PLANNING, ANALYZING AND OPTIMIZING EXPERIMENTS

Branko POPOVIĆ[1]

[1]University of Belgrade, Mechanical Faculty, Kraljice Marije 16, 11000 Belgrade, Republic of Serbia,
E-mail: branko@popovic.org

Research and Development (R&D) encompasses all the innovative activities of organizations engaged in the development of new products (recycled materials, hardware, software, services) or enhancement of existing ones. In the US, the Research & Development cost to revenue ratio averages 3.5% but reaches even higher levels of up to 40%. Research & Development applies scientific methods with iterative and cyclical processes through which information is constantly updated and replaced, using the Experimenting (Design of experiments, DOE). The experimenting is an empirical way to arbitrate competing models or hypotheses in order to test existing theories or new hypotheses. The experimenting not only contains the design as it is commonly thought but includes the sequential stages: creating designs, analyzing and constructing the plans of the experiments, optimizing the plans of the experiments and optimizing the response of the experiments. Experimental designs include different types, but today, factor designs, responsive surface designs, mixture design, and Taguchi designs are mostly used. This paper demonstrates the application of the methodology of experimentation with the creation, analysis, development of experimental plans, optimization of the experimental plan, and optimization of the experimentation response, in one specific example of a chemical reaction from practice, using the computer program Minitab®

Keywords: Design of Experiments (DoE).

INTRODUCTION

Today, every major industrial production in the world is based on Research and Development (R&D), which encompasses the innovative activities of organizations undertaken to develop new products (processed materials, hardware, software, services) or to improve existing ones. Typically, two basic organizational unit models have been encountered: R&D Departments with new process results development tasks and Industrial Research Departments with applied research tasks in scientific or technological fields that facilitate the development of process results. These departments are different from the vast majority of others in that they do not aim to profit immediately, because they bring greater risk and uncertainty to return on investment, but are key to gaining large market benefits by selling new products. In the US, the R&D cost and overall profit of organizations reach an average of 3.5% but reaches even higher levels (14.1% Merck & Co, 15.1% Novartis, 24.9% Ericsson, 43, 4% Allergan). Research and development apply scientific methods to iterative and cyclical processes through which information is constantly updated and replaced by Experimenting (Design of Experiments, DoE). Experimenting is an empirical way of arbitrating competing models or hypotheses, in order to test existing theories or new hypotheses. Major problems in Experimenting include achieving randomization, reliability, and replication to obtain the required statistical power, sensitivity, and orthogonality. Considerable advances in experimentation occurred in the early 20th century, with contributions from statisticians such as Ronald Fisher (1890–1962), Jerzy Neyman (1894–1981), Oscar Kempthorne (1919–2000), Gertrude Mary Cox (1900–1978), and William Gemmell Cochran (1909–1980), (Popović, & Ivanović, 2018, 2019).

The experimenting not only contains planning as it is usually thought, but it also includes the

sequential stages: creating designs, analyzing and constructing the plans of the experiments, optimizing the plans of the experiments and optimizing the response of the experiments. Creating Designs plans is the stage of forming the conceptual model and scheme of the experiment plan, with the number and sequence of individual experiments, with validity, reliability, and replicability established. Analyzing Designs involves considering the results of experimentation obtained, with repetitions, preliminary results, and deviations. Displaying Plots is a stage of automatic creation of outline diagrams of analyzed experiment plans, which facilitate the consideration and interpretation of the results obtained. Optimal Design is the process of refining a plan by reducing or increasing the number of experimental cycles with factors, surfaces, or mixtures. Response Optimization is the process of obtaining the optimum result of experimentation on an initially constructed optimization diagram. Experimenting involves many different types, but today they have mostly applied: Factorial designs, Response surface designs, Mixture designs, and Taguchi designs, (Bisgaard, 2008; Box, Hunter, & Hunter, & Stuart, 2005; Levin & Ramsey & Smodz, 2018; Montgomery & Runger, 2010; Nelson, 2004).

This paper presents the application of the methodology of experimentation with designing, analyzing and constructing experimental plans, optimizing experimental plans and optimizing the results of experimentation, on one specific example of a chemical reaction from practice. An especially recommended selection of experimentation procedures was used, depending on the set requirements and needs, using the Minitab® computer commercial software.

METHODOLOGY

The experimentation methodology covers the known basic elements of experimentation, the main types of experimentation, and my own recommended choice of the type of experimentation required. The basic elements of experimentation are based on the known concepts: input and output quantities, input size levels, the interaction of input quantities, the number of experiments required, the schema of the planned experiment, the model and the matrix of experimentation, as well as variance and regression, analyzes. The input magnitudes of experimentation are independent Factors (x_i: i =

1,2, n) with main effects A, B, .., N, e.g. process temperatures (A), which have several different values, e.g. Factor levels 90 ℃ and 100 ℃. The output magnitudes of experimentation are the dependent magnitudes of experimentation, e.g. Response (y) 98.7 ℃ with a confidence risk of α= 0.05. The interactions of input quantities are the interaction effects of particular factors, e.g. with the main effects of the interactions AB, AC, BC. The number of experiments required depends on the number of factors and their levels, e.g. the total number of experiments for 3 two-level factors is 2^3 = 8. The design point scheme is a 2- or 3-dimensional representation of the experiment points. The model is a mathematical representation of the relationship between experimentation results and factor sizes. A design matrix is a matrix description of a plan used to analyze experimentation. The main types of experimentation are with factors, with result surfaces, with mixtures, and with Taguchi experimentation. Factorial designs give the estimated results of experimenting with individual values. Response surface designs give the estimated results of experimenting with a series of values, which are located on the result surfaces. Mixture designs give estimated results of experimenting with individual components of mixtures. Taguchi designs provide estimated experimentation results for individual input sizes, according to the findings of an engineer and statistician Dr. Genichi Taguchi (1924 - 2012). The recommended choice of experiment type in Figure 1 includes all the major types of experimentation (designing, analyzing, constructing a plan, optimizing the plan, and optimizing the experimentation response), as well as practical experimentation procedures. It starts with the required type and successive phase of experimentation in order to select the necessary practical experimentation procedure, (Burman, Robert & Alm 2010; Hani, 2009).

CREATING DESIGNS PHASE

In the Creating Designs phase, an ideal schematic of the experimental design model is designed, with the number and sequence of individual experiments, with established validity, reliability, and reproducibility. E.g. When considering the effect of 3 factors at two levels, on a certain chemical reaction, two more factors need to be added, using the following steps: Step 1 - (Stat> DOE> Factorial> Create Factorial Design), step 2 - Select a special generation of experiments with 3

