Scientific paper

Technology Transfer Criteria from Pilot to Multipurpose Production Plant – a Case of Total Production Costs

Robert Agnič

Lek – A Sandoz company, Kolodvorska 27, 1234 Mengeš, Slovenia. Tel.: +386 1 7217 543, Fax: +386 1 7217 678, E-mail: robert.agnic@sandoz.com.

Received: 06-10-2006

Abstract

The need for optimizing chemical and fine chemical production processes in industry by taking economic issues into account has increased more than ever before. This work deals with the evaluation of processing data variables with multiple regression analysis in an existent production processes. The defined regression model is afterwards used to predict total production costs – TPC in transfer procedure of new processes from pilot to multipurpose production plant. Decision-tree diagram was also developed to simplify transfer procedure. Both, decision-tree diagram and defined regression model considerably affect the lower uncertainty of confidence for transferring new product processes, which can obviously lead to increased cost efficiency and competitiveness at the market.

Keywords: Multiple regression analysis, technology transfer procedure, multipurpose production plant, total production costs

1. Introduction

Chemical and fine chemicals production processes are often controlled by 'experience factors' and in applying new products with new technology procedures no firm basis is available for predicting process trends. Performed data of operating variables through pilot plant batches can prevent higher failures, but there is always a level of uncertainty as to what will happen with the predicted values of factors during scale-up to production plant. It seems logical then to obtain useful relations during production plant operations, which can be developed from the operating data. Regression analysis with associated significance tests are most valuable tools for evaluating operating data to achieve regression equation, which can help to optimize the existing processes and those which will be transferred from the pilot plant.

Kravanja and Iršič-Benedik optimized chemical processes with analyses of kinetic reactions.[1] Their results show that the optimal process scheme gave most economical benefit outcome. Fisher found out that evaluating plant operating data on three products in polymerization process with multiple regression analysis, yielded optimal production parameters, which can be used to perform higher yields and consecutively higher conversions.[2] Another process shows that the use of regression analysis lead to mathematical models, which indicates that important economic improvement in processing can be realized with relatively minor modifications to the circuits.[3] Also, waste water treatment plants use mentioned analysis for evaluation of process efficiency according to the removal of toxic and harmful chemicals.[4–6] Schweigert et.al. demonstrate the weather affects the concentrations of nitrate in soil and groundwater and how the use of multiple regression analysis is crucial in future projects, because of high sampling costs.[7] Modeling with multiple regression analysis reduces interferences in atomic absorption spectrometry which confirms it's usage in analytical branch as well.[8] Swamidass and Avittathur studied matching of plant and supplier flexibility with regression analysis.[9] Regression models result in connection relations between the profitability and flexibility of plants and suppliers.

The present paper shows a regression model based on operating data of several production processes performed on multipurpose production plants. This model was afterwards used to establish a decision-tree diagram, which can predict efficiency and justification for new technology process transfer from pilot to multipurpose production plant.

2. Data Selection

As mentioned before in the introduction section the data for regression analysis were selected from eight different production processes (that also means eight different products). During selection it was found out that eight variables (presented in Table 1) significantly influence production processes. Before any analysis, it is important to distinguish between dependent and independent variables, because only independent variables should be used to predict variation of the dependent variable. Table 1 shows defined variables, by which regression analysis has been performed.

Variable material consists of a combination of raw materials and intermediates. Raw materials are defined as auxiliary materials to realize the production process, while intermediates are starting materials, i.e. molecules or parts of final molecules. The work demonstrates the costs of employees to executing the tasks according to the production processes. Equipment has been evaluated through it depreciation and maintenance costs used in the selected production processes. Quality includes costs of analysis, laboratory equipment etc.

Table 1. Definition of dependent and independent variables.

