The use of Monte Carlo simulation to support management decisions of industrial batch processes

Jacek Stelmach[1]

Abstract

The main goal for any company is its survival and development which consists of a set of specific objectives. However, a prerequisite for achieving these objectives is to achieve a net profit, which is the result of proper operational decisions. Currently, managers can benefit by taking such decisions with the support of various tools of modeling the behavior of company in the market. In contrast to continuous processes, for which a number of optimization methods was developed and published, batch optimization techniques are not so common. The major obstacle is the complexity of these processes - in which a huge impact on the quality and profitability of the final product have even slight variations in recipes and parameters of the raw materials, additionally their analytical description is very complicated. A further complication is the fact that some of the independent variables are random whose distributions may be different from a normal distribution. In addition, there is the interaction between basic products, additional products and waste in batch processes.

The article presents a business model of multi-batch processes in which a part of the independent variables are random. There have been carried out Monte Carlo simulation of the short-term impact of changes in these variables on the objective function: first coverage margin and indicated the possibility to use statistical inference to support management decisions on the example of the real production process.

Keywords: *batch process, Monte Carlo simulation, net profit.*
JEL Classification: C580.

1. Introduction

Management decisions in companies often use information obtained from the controlling department. Most are historical information, rarely forecast. In both cases, however, these data are point information (point estimators like expected value or point prediction). They forget that part of the input is stochastic in nature and should be treated as random variables. Sometimes there is the prediction based on worst-case and best-case assumptions but there is no knowledge of the dispersion or confidence intervals. As a result it is very difficult to estimate the risk of such decisions. Of course it is possible to calculate the intervals assuming knowledge of the distribution of the input variables. It needs to build a model. Often, however, it is difficult to match the known distribution function to empirical data. And as often empirical distributions are far from normal distribution. And most importantly – in batch processes, it is difficult to create such models (Abel et al., 2000).

[1] Corresponding author: University of Economics in Katowice, Department of Statistics, jacek.stelmach@polwax.pl.

Thus appears problem – how the information provided to management by using the available data can be enriched. The known possibility is to carry out Monte Carlo simulation. This method is known since the mid-twentieth century. It allows us to solve the statistical problems using repeating sampling to determine the properties and behavior of certain phenomenon what reducing the computations (Good, 2004). Additional benefit is that in such method, empirical data with no dedicated experiment can be used (Robert and Casella, 2010), (Kroese et al., 2011).

2. Batch processes

Batch processes occur most frequently in the chemical, refining, pharmaceutical and food companies. These processes as opposed to continuous processes are periodically repeated in separate parts of the system in a defined sequence of events called the recipe. Frequently in such processes operations of: blending, separation and reaction are repeated (Korovessi and Linninger, 2006) in a sequence (see Fig. 1):

1. initial operations and filling the reactor,
2. mixing the raw materials and the stabilization physical conditions inside the reactor,
3. start-up of batch process,
4. main operations,
5. operations ending, stabilization physical conditions inside the reactor,
6. removing products and wastes.

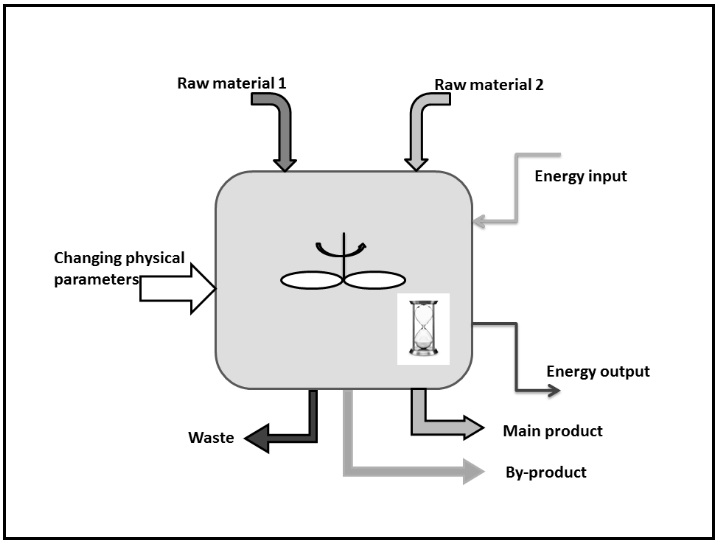

Fig. 1. Batch process scheme.