2-level factors, step 3 - of the plan with $2^3 = 8$ experiments and full resolution, which has 0 center points per plan and 1 repetition at the center point of the plan, step 4 - Generator selection in Submitted (Adding Generator Generators) and Enter Generation Labels factors (D = AB, E = AC) and step 5 - analysis of the results obtained.
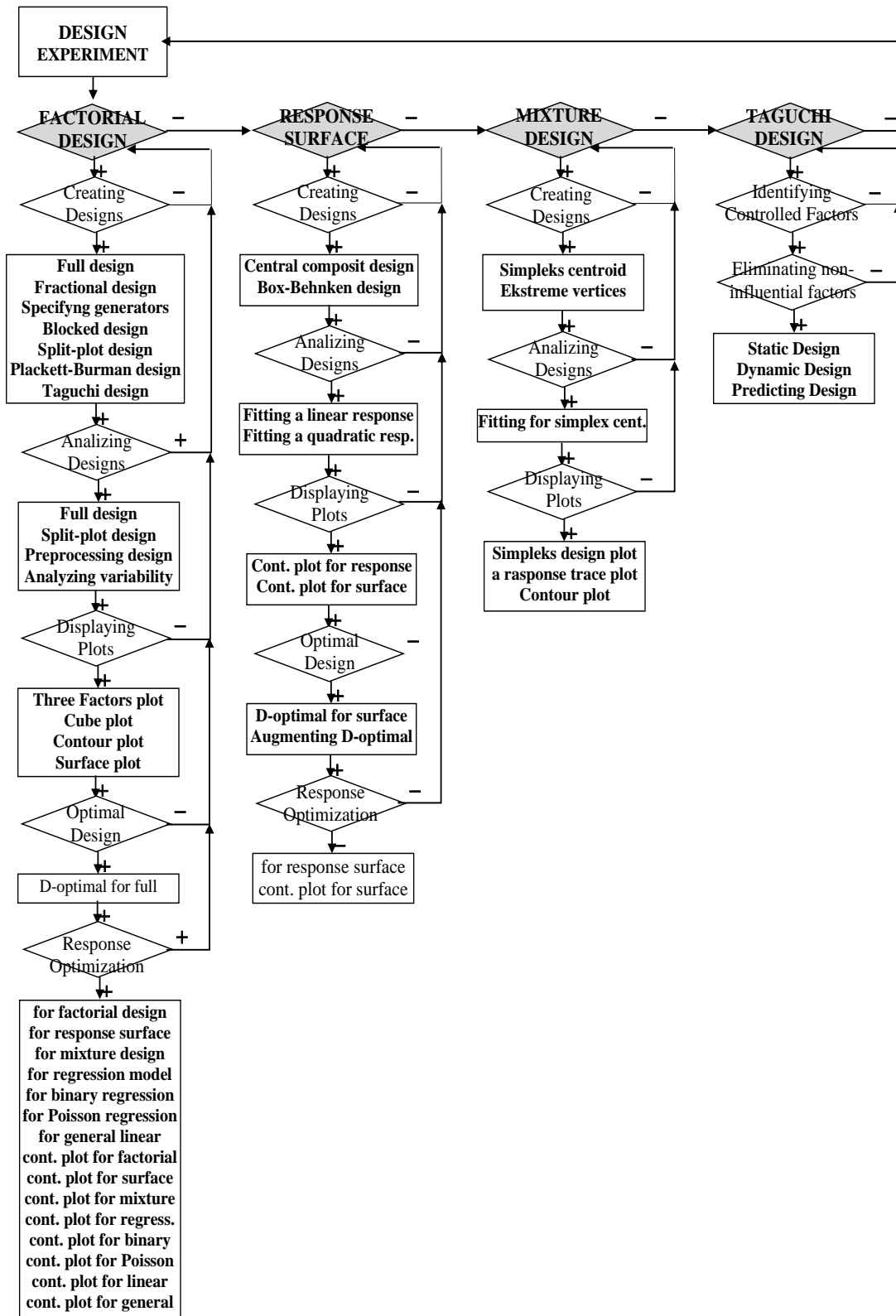

Figure 1: Recommended choice of experimentation procedures

The obtained special inspection of the generation in Figure 2 shows the following results: a) total number of factors (5), number of experiments performed (8), number of established test plans (1), basic plan (3.8), number of repetitions (1), total number of plan midpoints (0), applied to resolution (III), partial plan (1/4), b) factors generated (D = AB, E = AC) and c) estimation of substitution structure or influence of higher-order interaction factors. The experiment plan obtained in Figure 3 shows the following results: a) the sequence numbers of the experiments, b) the random order of the experiments, c) the number of central points of the plan, d) the number of plans of the identified experiments, and e) the code designation of all the factors considered, (Hazewinkel, 2001; Montgomery, & Runger, 2010; Sen, & Srivastava, 2011).

Fractional Factorial Design

```
Factors:  5    Base Design:        3, 8   Resolution:  III
Runs:     8    Replicates:         1      Fraction:    1/4
Blocks:   1    Center pts (total): 0

* NOTE * Some main effects are confounded with two-way interactions.


Design Generators: D = AB, E = AC


Defining Relation:  I = ABD = ACE = BCDE


Alias Structure

I + ABD + ACE

A + BD + CE
B + AD + CDE
C + AE + BDE
D + AB + BCE
E + AC + BCD
BC + DE + ABE + ACD
BE + CD + ABC + ADE


Design Table (randomized)

Run  A  B  C  D  E
  1  +  +  +  +  +
  2  -  +  -  -  +
  3  +  +  -  +  -
  4  -  +  +  -  -
  5  -  -  -  +  +
  6  +  -  +  -  +
  7  -  -  +  +  -
  8  +  -  -  -  -
```

Figure 2: Overview of special plan generation for 2-factor levels

↓	C1	C2	C3	C4	C5	C6	C7	C8	C9
	StdOrder	RunOrder	CenterPt	Blocks	A	B	C	D	E
1	3	1	1	1	-1	1	-1	-1	1
2	6	2	1	1	1	-1	1	-1	1
3	4	3	1	1	1	1	-1	1	-1
4	1	4	1	1	-1	-1	-1	1	1
5	8	5	1	1	1	1	1	1	1
6	2	6	1	1	1	-1	-1	-1	-1
7	5	7	1	1	-1	-1	1	1	-1
8	7	8	1	1	-1	1	1	-1	-1

Figure 3: Generating a plan for 2-factor levels

ANALYZING DESIGNS PHASE

In the experimental plan analysis phase, the results of experimentation with repetition are considered: full factor experiments, pre-processing reactions, and variability of the results obtained. E.g. When considering examples of the effect of 3 factors (time, temperature, catalyst) on a given chemical reaction, wherein 8 experiments only 16 cycles are performed in a single day, the following steps of analyzing full factor experiments are applied: Step 1 - Collect the results shown in figure 4 with Catalyst, Yield, and Cost results, step 2 - Select a program (Stat> DOE> Factor> Analyze Factor Design), step 3 - enter the data type (Yield) in the

(Responses) field, step 4 - Select (Charts) by entering type diagrams (Pareto, Normal, Half Normal) and 5 - analyzing the results obtained.

↓	C1	C2	C3	C4	C5	C6	C7-T	C8	C9
	StdOrder	RunOrder	CenterPt	Blocks	Time	Temp	Catalyst	Yield	Cost
1	4	1	1	1	50	200	A	48.4665	31.7457
2	3	2	1	1	20	200	A	45.1931	31.0513
3	8	3	1	1	50	200	B	49.2040	36.8941
4	6	4	1	1	50	150	B	45.5991	32.6394
5	1	5	1	1	20	150	A	42.7636	27.5306
6	2	6	1	1	50	150	A	44.7592	29.3841
7	7	7	1	1	20	200	B	44.7077	34.6241
8	5	8	1	1	20	150	B	43.3937	30.5424
9	12	9	1	2	50	200	A	49.0645	32.3437
10	14	10	1	2	50	150	B	45.1531	33.0854
11	16	11	1	2	50	200	B	48.6720	37.4261
12	15	12	1	2	20	200	B	45.3297	35.2461
13	10	13	1	2	50	150	A	45.3932	28.7501
14	13	14	1	2	20	150	B	43.0617	30.2104
15	9	15	1	2	20	150	A	43.2976	28.0646
16	11	16	1	2	20	200	A	44.8891	30.7473

Figure 4: Collected results of chemical reactions

The obtained overview of the complete plan of experiments in Figure 5 shows the results of a) the results of the analysis of variance: degrees of freedom (DF), sequential sums (Seq SS), adjusted squared averages (Adj MS), F- and P- values for all model sources, where no significant differences were observed for main (linear) effects (P= 0.000 <0.05) and two-way interactions (P= 0.017 <0.05) and only one Time*Temp interaction (P= 0.003 <0.05), b) model results: standard deviation (S), adjusted and predicted coefficients of obtained experiments (R-sq, R-sq adj, R-sq before) with all coded coefficient factors and c) regression of functions and structure of factor replacement:

Yield=39.48−0.1026Time+0.01502Temp+0.49Catalyst+0.001150Time*Temp−0.0029-Time*Catalyst−0.00281Temp*Catalyst+0.000031 Time*Temp*Catalyst.