ENABLERS = independent variables
Material
Work
Equipment
Quality
Energy
Other costs
Capacity
RESULTS = dependent variable
Total production costs ↔ *TPC*

The energy part covers heat transfer media, demineralised water and technical gases, together with all manipulation steps connected to them. Other costs include ecology, marketing, sales and management cost incurred by the activities related to these functions. Capacity defines the capacity of equipment for appointed production process. Finally, TPC represents total production costs caused by production processes, depending on products, produced on multipurpose production plant.

Defined variables are represented as average costs in EUR (except capacity), since all of them can be evaluated through costs. Costs data were collected for two periods: one includes all batches from year 2005 and the other half of the year 2006, which means 128 data points in total. A rule of thumb often used is that at least ten times more data points should be taken than the number of independent variables. Table 2 presents average costs data, with standard deviations (in bracket) for eight different products and eight defined variables in the periods mentioned. It is important to note that the data in Table 2 include four so-called 'old' and four 'new' products. Old products mean production processes introduced before year 2004 (labelled A, B, C, D) and new products after that date (labelled F, G, H, I).

3. Multiple Regression Analysis

Operations data from the production processes were analyzed with multiple regression analysis, using SPSS 14.0 program (Statistical Package for Social Sciences) for analyzing data. For all analyses 5% significance level was considered. T – criteria and F – criteria were chosen to eliminate statistically insignificant independent variables in regression equation. Adjusted coefficient of determination (R^2) was used to define the amount of dispersed data covered by evaluated regression model. Pearson's method was used to find out linear correlations between variables. The variables equipment and work showed lower linear correlation with variable TPC then others. According to the experience, these variables (equipment and work) significantly influence the variable TPC and were included into analysis irrespective of Person's method results.

4. Regression Model

During searching for a suitable regression model it was decided that the model should be able to evaluate dependent variable (TPC) versus real values of TPC based on multipurpose production plant data in range of 10%. SPSS program uses multiple regression analysis in *Stepwise* method to evaluate the relationships with suitable regression model (Stepwise method evaluates independent variables according to linear correlation with depent variables one by one to regression model and exclude variables according the probability level).[10] Table 3 shows the results of regression analysis, where operating data from real production processes were used.

From the relationships presented in Table 3, it is inferred that the regression analysis has led to the formulation of the following equation or regression model:

$$\text{TPC} = 1.103 \cdot \text{Material} + 2.059 \cdot \text{Equipment} + 1.904 \cdot \text{Work}. \quad (1)$$

It was found out that the regression model is statistically significant, meaning that t-criteria indicate that all variables in equation are significant at the probability level of 5% (independent variables different from zero value. according to zero hypothesis – H_0).[11] Moreover, adjusted coefficient of determination covers 99% of all data, which means high satisfactory level ($R^2 = 0.99$). Thus achieved regression model was used for calculating the predicted values of dependent variable TPC for all production processes. The reason of this evaluation was to find out the accuracy and applicability of the regression model. The comparison of predicted (calculated) and real values of TPC is presented graphically in Figure 1, which demostrates acceptable accordance of curves.

Table 2. Average results of measured variables with standard deviations.

