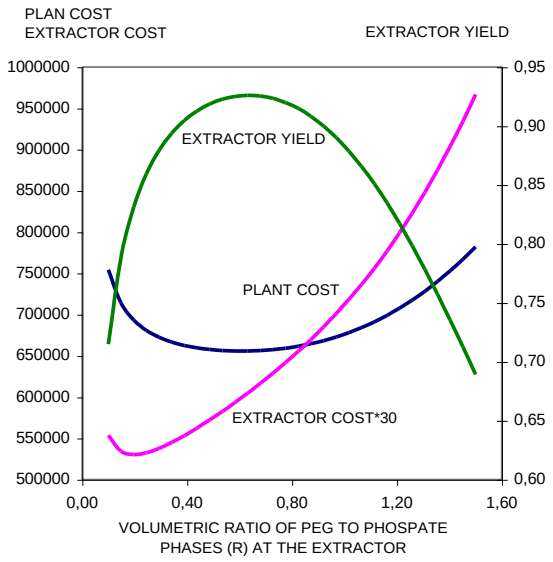


Fig 6. Tradeoff in the selection of the volumetric ratio of PEG to Phosphate phases (R) at the extractor

7. Analysis of the Tradeoffs that occur when setting the Process Variables

The descriptions just done about the economic tradeoffs involved in the assignment of a value to the process variables, very much resemble the ones presented in the hierarchical approach by Douglas (1988), for designing continuous processes. As a matter of fact the first attempts to extend this kind of analysis to batch processes were done by Malone and coworkers (Swami, 1985; Iribarren 1985), who used simplified sizing equations. This approach lacked a systematic linking with the geometric program model which incorporates the size and time constraints and so was very much limited to dedicated plants with storage at every batch – batch interface.

Barrera and Evans (1989) addressed the role of process variables in batch plants. They named tradeoffs of the first type to those that occur within a single stage, of the second type to those that occur between or among stages, and of the third type to those that are a combination of the first two types.

The nested algorithm that they proposed suffered feasibility problems because the size of equipment was fixed at an upper level while the process variables were fixed at the lower level. Then, very often a set of process variables was selected such that the resulting operating times did not satisfy the production target. Even so, the description of the tradeoffs is completely right for our case and we think, in general.

The X_{mf1} and the W_{mf2} are involved in tradeoffs of the second type, an increase in these variables increases the costs of these units, but decreases the cost of the units down or up stream respectively.

The X_{fer}, the N_p and the R are involved in tradeoffs of the third type. There is a particular value for these variables that produces a minimum cost of the respective stages, but they also affect the cost of other up or down stream units.

Bathia and Biegler (1996) pose the optimization of process variables as a large system of algebraic and differential equations. The handled processes consisted of a modest number of stages and products. Even so, large enough to show that optimizing the individual stages sequentially renders poor results as compared with simultaneous optimization. Barrera and Evans (1989) show that sequential optimization would only succeed with tradeoffs of the first type.

The shape of the overall cost in Figures 2 to 6 is enlightening. As the optimal X_{mf1} lies on its upper bound, this leads to fixing it at that value and deletes it from the list of optimization variables in any further step. In the other cases, when the variables produce a minimum cost at a value within its range, the shape of this overall costs are rather flat around the optima. That the gradient of the objective function be zero at this unconstrained optimum, is a condition. However,

the observation that it remains small over a quite large neighborhood of the optimum is very good news. For any particular structure of the plant, the feasible design and or operation will include the corresponding size and time constraints, and extra cost items in the objective function: the cost of storage tanks and the cost penalty of duplications. But a lot of the cost of a batch process is due to the idle times and volumetric under utilization of the process units. The possibility of moving the process variables (which in turn moves the size and time factors) over a wide range without a major penalty on the overall process unit cost, permits the size and time factors move to accommodate themselves to the imposed structure of the plant. These movements will be on the directions such to minimize idle times and volumetric under occupancy.