The obtained Pareto diagram, shown in Figure 6, shows three significant influencing effects of factors, which satisfy the limit of standardized effects (2.36): time (A), temperature (B), and time-temperature interaction AB), (Ghosh, & Rao, 1996; Montgomery, 2013; Sifri, 2014; Spall, 2010).

Screening design results are obtained after performing full factor experiments when the variability of the results can be analyzed and when certain factors are believed to have a stronger influence. E.g. In the example shown, chemical reaction data were collected using the following steps: Step 1 - data collection, step 2 - choice (Stat> DOE> Factorial> Preprocess Responses for Analyze Variability), step 3 - input of type of standard deviation (Compute for repeat responses across rows) in the field (Standard deviations to use for analysis), step 4 - entry ('Yield_1' - 'Yield_8'.) in the field (Repeat responses across rows), step 5 - the input of the data type (StdYield) in the field (Store standard deviations in), step 6 - entering a data type (NYield) in the field (Store number of repeats in) and step 7 - analyzing the results obtained.

The results of the experimental design of the experiments in Figure 7 show the following results: a) standard deviation of column yield repeats (StdYield), b) a number of repeats in column (NYield), and c) 8 standard deviations and 8 number of repeats, with one standard deviation and the number of repetitions for each combination of factor settings, in the order where that combination first appears, filling in the remaining rows with the missing symbol (*), (Bisgaard, 2008; Montgomery, Douglas, 2013; Pronzato, 2008).

Factorial Regression: Yield versus Blocks, Time, Temp, Catalyst

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	8	68.7449	8.5931	58.93	0.000
Blocks	1	0.0374	0.0374	0.26	0.628
Linear	3	65.6780	21.8927	150.15	0.000
Time	1	35.0328	35.0328	240.27	0.000
Temp	1	30.5405	30.5405	209.46	0.000
Catalyst	1	0.1047	0.1047	0.72	0.425
2-Way Interactions	3	3.0273	1.0091	6.92	0.017
Time*Temp	1	2.9751	2.9751	20.40	0.003
Time*Catalyst	1	0.0222	0.0222	0.15	0.708
Temp*Catalyst	1	0.0301	0.0301	0.21	0.663
3-Way Interactions	1	0.0021	0.0021	0.01	0.907
Time*Temp*Catalyst	1	0.0021	0.0021	0.01	0.907
Error	7	1.0206	0.1458		
Total	15	69.7656			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.381847	98.54%	96.87%	92.36%

Coded Coefficients

Term	Effect	Coef	SE Coef	T-Value	P-Value	VIF
Constant		45.5592	0.0955	477.25	0.000	
Blocks						
1		-0.0484	0.0955	-0.51	0.628	1.00
Time	2.9594	1.4797	0.0955	15.50	0.000	1.00
Temp	2.7632	1.3816	0.0955	14.47	0.000	1.00
Catalyst	0.1618	0.0809	0.0955	0.85	0.425	1.00
Time*Temp	0.8624	0.4312	0.0955	4.52	0.003	1.00
Time*Catalyst	0.0744	0.0372	0.0955	0.39	0.708	1.00
Temp*Catalyst	-0.0867	-0.0434	0.0955	-0.45	0.663	1.00
Time*Temp*Catalyst	0.0230	0.0115	0.0955	0.12	0.907	1.00

Regression Equation in Uncoded Units

Yield = 39.48 − 0.1026 Time + 0.01502 Temp + 0.49 Catalyst + 0.001150 Time*Temp
        − 0.0029 Time*Catalyst − 0.00281 Temp*Catalyst + 0.000031 Time*Temp*Catalyst

Figure 5: Chemical reaction regression analysis table

Pareto Chart of the Standardized Effects
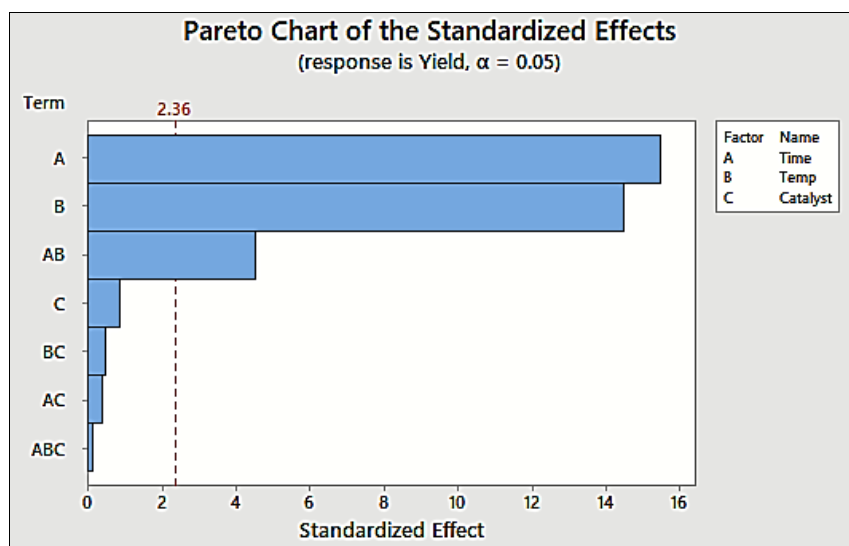(response is Yield, α = 0.05)

Figure 6: Pareto diagram with the influencing effects of a chemical reaction factor

Analyzing the variability of the results obtained is applied when certain factors are believed to have a stronger influence and is performed in two stages: 1) calculating least squares regressions to fit the reduced model and 2) analyzing the reduced model using maximum likelihood estimation to obtain the final regression coefficients. E.g. the example shown gives the following analysis of the variability of the results in phase 1 using the following steps: Step 1 - program selection (Stat> DOE> Factorial> Analyze Variability), step 2 - input (StdYield) in the field (Response, standard deviations), step 3 - (Terms) and input 2 in the field (Include terms from the model up through

order), step 4 - choice 4 Graphs and input (Pareto, Normal, Half Normal) in the field (Effects Plots)

and step 5 - analyzing the results obtained.

↓	C1	C2	C3	C4	C5	C6	C7-T	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17
	StdOrder	RunOrder	CenterPt	Blocks	Time	Temp	Catalyst	Yield_1	Yield_2	Yield_3	Yield_4	Yield_5	Yield_6	Yield_7	Yield_8	StdYild	NYild
1	1	3	1	1	20	150	A	42.7636	43.2976	43.2711	42.9325	43.2365	43.0210	43.6365	42.9112	0.2800	8
2	2	2	1	1	50	150	A	44.7592	47.3932	43.2094	41.8365	49.7236	49.7662	48.5013	48.4485	3.0456	8
3	3	4	1	1	20	200	A	45.1931	43.1891	44.8547	43.8015	44.8025	45.7207	46.3449	45.4934	1.0240	8
4	4	7	1	1	50	200	A	48.4665	53.0645	40.4579	59.4553	40.3418	39.8702	55.2384	56.4470	8.0317	8
5	5	8	1	1	20	150	B	43.3937	43.0617	43.5734	42.3529	44.0803	43.3958	43.1371	43.3953	0.4915	8
6	6	5	1	1	50	150	B	45.5991	42.1531	48.3112	40.7675	42.1684	51.2708	49.1959	49.1860	3.9723	8
7	7	1	1	1	20	200	B	44.7077	47.3297	42.3855	45.8798	42.2025	47.2578	46.2602	45.9203	2.0003	8
8	8	6	1	1	50	200	B	49.2040	41.6720	59.6118	60.0008	59.7650	38.0484	39.9795	38.8697	10.0303	8

Figure 7: Obtained results of the trial experiments

The obtained overview of the complete plan of experiments in Fig. 8 shows the results of a) the results of the analysis of variance (DF), (Seq SS), (Adj MS), F- and P values for all model sources, where significance for factors (Time, Temp) and not for interactions, b) model results: (S), (R-sq, R-sq adj, R-sq before) with coded factor coefficients and c) regression function and factor replacement structure:

Ln(StdYield)=−7.339+0.11482Time+0.03237Temp+0.377Catalyst−0.000268Time*Temp-−0.00620Time*Catalyst+0.000318Temp*Catalyst.