Material	Work	Equipment	Quality	Energy	Other costs	Capacity	TPC	Product	Year
269.58 (6.53)	209.55 (4.11)	95.76 (4.87)	18.23 (2.01)	113.91 (5.12)	196.92 (5.53)	2600 (31)	903.94 (13.33)	A	2005
4778.02 (137.24)	12.00 (0.82)	15.66 (1.05)	85.27 (3.92)	59.29 (2.45)	315.80 (7.18)	785 (25)	5266.04 (102.54)	F	
234.95 (5.78)	69.94 (2.33)	23.09 (1.12)	9.56 (0.81)	9.74 (1.12)	108.39 (4.32)	1500 (28)	455.67 (10.83)	B	
870.69 (8.32)	119.40 (3.21)	126.62 (3.15)	7.16 (0.80)	23.37 (1.98)	183.45 (4.93)	6237 (43)	1330.69 (19.21)	G	
364.07 (7.21)	1.42 (0.12)	1.03 (0.27)	3.11 (0.32)	0.37 (0.05)	22.27 (0.87)	8881 (45)	392.27 (8.89)	C	
1032.52 (53.62)	88.49 (2.89)	32.90 (1.57)	49.61 (2.36)	9.13 (1.06)	158.51 (3.91)	2500 (30)	1371.16 (18.72)	H	
254.49 (5.78)	70.27 (2.54)	70.14 (2.32)	15.46 (1.89)	29.96 (2.05)	127.69 (3.18)	4500 (36)	568.01 (11.47)	D	
1475.00 (61.50)	111.67 (3.67)	504.88 (11.83)	22.31 (2.17)	15.43 (1.43)	336.69 (7.30)	5000 (38)	2898.71 (45.31)	I	
289.34 (6.43)	183.51 (3.88)	42.71 (2.75)	16.64 (1.92)	49.57 (4.21)	111.17 (4.41)	2127 (29)	692.94 (12.02)	A	2006
2639.96 (78.12)	139.69 (7.22)	34.43 (1.12)	5.61 (0.71)	122.43 (2.94)	384.95 (7.42)	2000 (25)	3327.07 (73.15)	F	
282.62 (5.92)	15.97 (0.42)	8.02 (0.92)	1.74 (0.32)	4.42 (0.83)	28.87 (4.10)	961 (26)	341.64 (9.87)	B	
496.79 (6.74)	123.59 (3.15)	246.67 (4.85)	7.85 (0.81)	50.89 (2.31)	288.04 (5.16)	6000 (41)	1213.83 (17.31)	G	
231.55 (6.10)	16.31 (1.18)	0.74 (0.13)	3.41 (0.31)	1.63 (0.37)	29.86 (0.96)	5000 (38)	283.49 (8.03)	C	
971.19 (42.78)	71.85 (2.53)	50.12 (1.71)	34.18 (2.18)	10.54 (1.15)	105.70 (3.63)	2295 (30)	1243.58 (16.52)	H	
298.92 (6.02)	50.72 (2.06)	24.47 (1.63)	14.11 (1.84)	12.55 (1.47)	60.51 (2.51)	4453 (35)	461.28 (8.74)	D	
1515.09 (64.11)	189.11 (3.97)	0.00 (0.00)	38.19 (2.43)	31.36 (2.03)	226.05 (6.17)	3572 (34)	1999.8 (38.86)	I	

Table 3. Results of multiple regression analysis.

Model		Unstandardized Coefficients B	Std. Error	T	Sig.	95% Confidence Interval for B Lower Bound	Upper Bound
1	(Constant)	325.210	110.680	2.938	.011	87.825	562.594
	Material	1.096	.072	15.249	.000	.942	1.251
2	(Constant)	143.762	50.929	2.823	.014	33.737	253.787
	Material	1.094	.030	36.660	.000	1.030	1.159
	Equipment	2.299	.278	8.260	.000	1.698	2.900
3	(Constant)	–21.515	26.785	–.803	.437	–79.874	36.845
	Material	**1.103**	.011	**97.328**	**.000**	1.079	1.128
	Equipment	**2.059**	.109	**18.941**	**.000**	1.822	2.296
	Work	**1.904**	.214	**8.881**	**.000**	1.437	2.371

a) Dependent Variable: TPC, $R^2 = 0.99$, F = 3328.7.

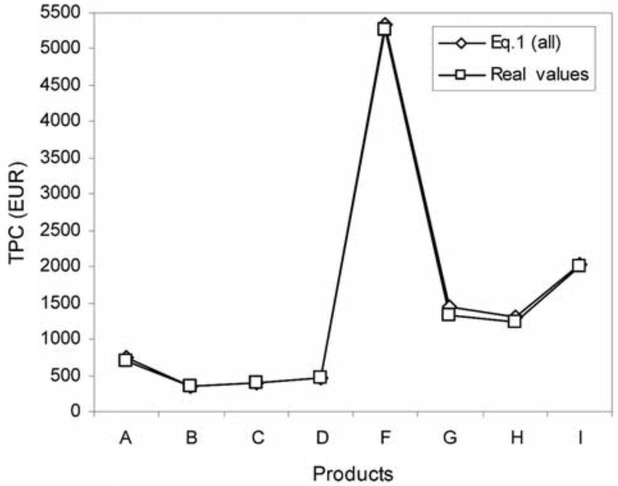

Figure 1. Comparison of predicted values by regression model, Eq.1, and real values.