8.Conclusions

A process performance optimization model for the design of batch plants has been developed. The model is a MINLP having constraints for units sizing and discrete 0-1 variables for the structural design: parallel units (in-phase and out-of-phase) and storage tank locations. The model considers composite units (batch units combined with semicontinuous items considered here as a single stage).

We used the model to design a protein production plant. It is not trivial to estimate good constant time and size factors for the plant design. This difficulty is overcame by the proposed process performance models, which predict the size and time factors as a function of process variables. It is easier to guess good or reasonable values for the process variables than for time and size factors of the stages.

An even better set of initial factors was obtained by optimizing the process variables in SP-FISS. It is an unconstrained design problem, and, disregarding the plant structure constraints, process variables behave as in continuous processes trading off cost components with a smooth dependence of the Total Annual Cost on the process variables. This was done for each of the process performance variables considered. Afterwards, for any particular structure and relaxing the process variables, they move to accommodate the time and size factors reducing both idle times and under- utilization of equipment.

Probably, the major merit of the approach is its modular structure. The process performance models are additional constraints to the traditional geometric program, which remains unchanged. To set up higher level of detail models is highly facilitated.

The SP-FISS permits to get a very good set of values for the process variables, before solving the whole problem. Furthermore, they permit to get insight about the role of process variables, and delete those, which have not an important economic impact, or should be fixed at a bound.

Subscript or superscript

chr	Chromatography column
ext	Extractor
fer	Fermentor
hom	Homogenizer
mf1,2	Microfilter 1, 2
uf1,2	Ultrafilter 1,2

Aknowledgments. The authors are grateful for the financial support to Fundacion Vitae (Brazil) and Fundacion Antorchas(Argentina).

References

Asenjo J., J. Montagna, A. Vecchietti, O. Iribarren and J. Pinto “Optimization of Protein Production Plants with a Time and Size Factors Process Model” Biotechnology Progress 16, 228-237 (1999).

Asenjo Juan A. “Separation Processes in Biotechnology” Marcel Dekker New York (1990).

Barrera M. and L. B. Evans “Optimal Design and Operation of Batch Processes” Chem. Eng. Comm. 82, 45-66 (1989).

Bhatia T. and L. Biegler “Dynamic optimization in the design and scheduling of multiproduct batch plants” Ind. Eng. Chem. Res. 35, 2234-2246 (1996).

Belter P. A., E. L. Cussler and Wei-Shou Hu “Bioseparations. Downstream Processing for Biotechnology” John Wiley New York (1988).

Biegler L. T. , I. E. Grossmann and A. W. Westerberg “Systematic Methods for Chemical Process Design” Prentice – Hall (1997).

Douglas J. M. “Conceptual Design of Chemical Processes” McGraw-Hill NewYork (1988).

Flatz, W. “Equipment Sizing for Multiproduct Plants” Chem. Eng. July 13, 1981 pp 105

Floudas, C. A. and P. M. Pardalos “Recent Developments in Global Optimization andtheir Relevance to Process Design” Paper I09 at FOCAPD 99 Breckenridge,Colorado July 1999

Iribarren, Oscar A. “Batch Chemical Processes Design” Ph. D. Thesis University of Massachusetts, Amherst, 1985.

Modi A. and I. A. Karimi; “Design of Multiproduct Batch Processes with Finite Intermediate Storage”. Comp. Chem. Engng. Vol 13, p. 127-139 (1989).

Salomone H. E. and O. A. Iribarren “Posynomial Modeling of Batch Plants” Comp. Chem. Engng. Vol. 16, p. 173-184 (1992).

Salomone H., J. Montagna and O. Iribarren “Dynamic Simulations in the Design of Batch Processes” Computers Chem Engng 18 (3), 191-204 (1994).

Swami, Sunil “Preliminary Design and Optimization of Batch Processes” M.S. Thesis University of Massachusetts, Amherst, 1985.