The obtained Pareto diagram shown in Figure 9 shows the results of the significant effects of factors (A, B, C), which satisfy the limit of standardized effects (12.71). At this point, it is observed that the model should be reduced using the least-squares regression method to obtain factors (Time, Temp, Catalyst) that should be retained in the model. Of course, this model is just one of the possible scaled-down models to choose from, so you may need to fit several models to find the right model.

Analyzing the experiment plan, using phase 2 least-squares regression estimation using reduced model analysis, using maximum likelihood estimation to obtain the final model coefficients, the following results are obtained using the following steps: Step 1 - Program Selection (Stat> DOE> Factorial> Variability Analysis), step 2 - (StdYield) entry in the (Answer, standard deviations) field, step 3 - choice (Option) and entry (Maximum likelihood) in the (Estimate) method field, step 4 - choice (Conditions) and scroll (BC) from (Selected Conditions) to (Available Conditions), step 5 - Selections (Charts) and Delete (Pareto, Normal, Half Normal) in the Table Effects field and Three in One in the Other Plans field and step 6 - analysis of the results obtained, (Hazewinkel, 2001; Sen, & Srivastava, 2011).

The obtained overview of the complete plan of experiments in Figure 10 shows the results of a) the results of the estimate (Effect), (Ratio Effect), (Coef), (SE Coef), Z- and P- values and VIF, b) the influence of the factor (Time) with the strongest effect (2.0365) since the effect ratio shows an increased standard deviation (7.6636) when the factor level (Time) changes, c) the factor (Temp) has the next strongest effect (− 1.1491) with the standard deviation increasing to (3.1552) when the factor (Temp) changes from low to high, d) the factor (Catalyst) has the smallest main effect (4300) and the effect ratio shows that the standard deviation increases to (1.5373) when the factor (Catalyst) changes from low to high level until the interactions are statistically significant and e) the regression function and structure of factor replacement:

Ln(StdYield)= −7.34+0.1148Time+0.03237Temp+0.432Catalyst−0.000268Time*Temp−0.00620Time*Catalyst.

```
Analysis of Variability: StdYield versus Time, Temp, Catalyst


Method

Estimation  Least squares


Analysis of Variance for Ln(StdYield)

Source                    DF   Adj SS    Adj MS   F-Value   P-Value
Model                      6  138.766    23.128    675.92     0.029
  Linear                   3  136.942    45.647   1334.07     0.020
    Time                   1  100.490   100.490   2936.88     0.012
    Temp                   1   31.974    31.974    934.45     0.021
    Catalyst               1    4.478     4.478    130.88     0.056
  2-Way Interactions       3    1.824     0.608     17.77     0.172
    Time*Temp              1    0.979     0.979     28.62     0.118
    Time*Catalyst          1    0.839     0.839     24.51     0.127
    Temp*Catalyst          1    0.006     0.006      0.18     0.746
Error                      1    0.034     0.034
Total                      7  138.800


Model Summary for Ln(StdYield)

       S    R-sq  R-sq(adj)  R-sq(pred)
0.184977  99.98%     99.83%      98.42%

Coded Coefficients for Ln(StdYield)

                         Ratio
Term             Effect  Effect      Coef  SE Coef  T-Value  P-Value   VIF
Constant                           0.7020   0.0188    37.35    0.017
Time             2.0371  7.6682   1.0185   0.0188    54.19    0.012  1.00
Temp             1.1491  3.1552   0.5745   0.0188    30.57    0.021  1.00
Catalyst         0.4300  1.5373   0.2150   0.0188    11.44    0.056  1.00
Time*Temp       -0.2011  0.8178  -0.1005   0.0188    -5.35    0.118  1.00
Time*Catalyst   -0.1861  0.8302  -0.0931   0.0188    -4.95    0.127  1.00
Temp*Catalyst    0.0159  1.0160   0.0079   0.0188     0.42    0.746  1.00


Regression Equation in Uncoded Units

Ln(StdYield) = -7.339 + 0.11482 Time + 0.03237 Temp + 0.377 Catalyst -
  0.000268 Time*Temp
                - 0.00620 Time*Catalyst + 0.000318 Temp*Catalyst


Alias Structure

Factor  Name

A       Time
B       Temp
C       Catalyst


Aliases

I
A
B
C
AB
AC
BC
```

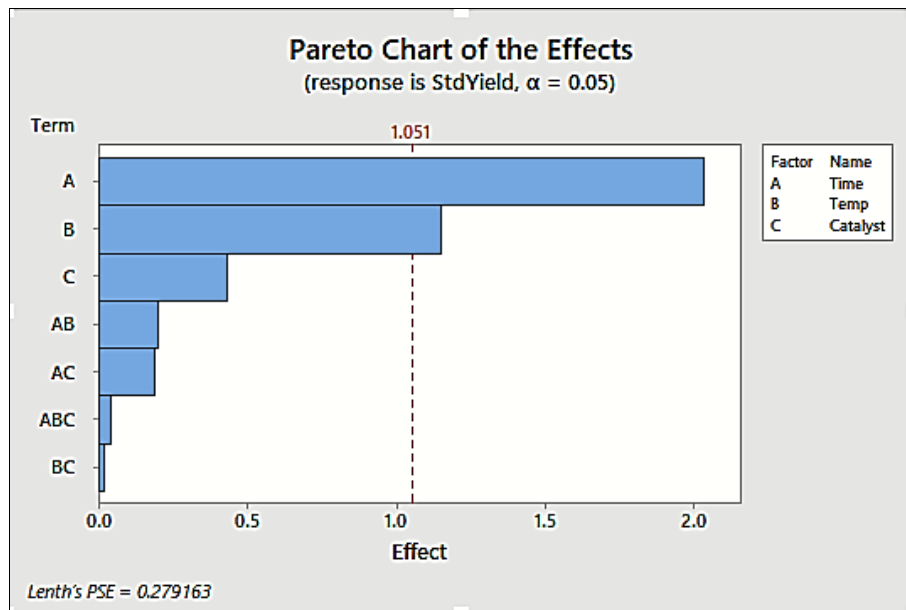*Figure 8: Obtained results of phase 1 variability analysis*

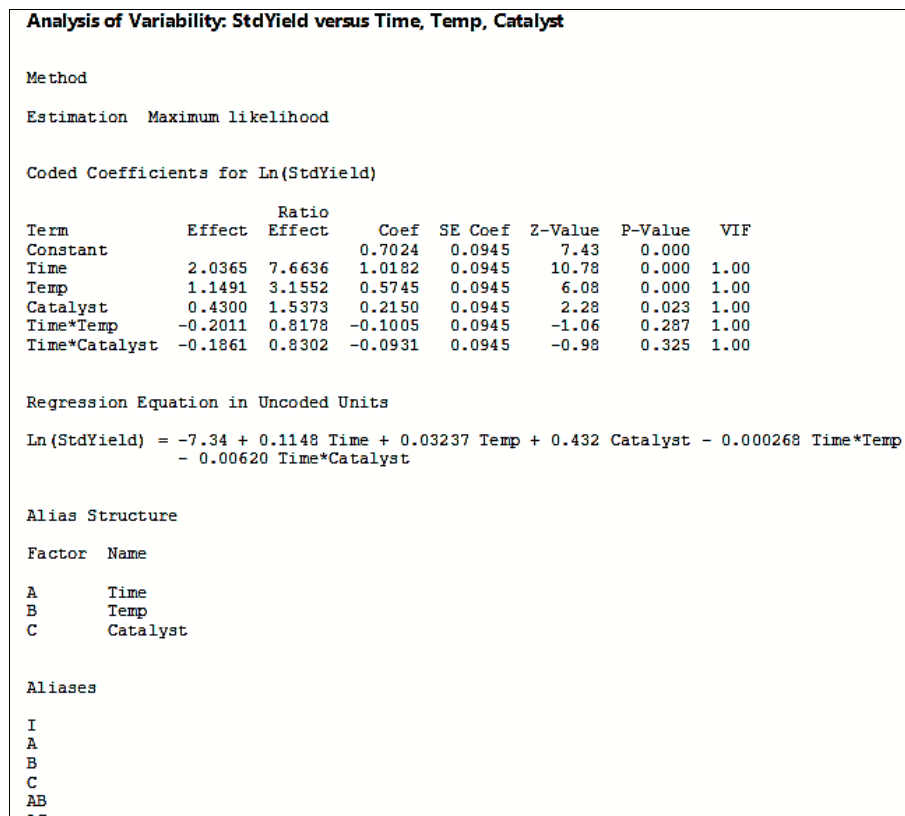*Figure 9: Obtained results analyzing phase 1 variability*