3. Problem description

One of the crucial decisions in the companies with batch processes is to decide about product portfolio that gives the highest net profit taking into consideration limitations in volumes of installation, known (or estimated) quality of raw materials, desirable quality of products and by-products, cost of waste, operational expenses. Unfortunately analytical optimization using known linear and nonlinear methods is almost impossible because (Abel et al., 2000), (Li et al., 1998), (Thokozani, 2010), (Barker and Rawtani, 2005), (Nomikos and MacGregor, 1995):

- the equations that describes phenomenon are very complex and nonlinear, and usually needs additional empirical corrections,
- real raw materials are not homogenic but a mixture of substances,
- the process does not proceed uniformly throughout the reactor volume and is disturbed by the events which are not always possible to identify and measure,
- there is an interaction concerning the volumes and quality of main product and by-products and waste,
- performed measurements and tests contain errors.

In such a situation commonly used expert knowledge is supported with the analyzes and simulations performed by specialized computer programs. Unfortunately, most often, these programs do not use the available information on distributions of random variables, bases on average values. It can lead to increased risk of taking a wrong decision if there is a particularly high dispersion of variables.

The problem is defined to propose a method that allows to verify the optimized product portfolio in a company with batch processes using the knowledge of the empirical distributions of variables included in the model. The method will be verified with empirical data of real batch processes.

4. Model presentation

The model was prepared under certain assumptions, see (Szczepankowski, 2004), (Dittmann, 2008), (AlGhazzawi and Lennox, 2009), (Lennox et al., 2001):

- the production bases on multiproduct batch processes with shared installations,
- a time and costs of any operation does not dependent on previous operation,
- each batch process has got own recipe including raw material, quality of product, process costs and efficiency,
- the waste are treated as by-products,

- there is implemented additional linear loss function for products with reduced quality,
- objective function is defined as a margin I less loss function.

Taking into account the following **variables**:

i = 1…I	raw material index
j = 1…J	product index
m = 1…M	batch process index
n = 1…N	installation index

p_{si} – unit price of i^{th} raw material	$\mathbf{p_s}$	[PLN/t]
p_{pj} – unit price of j^{th} product	$\mathbf{p_p}$	[PLN/t]
k_{zm} – unit variable costs of process m	$\mathbf{k_z}$	[PLN/t]
k_{sm} – fixed costs of process m in a period	$\mathbf{k_s}$	[PLN]
b_m – batch process	$\mathbf{b}$	
k_{wj} – unit costs of product j (weighted for all processes)	$\mathbf{k_w}$	[PLN/t]
f_{pj} – marker of quality requirements (0<f_{pj}<1 – fulfilled requirements; f_{pj}>1 - not fulfilled requirements)	$\mathbf{f_p}$	
b_{0j}, b_{1j} – coefficients of loss function for j^{th} product	$\mathbf{b_0}$, $\mathbf{b_1}$	
r_{smi} – recipe of process m (a volume of i^{th} raw material per 1 ton of input)	$\mathbf{R_s}$	
w_{pmj} – efficiency of process m (share of j^{th} product in 1 ton of output)	$\mathbf{W_p}$	
x_{mn} – assignment of pprocess b_m to the installation a_n (0 – no assignment, 1 – assignment)	$\mathbf{X}$	
q_j – volume of j^{th}product in a period	$\mathbf{q}$	[t]
u_i – delivery of i^{th} raw material in a period	$\mathbf{u}$	[t]
y_{mj} – volume of production in process m necessary to produce j^{th} product in a period	$\mathbf{Y}$	[t]

with the **limitations**:

s_{li} – availability of i^{th} raw material in a period	$\mathbf{s_l}$	[t]
a_{ln} – max unit production volume of installation n	$\mathbf{A_l}$	[t/h]
d_{lj} – demand on j^{th} product in a period	$\mathbf{d_l}$	[t]

and **interactions**:

z_{mj} – necesstity of production j^{th} product in process m	$\mathbf{z_m}$

M_j set contains all the processes m in which j^{th} product is created.