Figure 2 shows differences between the predicted (calculated) and real values in percents for all products.

As shown in Figure 2, the predicted values with regression model evaluate the dependent value – TPC in range of 10% as it was demanded at the beginning of this

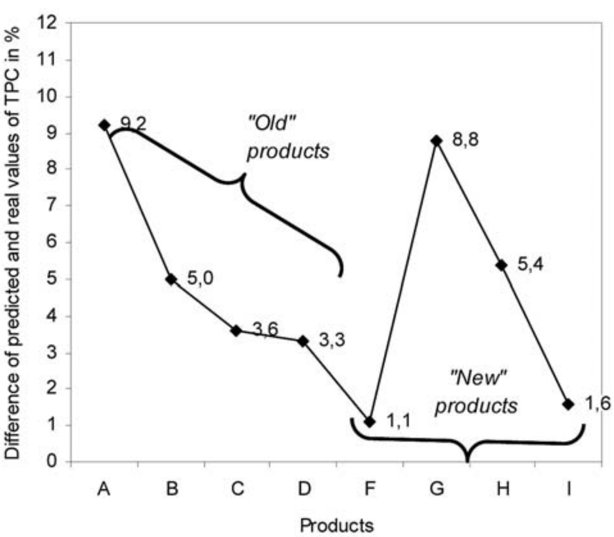

Figure 2. Comparison of predicted and real values of TPC in %.

analysis. It is interesting to observe, that the so-called old products have on the average 5.3% difference between the observed values of TPC, and the new ones almost cover the values of the predicted and real TPC with only 4.2% difference in average. This effect can be explained through main costs, which influence new product prices, which have independent variables accounted in the regression model as material, work and equipment. As opposed to old products, the intermediates and raw materials are still leading in high cost shares, but there are also significant other variables, which were excluded from the model during multiple analysis. This is why higher differences between calculated and realistic values are presented. Nevertheless, according to the aim to find out a useful regression model to assure successful new technology transfer from pilot to multipurpose production plant, the achieved difference for new products (in average 4.2%) between predicted and real TPC values is more than suitable.

5. Regression Model for Old and New Products

During regression analysis it was noticed that old and new products have different cost values. The idea was to analyse cost data separately for new and old products by multiple regression analysis. The results are shown in Tables 4 and 5.

The regression model for old products in defined significance level of 5% includes variables energy and quality also with constant value:

$$TPC = 305.882 + 3.768 \cdot Energy + 9.909 \cdot Quality. \quad (2)$$

Multiple regression analysis of new products data leads to regression model:

$$TPC = 1.129 \cdot Material + 2.141 \cdot Equipment + 2.544 \cdot Work \quad (3)$$

It was shown that Eq. 2 (old products) includes different variables in the regression model than Eq. 1 (all products) and Eq. 3 (new products). The reason lies in different cost data. Old products have already passed lear-

Table 4. Results of multiple regression analysis for old products.

Model		Unstandardized Coefficients		T	Sig.	95% Confidence Interval for B	
		B	Std. Error			Lower Bound	Upper Bound
1	(Constant)	371.785	27.726	13.409	.000	303.942	439.629
	Energy	5.064	.609	8.321	.000	3.575	6.553
2	(Constant)	**305.882**	31.486	**9.715**	**.000**	224.945	386.818
	Energy	**3.768**	.648	**5.817**	**.002**	2.103	5.433
	Quality	**9.909**	3.715	**2.667**	**.045**	.359	19.460

a) Dependent Variable: TPC, R^2 = 0.95, F = 73.5

Table 5. Results of multiple regression analysis for new products.