Figure 10: Obtained results of phase 2 variability analysis

DISPLAYING PLOTS PHASE

In the Displaying Plots phase, flow charts of the analyzed experiment plans are automatically created, which makes it easy to consider and interpret the results obtained. The relationship is shown: the main effect = f(factor level) of the interplay of factor effects, characterized by the main effects and their level. If the line is parallel to the axis (x) then there is no main effect and when the line is not then the main effect is present and the steeper the slope of the line, the greater the main effect. E.g. in the example shown, a diagram of the experiment plan is constructed using the following steps: Step 1 - data collection, step 2 - selection (Stat> DOE> Factorial> Factorial Plots),

step 3 - data type Yield in the field Responses, step 4 - choice Variables to Include in Plots and input of factors considered (Time, Temp, Catalyst) and step 5 - analyzing the results obtained.

The obtained diagram of the main effects of the interactions in Figure 11 shows the following results: (Time) and (Temp) factors have similar

effects that increase with the transition from low to high.

The obtained diagram of factor interactions (Temp) in Figure 12 shows the following results: the increase in the factor effect is larger when both the factor (Time) is larger (50) and the effect of the factor decreases when the factor is small (20).
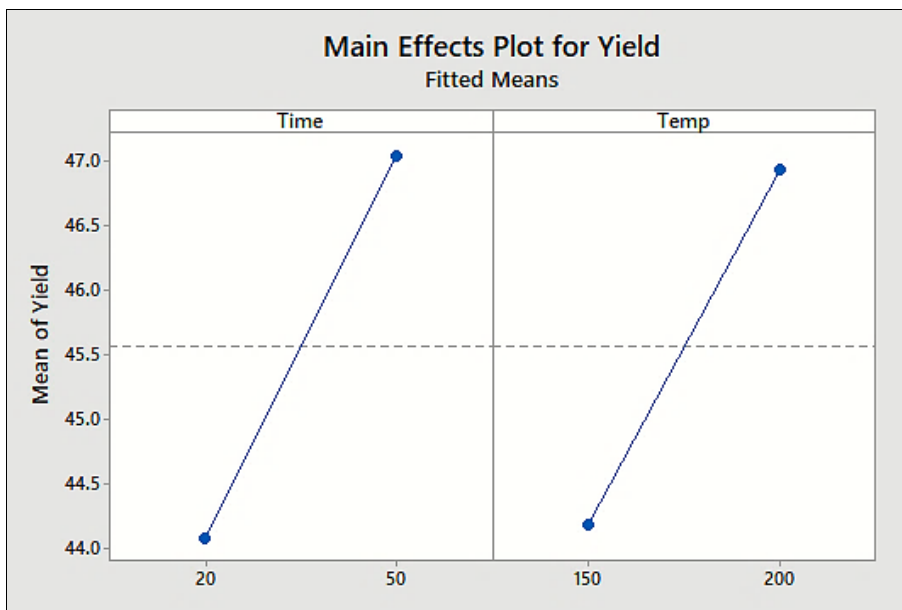
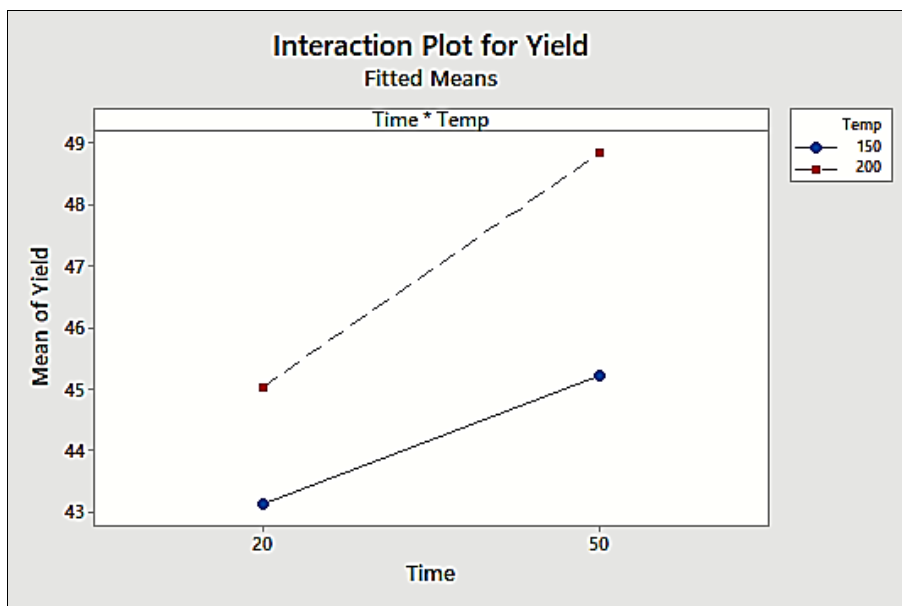

Figure 11: Diagrams of the main effects of interactions

Figure 12: Diagram of factor interactions (Temp)

OPTIMAL DESIGN PHASE

In the optimization phase of the experimental design (Optimal Design), a procedure for reducing or increasing the number of experimental cycles

with factors, surfaces or mixtures is applied, with the following criteria: minimizing deviations in regression coefficients (D-optimality) and even distribution of plan points in space (Distance-based optimality) spreads), which can be used when it is

not possible or desirable to select a model in advance. In the example of considering the influence of 3 factors with 2 levels on a certain chemical reaction, there is no such need, so a boundary contour plan diagram can be constructed. E.g. In the example shown, there is a significant interaction between reaction time and temperature, so that the dependence of the experiments on the input factors with the maximum reaction results and the minimum cost is determined by applying the following steps in phase 1): Step 1 - data collection, step 2 - Program selection (Stat > DOE> Factorial> Overlaid Contour Plot for Catalyst A, step 3 - Input (Cost) and (Yield) in Field (Selected), Input (Time) in Field (X-Axis) and Input (Temp) in Field (Y-Axis), step 4 - Choice (Contours) and enter preferred cost (28, 35) in the field (Low) and (High) and enter preferred yield (35, 45) in the field (Low) and (High), then in phase 2): step 5 - selection (Settings) and entry (A) in the (Setting) field for (Catalyst), step 6 - program selection (Stat> DOE> Factorial>

Overlaid Contour Plot) for catalyst B, step 7 - entry (Cost) and (Yield) in the field (Selected), entry (Time) in the field (X-Axis) and entry (Temp) in the field (Y-Axis), step 8 - selection (Contours) and entry (28, 35) in the field (Low) and (High) and input (35, 45) in the field (Low) and (High), step 9 - selection (Settings) and entry (B) in the field (Setting) for (Catalyst) and step 10 - analyzing the results obtained, (Chernoff, 1972; Fedorov, 1972; Goos, 2002; Goos, Jones & Bradley, 2011; Kôno, 1962; Pronzato, 2008; Pukelsheim, 2006; Shah, Sinha, & Bikas, 1989).