Based on an analysis of empirical data coming from one year period, several of them: p_{si}, p_{pj}, d_{lj}, w_{pmj}, f_{pj} has been classified as random variables, that cannot be treated as fixed. As a criterion for optimization of short-term profit and loss account – margin I was chosen, taking into account the costs of reduced quality less fixed costs allocated to the products:

$$F = \sum_{j=1}^{J} m_{Ij} - \sum_{m=1}^{M} k_{sm} \tag{1}$$

where margin I is calculated for j[th] product according to:

$$m_{Ij} = \sum_{m \in M_j} (y_{mj} \times w_{pmj}) \times (p_{pj} - k_{wj}) - c_j \tag{2}$$

and operational costs are:

$$m_{Ij} = \sum_{m \in M_j} (y_{mj} \times w_{pmj}) \times (p_{pj} - k_{wj}) - c_j \tag{3}$$

with manufacturing costs below.

$$k_{wj} = \frac{\sum_{m \in M_j} k_{zm} \times y_{mj} \times w_{pmj}}{q_j} + \frac{\sum_{m \in M_j} y_{mj} \times \sum_{i=1}^{I} p_{si} \times r_{smi}}{q_j}\,. \tag{4}$$

Depends on the agreements, the penalty c_j can be paid either as a fixed value or a value proportional to product volume, so is to be calculated according to the formula below.

$$c_j = f_{pj} \times (b_{0j} + q_j \times b_{1j}) \times \mathbf{1}(f_{pj} - 1)\,. \tag{5}$$

This model was used to carry out simulation Monte Carlo investigations, to find the effect of the variability of certain parameters of the model.

5. Monte Carlo simulation

As mentioned above certain variables: p_{si}, p_{pj}, d_{lj}, w_{pmj}, f_{pj} were classified as random variables. A simulation was carried out with empirical data coming from one of chemical companies. A portfolio of the company contains many products often similar with different package or method of application, so it was decided to choose four dominant products and to make the investigations concerning the distribution of margin I for each product and the distribution of whole objective function. The real data come from the interval January 2012-December 2013. The study was performed for three cases with 1000 simulations:

1. all random variables were treated as ‘random’ with the known empirical distribution,
2. all random variables were treated as ‘random’ with the distribution approximated to normal distribution - what exists in practice,

3. external variables – i.e. the variables affected by the external environment ($\mathbf{d_l}$, $\mathbf{p_p}$ and $\mathbf{p_s}$) were treated as fixed and replaced with the median.

Very interesting are the distributions of prices: for raw materials and products because there are far from normal distribution: skewed and relatively heavy-tailed as presented in fig. 2 and fig. 3. And it is not unusual in real life, although often any statistical inference is carried out assuming a normal distribution.

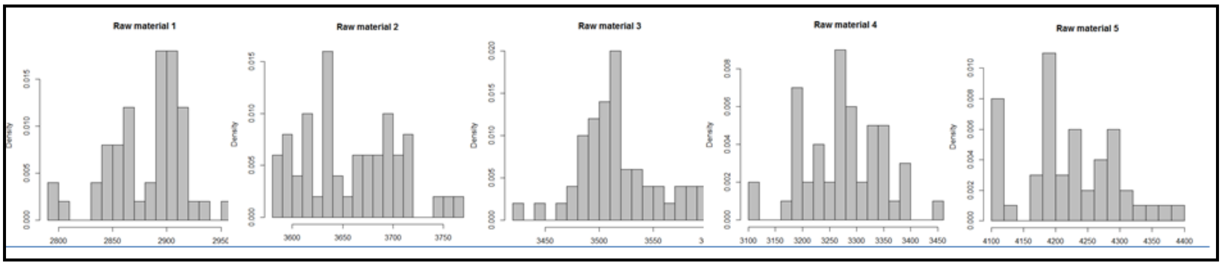

Fig. 2. Histograms of prices of raw materials.