Model		Unstandardized Coefficients		T	Sig.	95% Confidence Interval for B	
		B	Std. Error			Lower Bound	Upper Bound
1	(Constant)	614.027	229.656	2.674	.037	52.079	1175.974
	Material	.997	.106	9.373	.000	.737	1.257
2	(Constant)	215.813	106.691	2.023	.099	–58.446	490.072
	Material	1.076	.041	25.935	.000	.969	1.183
	Equipment	2.074	.334	6.208	.002	1.215	2.933
3	(Constant)	–156.500	101.711	–1.539	.199	–438.896	125.896
	Material	**1.129**	.023	**48.082**	**.000**	1.064	1.195
	Equipment	**2.141**	.160	**13.344**	**.000**	1.695	2.586
	Work	**2.544**	.601	**4.230**	**.013**	.874	4.213

a) Dependent Variable: TPC, R^2 = 0.99, F = 992.7.

ning curves, material costs do not affect the production costs as it is a case with new products (higher number of tendering firms), equipment depreciation costs are payed off, etc. Therefore, other variables have greater impact on the production costs.

Figure 3 represents evaluation of regression models (Eqs. 1–3) used for calculation of TPC predicted values with real TPC data.

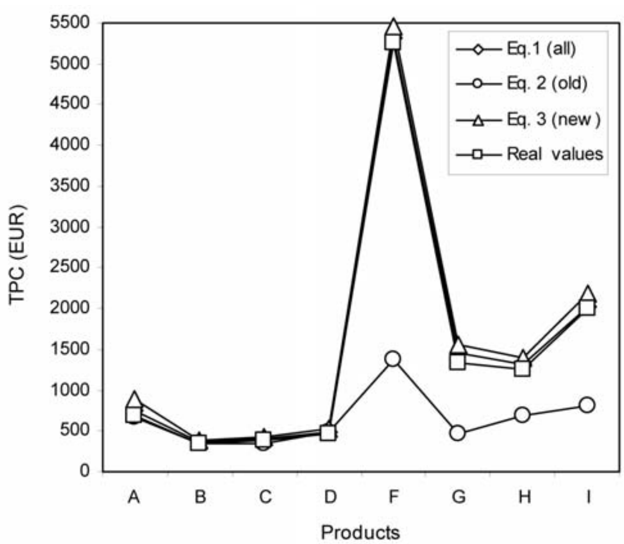

Figure 3. Comparison of predicted values by regression models (Eqs. 1–3) and real values.

Figure 3 shows that Eq. 2 is not appropriate for the prediction of TPC prior to technology transfer procedure. Eq. 3 also predicts TPC values with high accuracy level as Eq. 1, which was expected (because of the similarity between regression models). Nevetheless, Eq. 1 shows higher accordance of predicted TPC values with real TPC data than Eq. 3. For this reason it was chosen as the most valuable and useful way to predict TPC variable.

6. Decision-Tree Diagram

According to the findings it is now possible to establish decision-tree diagram, which could be used as a manual or guide for technology transfers of new processes from pilot to production plant (Figure 4).

The decision-tree digram shows a simple picture for defining the status of a new technological process. Synthesis development starts in lab, after succesfull trials it proceeds to pilot plant. Pilot batches confirm the previously defined parameters of new technological process. Cost evaluation of independent variables is introduced and these data are used for calculating the TPC predicted value.

At the first decision point, values of TPC received from pilot plant batches are compared with the predicted – calculated TPC from the regression model. If lower pilot TPC values are achieved compared to the calculated TPC, it means that there will be lower costs during scale-up procedure and later during full scale production. Therefore, technology transfer of new process can be approved.