The obtained optimization diagram for catalyst A in Figure 13 shows the following results: a) white area with results within which time factor (Time) and temperature (Temp) have maxima of chemical reaction results (Yield) and b) minimum cost (Cost) around the point of the diagram (Time = 25; Temp = 185).
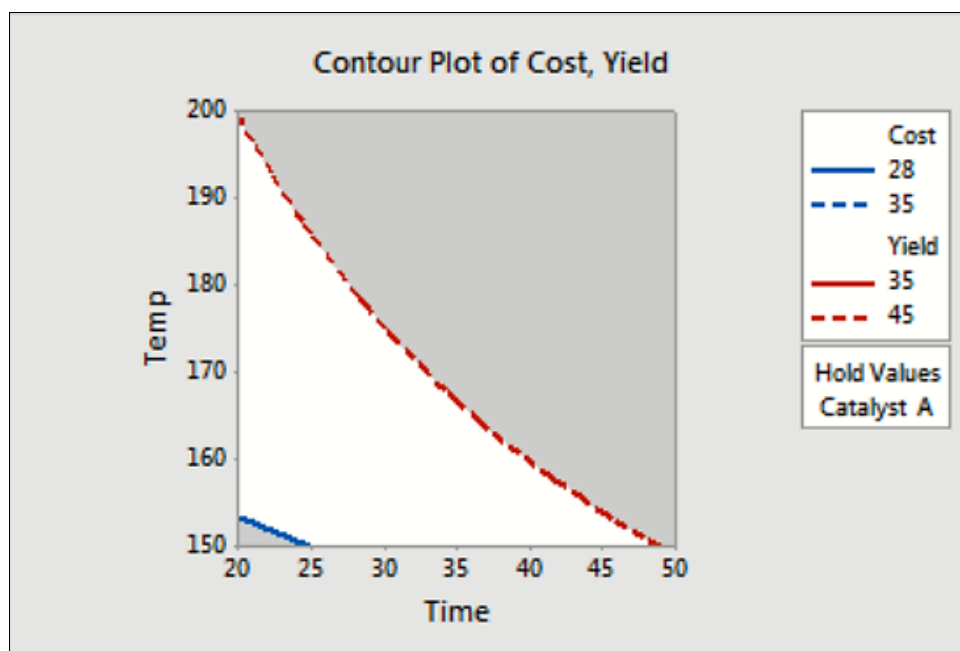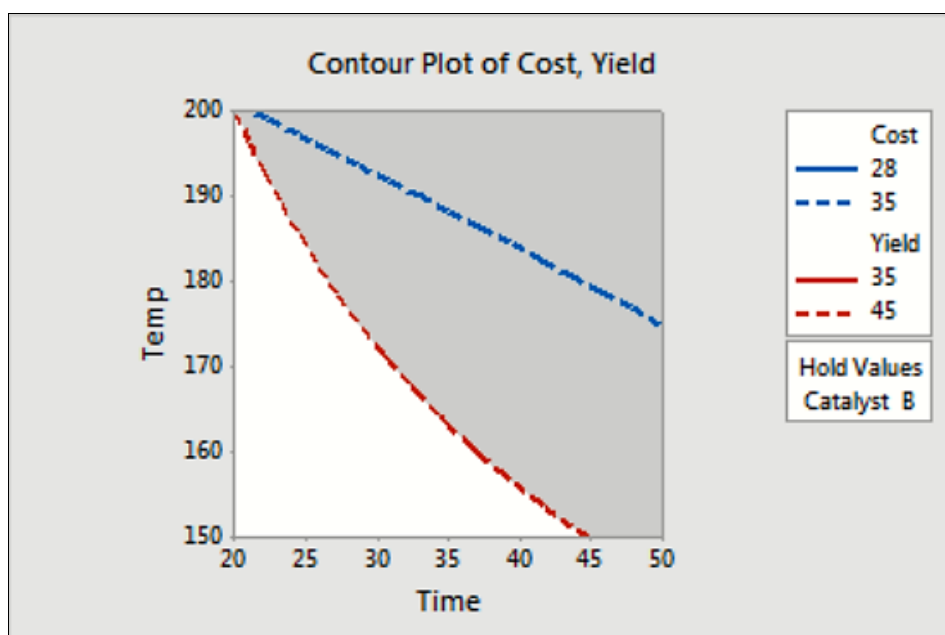

Figure 13: Contour plan optimization diagram for catalyst A

The obtained optimization diagram for catalyst B in Figure 14 shows the following results: a) white area with results within which time factor (Time) and temperature (Temp) have maxima of chemical reaction results (Yield) and b) minimum cost (Cost) around the point of the diagram (Time = 20; Temp = 200).

Figure 14: Contour plan optimization diagram for catalyst B

RESPONSE OPTIMIZATION PHASE

In the Response Optimization phase, the process of obtaining the optimal experiment result in the initially constructed optimization diagram is applied when the result lines are interactively changed to obtain the desired experiment results. The optimization of the results of the experimentation is based on the determination of the estimated values by using certain weights (Weight), according to the diagrams of the minimization, maximization, and determination of a certain value according to Figure 15, with lower and upper bounds. In the example shown, the results of the experiment are optimized with the maximum of the reaction results and the minimum required costs, using the following steps: Step 1 - data collection, step 2 - program selection (Stat> DOE> Factorial> Response Optimizer), step 3 - input (Minimize) into the field (Goal) for Cost and Maximize in Yield Goal, step 4 - Setup for Cost Costs (-, 28, 35) in Lower, Target, Upper) and for yield (Yield) input of values (35, 45, -) into the fields (Lower, Target, Upper) and step 5 - analyzing the obtained results.

The obtained optimal experimentation results in Figure 16 with the maximum of the chemical reaction results and the minimum of required costs show the results: a) applied parameters for (Cost= − 28, 35) and (Yield= 35, 45, −), b) solutions obtained Time= 46.3626 and Yield= 150 with catalyst A ic) composite desirability of 0.924449 which is quite close to 1, indicating favorable results overall.

The obtained interaction optimization diagrams in Figure 17 with individual factors and possible individual (d) and composite desires show the following results: a) an acceptable high composite desirability of D = 0.924449 which is quite close to 1, indicating favorable overall results, b) individual cost desirability (Cost) with maximum value y= 28.9132 (d= 0.86954) obtained in cross-section shown with a dashed line in the diagram, c) individual yield desirability (Yield) with maximum value y = 44.8282 (d = 0.86954) obtained in cross-section shown with the dashed line in the diagram and d) allowing adjustment of the factor settings of the initial solution by moving the (red) vertical lines.
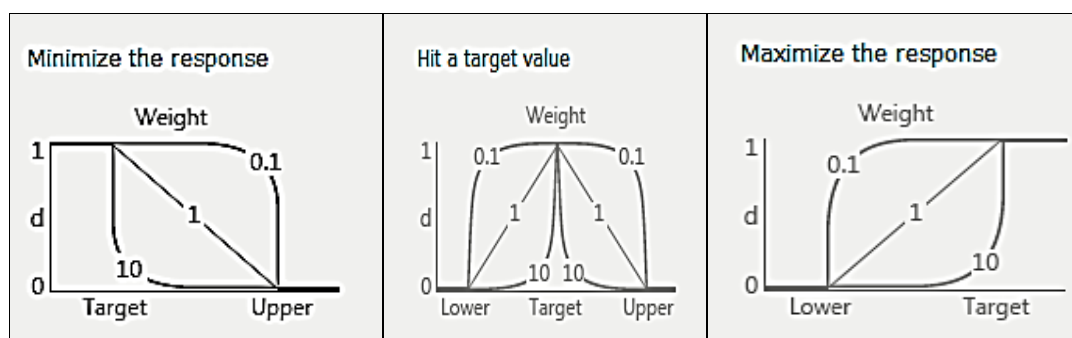
Figure 15: Determination of estimated values using certain weights

Figure 16: Optimal chemical reaction results obtained

If one now wants to reduce costs exactly by value (y= 28) by moving (red) vertical lines, then the optimization diagrams of interactions in Figure 18 are obtained with individual factors and possible individual (d) and composite desires, which show the following results: a) an acceptable high composite desirability of D= 0.9130, which is quite close to 1 indicating favorable overall results, b) individual desirability of the factor (Cost) with a maximum value of y≅ 28 (d= 1,000) obtained in cross-section with the dashed line in the diagram, c) the individual desirability of the factor (Yield) with the maximum value of y= 43.3356 (d= 0.83356) obtained in cross-section with the dashed line in the diagram, and d) allowing adjustments to the factory settings of the initial solution.