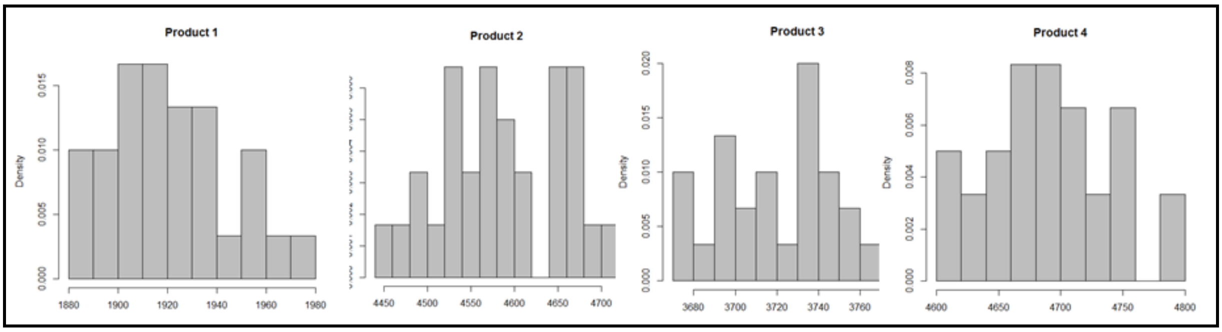

Fig. 3. Histograms of prices of products.

The distributions of objective function for all three cases are presented in fig. 4. The distributions of margin I for each product are placed in fig. 5 – fig. 7.

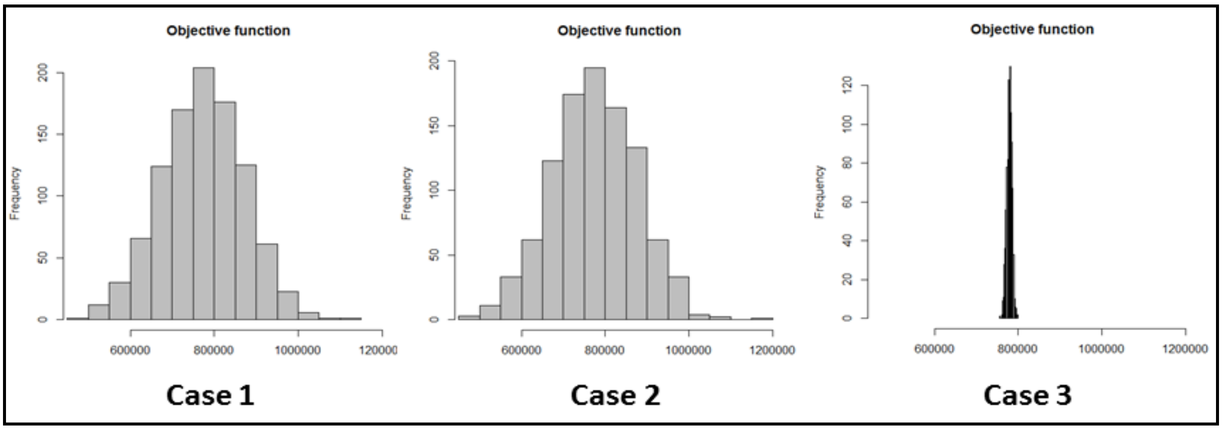

Fig. 4. Histograms of objective function.