If there is an opposite situation, the existing process should be considered according to the differences in costs, and decision made, where to return this process. In case that the process needs only few modifications (e.p. temperature, or pH range set up), this can be settled at the pilot plant. If calculations indicate that the process needs much more corrigenda, it must be returned to the roots, into i.e. the synthesis department. Why so? A simple look on the costs produced on lab or pilot scale in comparison with the production plant for a modification of the specifical process shows, that this costs can be ten times or higher on the production plant. The reasons are higher capacities, cleaning procedures necessary for manufacturing another product, knowledge and education of process operators, occupation of multipurpose production plant with other products (time delays, deficit of other products). When there is opinion, that the returned technological process can be re-evaluated, we should proceed according to the decision-tree diagram as before.

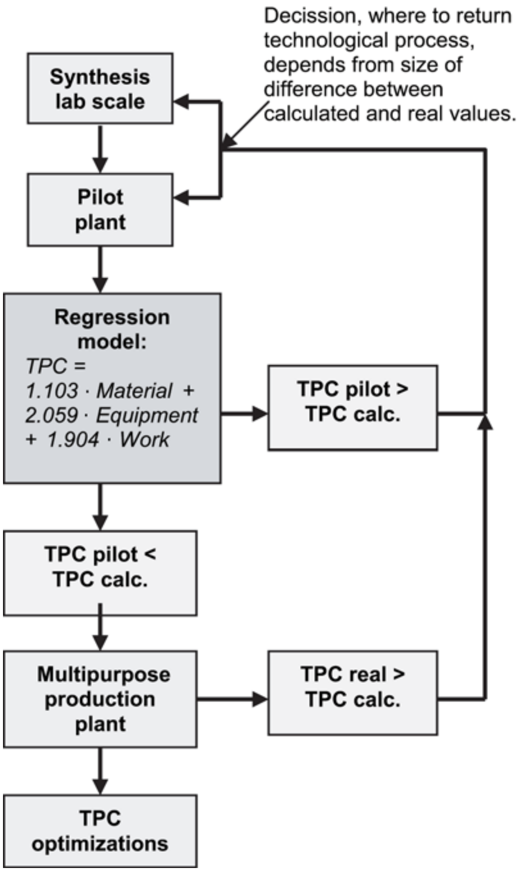

Figure 4. Scheme of the decision-tree diagram for technology transfer.

After approval for transferring technology of a new process to the production plant, production scale batches can started. After some of them have been produced, realistic values of TPC can, or must be examined against calculated ones. Higher values of real TPC indicate that something has happened with the process during transfer and scale-up procedures, so it must be considered to turn the process back to the development departments to check the reasons for costs failure. If the realistic TPC is lower, production campaigns can proceed, meaning that technology transfer of new process has been successful and can be approved. The only step after is to assure the life cycle management of this process and the product connected with it, i.e. polishing the process and consecutively, the cost efficiency.

7. Technology Transfer Case

Production process technology transfer for product X shows practical use of the decision-tree diagram. Table 4 presents cost data based on pilot plant batches for product X needed for the prediction of TPC values for multipurpose production plant. Variable equipment is not considered, because no new equipment has been purchased and preparation of the existing plant according to the production process of product X is accounted in variable work.

Table 6. Pilot plant costs – product X.

Pilot batches	Materials	Work	TPC
EUR	66748.60	14221.09	80969.69
EUR/kg*	370.83	79.01	**449.84**

* Campaign 180 kg.

The regression model presented data (Table 6) indicate that TPC calculated value is 559.45 EUR. Evaluation through decision-tree diagram shows that pilot TPC is **lower** than the TPC calculated. If this is a case, technology transfer for product X can be approved to start production batches on multipurpose plant.

Table 7 presents the costs of production batches on a multipurpose plant. Higher production costs according to calculated ones indicate transfer failure, which should be analyzed in the development scale according to the decision-tree diagram.

Table 7. Production plant costs – product X.

Production batches	TPC
EUR	122320.20
EUR/kg*	**821.55**

* Campaign 148.9 kg.