If a further increase in yield is exactly desired by the value (y= 45) by moving the (red) vertical line then an optimization diagram of the interactions in Figure 19 is obtained with the individual factors and possible individual (*d*) and composite desires showing the following results: a) acceptable high composite desirability of <u>D</u>= 0.9238 which is fairly close to 1, indicating favorable overall results, b) individual desirability of the factor (Cost) with a maximum value of *y*= 29.0263 (*d*= 0.85339) obtained in cross-section with the dashed line in the diagram shown, c) the individual desirability of the factor (Yield) with the maximum value of *y*≅ 45 (*d*= 1.0000) obtained in cross-section with the dashed line in the diagram shown, and d) allowing adjustments to the factory settings of the initial solution.
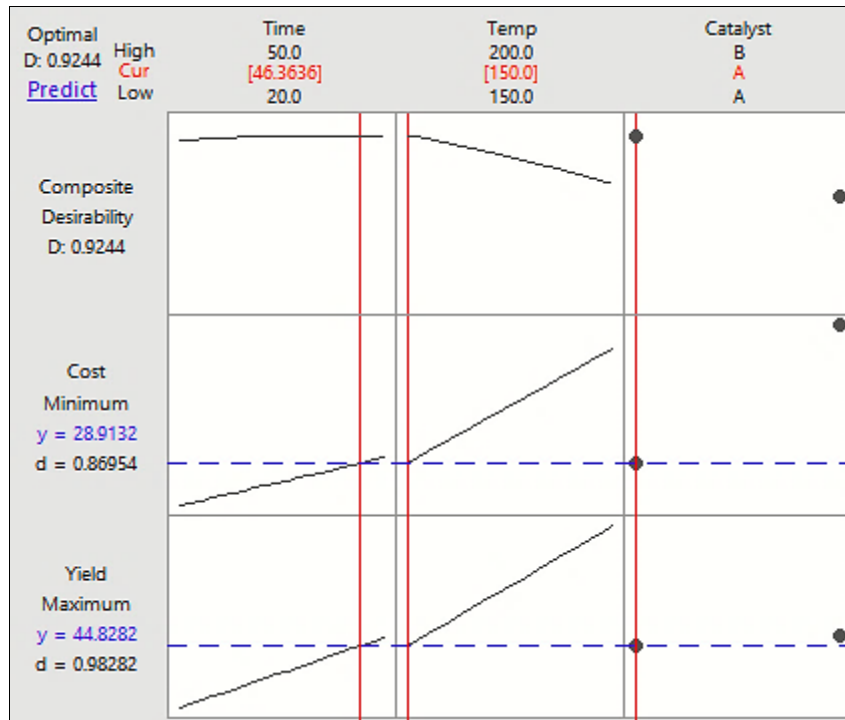
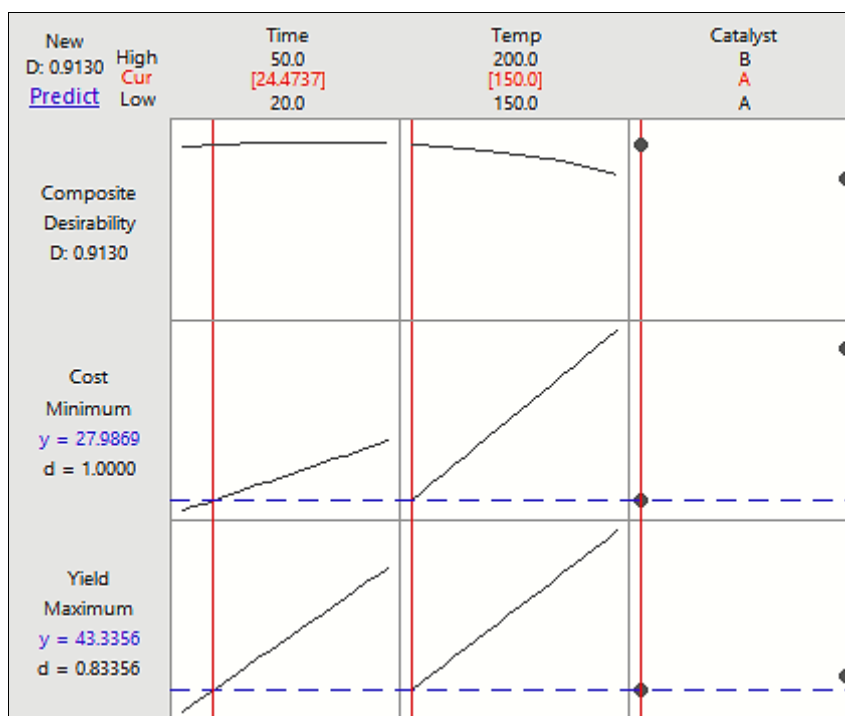*Figure 17: Interaction optimization diagrams obtained*

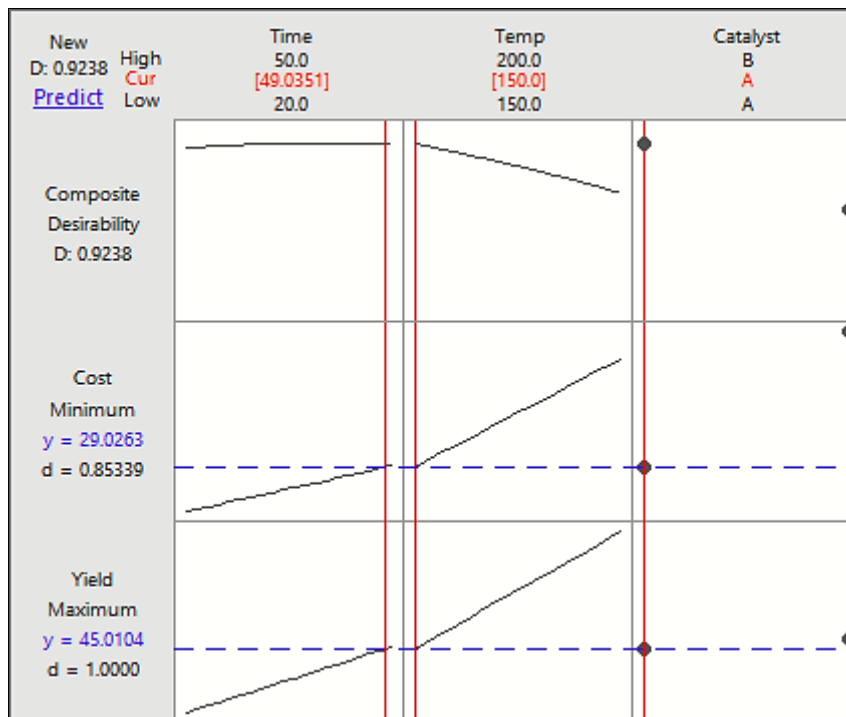*Figure 18: Cost optimization diagrams obtained for value (y = 25.00)*