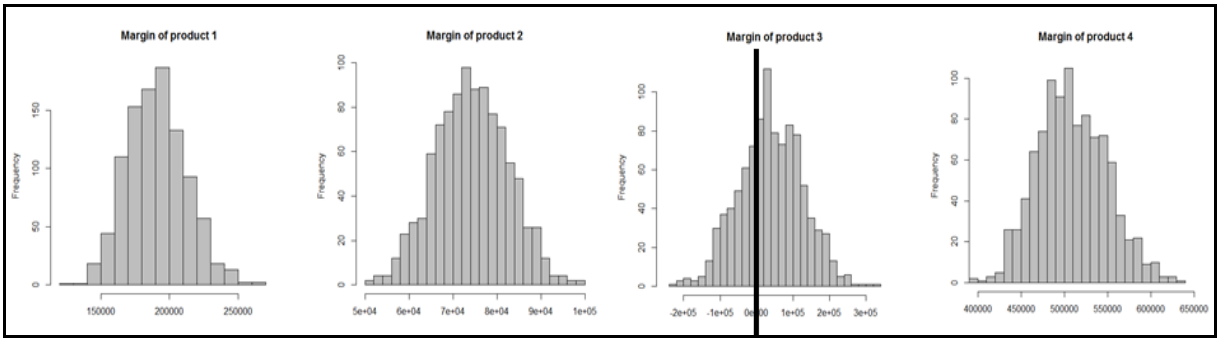

Fig. 5. Histograms of margin I – first case (black line indicates the level of zero).

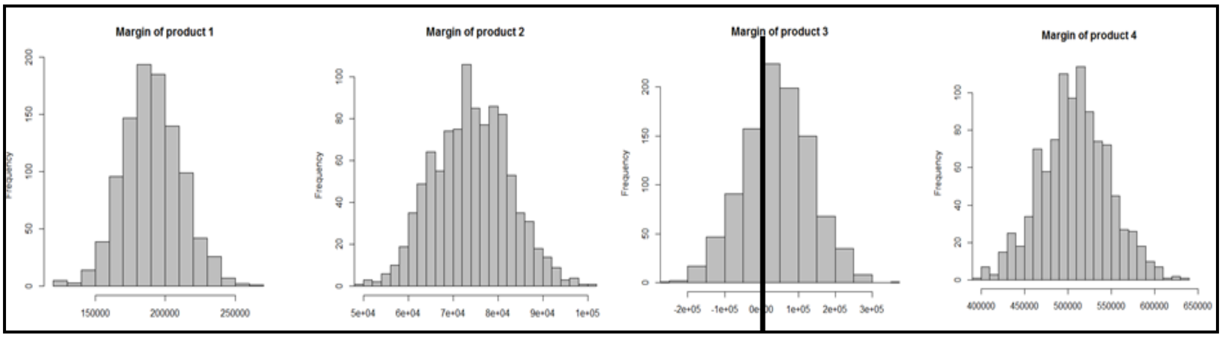

Fig. 6. Histograms of margin I – second case (black line indicates the level of zero).

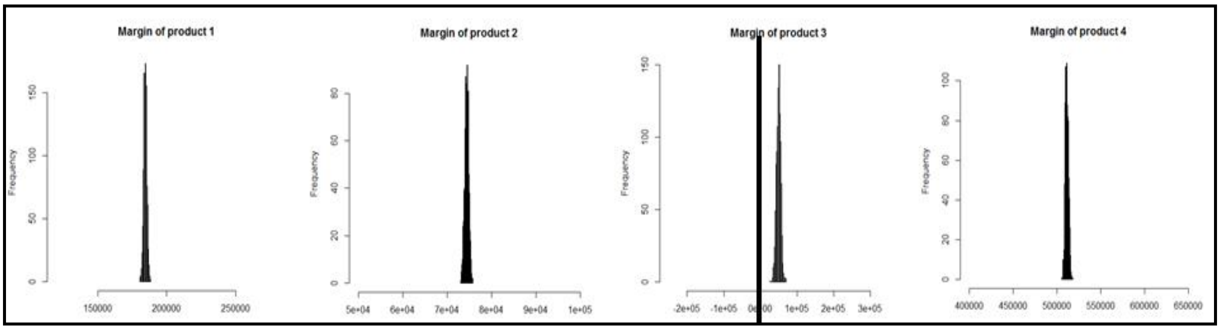

Fig. 7. Histograms of margin I – third case (black line indicates the level of zero).