Aditional development evaluation of the process shows differences between the pilot and the production scale during the heat transfer phase. During dissolving of solid NaOH on pilot plant two hours were needed to finalize operation. Production scale shows that two hours was not enough to dissolve all NaOH in solution due to lower heat transfer. This point is crucial for this reaction step to achive higher yields. so less NaOH in solution resulted in lower yields as expected and also higher consumption of materials to achieve expected yields (NaOH and solvent $SOCl_2$). Prolongation of the dissolving step and visual control was implemented.

Modified parameters resulted in the process with lower costs (TPC = 603.02 EUR). Achieved costs are acceptable and still slightly higher than the predicted ones, but considering the results from the regression model, which is based on old and new products, we can expect that cost optimization of process wil result lower TPC values.

8. Conclusions

This paper describes efforts to decrease the level of uncertainty or confidence and technology transfer costs of new processes at a production site in order to make the products price more competitive in an ever-increasing struggle for market share. It is shown that a simple and effective regression model can predict the key variable values, which can be used in decision making about new process transfer from the development to production plant. It is evident that this regression model (Eq. 1) predicts TPC in average in the failure range of 4.2% for new products and consecutively for new processes. The decision-tree scheme was developed, which is a simple and useful tool during technology transfer procedure as shown in the case of product process transfer. The results from this article can help to save the time and decrease costs, when transferring new product processes from pilot to multipurpose production scale.

9. Acknowledgement

Author gratefully thanks and acknowledges prof. dr. Valentin Koloini for the professional contributions and competent support at this article.

10. References

1. I. N. Benedik, Z. Kravanja, Slovenski kemijski dnevi, Maribor, **2002**, pp. 530–537.
2. F. P. Fischer, *Ind. Eng. Chem.* **1960**, *52*, 981–984.
3. G. Chi, M. C. Fuerstenau, P. A. Anderson, *Intern. J. Mineral Process.* **1997**, *49*, 185–192.
4. S. P. Golfinopoulos, G. B. Arhonditsis, *Chemosphere* **2002**, *47*, 1007–1018.
5. P. Geladi, L. Hadjiiski, P. Hopke, *Chemom. Intell. Lab. Syst.* **1999**, *47*, 165–173.
6. S. Morad, M. Shacham, A. Brenner, *Chem. Eng. Process.* **2007,** *46,* 22–229.
8. P. Schweigert, N. Pinter, R. R. van der Ploeg, *J. Plant Nutr. Soil Sci.* **2004**, *167*, 309–318.
9. M. Grotti, M. L. Abelmoschi, F. Soggia, C. Tiberiade, R. Frache, *Spectrochim. Acta B* **2000**, *55*, 1847–1860.
10. B. Avittathur, P. Swamidass, *J. Oper. Manag.* **2006**, in press.
11. J. H. Zar (Ed.), Biostatistical Analysis, 2nd, Prentice-Hall, Engelwood Clifs, New Jersey, **1984**.
12. SPSS® 14.0 for Windows, SPSS Inc. Chicago, **2005.**

Povzetek

Potreba po optimizaciji kemijskih in farmacevtskih proizvodnih procesov v industriji z uporabo ekonomskih orodij narašča bolj kot kdajkoli prej. Predstavljeno delo ocenjuje procesne spremenljivke iz obstoječih tehnoloških procesov s pomočjo multivariantne regresijske analize. Določen regresijski model je tako uporabljen za napovedovanje vrednosti TPC (celotnih proizvajalnih stroškov) med procesom prenosa novega izdelka iz razvoja na večnamenski proizvodni obrat. S temi podatki je bil izdelan diagram oziroma odločitveno drevo za poenostavitev prenosov. Tako regresijski model kot odločitveno drevo bistveno vplivata na zniževanje stopnje negotovosti ob prenosu procesov novih izdelkov, kar očitno vodi k izboljšani stroškovni učinkovitosti in konkurenčnosti na trgu.