Figure 19: Obtained yield optimization diagrams (y = 45.00)

RESULTS AND CONCLUSIONS

All the necessary results of the experimentation were obtained. In the design phase of the experimental design, in the case of a certain chemical reaction, with 3 factors (A, B, C) and 2 levels each, resolution (III) was applied and two more factors were generated (D = AB, E = AC) so that obtained from the partial plan (1/4) with 8 experiments. During the analysis phase of the experiment plan, we obtained: a review of the analysis of variance with suitable F- and P- values for all model sources, model results, favorable coefficients of determination of the obtained experiments (R-sq = 98.54%, R-sq adj = 96.87 %, R-sq pred = 92.26%,) and it was observed that all three significant influencing factors were: time (A), temperature (B), and the time-temperature interaction AB, satisfy the limit of standardized effects (2,36). The results of the experimental design of the experiments gave: good model coefficients: (R-sq = 99.98%, R-sq adj = 99.83%, R-sq pred = 98.42%, coded factor coefficients, regression function and the obtained results of the analysis of variability showed that the strongest effect (2.0365) had a factor (Time), then (1.1491) a factor (Temp) and the smallest (0.4300) factor (Catalyst). The diagrams show that factors (Time) and (Temp) have similar effects, which increase with the transition from low to a high level and that the increase in the effect of factor (Time*Temp) is larger when the factor is larger and the effect of the factor decreases when the factor is smaller. The example of chemical reaction of 3 factors with 2 levels, does not require reduction or increase of the number of experimental cycles. In the phase of optimization of the experimental plan, diagrams of boundary contour plans are constructed, which in the white area contain the optimal factors (Time = 25) and (Temp = 150), which satisfy both the results of experimenting Maximum (Yield) and Minimum (Cost). Finally, in the optimization phase of the experimentation results, the optimum results obtained with catalyst A: Cost minimum (28.9132) and maximum Yield (44.8282) with factors Time = 46.3626 and Temp = 150. If further reduction of costs is desired to a lower value (28) then moving the (red) vertical lines yields a minimum cost (Cost = 27.9869) and a maximum yield (Yield = 43.3356) with the factors Time = 24.4737 and Temp = 150, but if desired further increasing the yield to a higher value (Yield = 45), then moving the vertical lines yields a minimum (Cost = 29.0263) and a maximum (Yield = 45.0104) with factors Time = 49.0351 and Temp = 150.

REFERENCES

Bisgaard, S. (2008). Must a Process be in Statistical Control before Conducting Designed Experiments? *Quality Engineering, ASQ, 20*(2), 143 - 176.

Box, G., Hunter, E., William, P., Hunter, G., & Stuart J. (2005). *Statistics for Experimenters: Design, Innovation, and Discovery (2 ed.)*. Hoboken, N.J.: Wiley. ISBN 978-0471718130

Burman, L., E.; Reed, R.W., & Alm, J. (2010). A call for replication studies. *Public Finance Review, 38*(6), 787–793, doi: 10.1177/ 1091142110385210, retrieved 27 October 2011.

Chernoff, H., (1972). *Sequential Analysis and Optimal Design*. SIAM. ISBN 978-0-89871-006-9

Dr. Hani (2009). *Replication study*. Archived from the original on 2 June 2012. Retrieved 27 October 2011

Fedorov, V. V. (1972). *Theory of Optimal Experiments*. Academic Press

Ghosh, S., & Rao, C. R. (1996). *Design and Analysis of Experiments. Handbook of Statistics. 13*. North-Holland. ISBN 978-0-444-82061-7.

Goos, P. (2002). *The Optimal Design of Blocked and Split-plot Experiments*. Lecture Notes in Statistics. 164. Springer.

Goos, P., & Jones, B. (2011). *Optimal design of experiments: a case study approach*. Chichester Wiley. p. 304. ISBN 978-0-470-74461-1

Hazewinkel, M. (2001). *Regression analysis, Encyclopedia of Mathematics*. Springer. ISBN 978-1-55608-010-4

Kôno, K. (1962). Optimum designs for quadratic regression on k-cube (PDF). Memoirs of the Faculty of Science. Kyushu University. *Series A. Mathematics, 16*(2), 114–122. doi:10.2206/kyushumfs.16.114

Levin, D.M., Ramsey, P.P., & Smodz, R.K. (2018). *Applied Statistics: Using Microsoft Excel & Minitab*. Pearson.

Montgomery, D. (2013). *Design and analysis of experiments (8th ed.)*. Hoboken, NJ: John Wiley & Sons, Inc. ISBN 9781118146927

Montgomery, D.C., & Runger, G.C. (2010). *Applied Statistics and Probability for Engineers*. John Wiley & Sons.

Nelson, W. B. (2004). *Accelerated Testing—Statistical Models, Test Plans, and Data Analysis*. New York, John Wiley & Sons. ISBN 0-471-69736-2

Popović, B. (20180. Optimal costs of quality. *Journal of Engineering Management and Competitiveness - JEMC, 8*(1), 37-46.

Pronzato, L. (2008). Optimal experimental design and some related control problems. *Automatica, 44*(2): 303–325. arXiv: 0802.4381. doi: 10.1016/j. automatica 2007.05.016.

Pukelsheim, F. (2006). *Optimal Design of Experiments. Classics in Applied Mathematics*. 50 (republication with errata-list and new preface of Wiley (0-471-61971-X) 1993 ed.). Society for Industrial and Applied Mathematics. 454+xxxii. ISBN978-0-89871-604-7

Sen, A., & Srivastava, M. (2011). *Regression Analysis — Theory, Methods, and Applications*. Springer-Verlag, Berlin.

Shah, K., R. & Sinha, Bikas, K. (1989). *Theory of Optimal Designs*. Lecture Notes in Statistics. 54. Springer-Verlag. 171+viii. ISBN 978-0-387-96991-6

Sifri, J. (8 December 2014). *How to Use Design of Experiments to Create Robust Designs With High Yield*. youtube.com. Retrieved 11 February 2015

Spall, J. C. (2010). Factorial Design for Efficient Experimentation: Generating Informative Data for System Identification. *IEEE Control Systems Magazine, 30*(5), 38–53. doi: 10.1109/MCS.2010.937677

PLANIRANJE, ANALIZIRANJE I OPTIMIZACIJA EKSPERIMENATA

Istraživanje i razvoj (R&D) obuhvata sve inovativne aktivnosti organizacija koja se bave razvojem novih proizvoda (recikliranih materijala, hardvera, softvera, usluga) ili unapređenjem postojećih proizvoda. U SAD, procenat prihoda od istraživanja i razvoja prosečno iznosi 3.5% ali može dostići veće nivoe, čak 40%. Istraživanje i razvoj primenjuje naučne metode sa interaktivnim i cikličnim procesima kroz koje se informacije stalno ažuriraju i zamenjuju putem eksperimentisanja (dizajn eksperimenta). Eksperimentisanje je empirijski način arbitraže konkurenstkih modela ili hipoteza u cilju testiranja postojećih teorija ili novih hipoteza. Eksperiment ne sadrži samo dizajn već i sekvencijalne faze: stvaranje dizajna, analiza i konstrukcija planova eksperimenta, optizimacija plana eksperimenta i optimizacija odgovora eksperimenta. Eksperimentalni dizajni uključuju različite tipove, dok se danas najviše koriste faktoriski dizajni, prilagodljivi dizajni površina, dizajn površina i Taguchi dizajn. Ovaj rad prikazuje primenu metodologije eksperimentisanja sa stvaranjem, analizom i razvojem eksperimentalnih planova, optimizaciju eksperimentalnog plana, i optimizaciju eksperimentalnog odgovora, u jednom konkretnom primeru hemijske reakcije i primenom računarskog programa Minitab®.

Ključne reči: Planiranje eksperimenata.