Basic measures - position and dispersion of simulated margins are shown in table 1.

Case no	Measure	Objective function	Margin of product 1	Margin of product 2	Margin of product 3	Margin of product 4
First	Mean	772189	190777	74234	39274	509065
	Standard deviation	98004	21363	8240	86780	40094
Second	Mean	773671	190811	74172	41773	508075
	Standard deviation	101062	20835	8563	91481	39556
Third	Mean	778883	184522	74428	49531	511563
	Standard deviation	6265	1173	454	5747	1892

Table 1 Basic measures of simulated margins.

Although (case 2), the values of variables with non-normal distribution were simulated as normally distributed, the results compared to empirical data (case 1) are very similar. Number of factors that affect the value of the margins was so large that the differences in the distributions from the normal one had little effect on the final distributions. Much more important is the omission of the randomness of variables (case 3), which results in loss of information about the dispersion of margins (and a small bias). In the extreme case (product 3) this can lead to a lack of knowledge about the dangerous probability of loss of profitability.

Conclusion

Most often empirical variables are random. This is a result of measurement errors and the large number of influencing factors in their values, which are often not fully identified. Therefore, treatment of these variables in microeconomic models, as non-random variables leads to unnecessary loss of information and may be the cause of bad management decisions. In such cases, Monte Carlo simulations can be used to determine the distribution of economic indicators. This is especially important if the analytical model is very complex and does not allow to use these variables in the form of mathematical formulas – as for example in analyzed batch processes.

Acknowledgements

The work paper funded by resource on science period 2012-2015, National Science Centre Nr DEC-2011/03/B/HS-4/05630.

References

Abel, O., Helbig, A., Matquardt, W., Zwisk, H., & Daszkowski, T. (2000). Productivity optimization of an industrial semi-batch polymerization reactor under safety constrains. *Journal of Process Control*, *10*, 351-362.

AlGhazzawi, A., & Lennox, B. (2009). Model predictive control monitoring using multivariate statistics. *Journal of Process Control*, *19*, 314-327.

Barker, M., & Rawtani, J. (2005). *Practical batch process management*. Oxford: Elsevier.

Dittmann, P. (2008). *Prognozowanie w przedsiębiorstwie*. Kraków: Oficyna a Wolters Kluwer business.

Good, P. (2004). *Permutation, parametric and bootstrap tests of hypotheses.* New York: Springer Science Business Media B. V.

Korovessi, E., & Linninger, A. (2006). *Batch processes*. Boca Raton: Taylor@Francis Group.

Kroese, D., Taimre, T., & Botev, Z. (2011). *Handbook of Monte Carlo methods.* New York: Wiley Series in Probability and Statistics John Wiley&Sons.

Lennox, B., Montague, G., Hiden, H., Kornfeld, G., & Goulding, P. (2001). Process monitoring of an industrial fed-batch fermentation. *Biotechnology and Bioengineering*, *74*(20), 125-135.

Li, P., Arellano, H., Garcia, H., Wozny, G., & Reuter, E. (1998). Optimization of a semibatch distillation process with model validation on the industrial site. *Industrial & Engineering Chemistry Research*, *37*, 1341-1350.

Nomikos, P., & MacGregor, J. (1995). Multivariate SPC charts for monitoring batch processes. *Technometrics*, *37*(1), 41-59.

Robert, C. P., & Casella, G. (2010). *Introducing Monte Carlo with R*. New York: Springer.

Szczepankowski, P. J. (2004). *Zarządzanie finansami przedsiębiorstwa*. Warszawa: Wydawnictwo Wyższej Szkoły Przedsiębiorczości i Zarządzania im. Leona Koźmińskiego.

Thokozani, M. (2010). *Batch chemical process integration*. New York: Springer Science Business Media B. V.