

## Selection of Plan of Experiment by Statistical Analysis of the Parameters of Teletraffic Model with Voice Services

Iv. Balabanova<sup>1</sup>, G. Georgiev<sup>1</sup>, P. Kogias<sup>2,\*</sup> and St. Sadinov<sup>1</sup>

<sup>1</sup>Technical University of Gabrovo, Bulgaria

<sup>2</sup>Eastern Macedonia and Thrace Institute of Technology, Greece

Received 29 October 2016; Accepted 29 December 2016

### Abstract

A teletraffic model of voice services had been created. Simulations of the model of composite symmetric  $B_m$ , symmetric quasi -D optimal and non-symmetric quasi - D optimal types of experimental plans have been carried out. The plans of the experiment and the received experimental data are presented. Check on the suitability of the regression model of second degree has been done. The results of applied regression analysis are presented. The best plan of the experiment has been selected. The most suitable mathematical models on the average downtime of the system and the probability of losses have been found.

*Keywords:* plan of the experiment; teletraffic model; regression analysis; average downtime; probability of losses

### 1. Introduction

Originally the methods of the mathematical statistics have been used just at the end of the concrete experimental researches respectively during the processing, analysis and performance of the received test data. The object of research has been mostly simple systems in which processes and phenomenon depending on less number of variables could be defined. It was first shown the expediency of simultaneous modification of some or all control factors of an object at the end of the 20s in the last century. The mathematical statistics makes its first steps in the study of complex systems in which individual phenomena is difficult to be clearly distinguished. It appears the necessity to use methods for the optimal design of experiment and one of the most widely used is regression analysis. [1-5].

Each of the controllable factors of the studied object during the experiment can accept one or more vales, named levels. Combination of the levels of controllable factors responses to each possible status of the object. The set of all possible combinations of factors is the one which determines the total number of the possible tests. [5, 6]. One of the main tasks of planning of the experiment is to define which part of all possible tests has to be performed. Others are connected to quantify the impact of factors, finding mathematical models, determination of optimal flow conditions of a process and others. [7-8].

The goal of the present work is to select a plan of an experiment for conduction of a simulation process of teletraffic model of voice services for reduction of the number of tests and finding appropriate mathematical models representing completely experimental data based on the results obtained after application of the apparatus of

regression analysis.

### 2. Object of study

The object of study is a tele traffic model of voice services On-off+H/M/1/k with a tail FIFO, created using interrupted Poisson process and hyper exponential inflow. After generating of two independent inflows with different intensities of coming calls, the packages are processing by different way depending if they are data or voice information.

The following controllable factors and parameters of the object have been defined:

*Controllable factors:*

$x_1$  – intensity of incoming calls  $\lambda_1$ , s;

$x_2$  – time  $t_d$ , ms;

*Parameters of the object:*

$y_1$  – average downtime in the system  $W$ , ms;

$y_2$  – probability of losses  $B$ , %.

The received levels of variation of the controllable factors of teletraffic model of voice services are presented in Table.

**Table 1.** Levels of variation of the controllable factors

Level of variation	$X_1$	$X_2$	$x_1$ , ms	$x_2$
Low level	-1	-1	0,2	50
Average level	0	0	0,5	100
Upper level	+1	+1	0,8	150

### 3. Plans of the experiment with simulation of teletraffic model of voice services

In simulation of teletraffic model of voice services on a full factorial experiment (FFE) the total number of the possible

\* E-mail address: kogias@teiemt.gr

tests is 27. The following plans of experiment have been applied to achieve the set goal:

✓ *Symmetric compositional plan of type  $B_m$* , where  $m$  is the number of the controllable factors – plans of that type consist of a core and star points, as the core is formed (FFE) in type  $2^m$  or by fractional factorial experiment (FRFE) of type  $2^{m-p}$ , while the star points have a shoulder  $\alpha = +/-1$ . They become better when increase the number of factors. Their advantages are in their compositionality, i.e. the possibility to receive plans tier through rebuilding, and symmetry, relieving the calculation of the factors of the model and determining their importance [9, 12];

✓ *Symmetric quasi-D- optimal plan (Pesochinski's plan)* for second order polynomial in  $m = 2$  – those are non compositional plans. In many cases the use of non compositional plans is associated with a higher efficiency, for example when it is known that in the area of the experiment the surface of responses has non-linear characteristic. When  $m = 2$  the plans of that type are obtained after joining of two FFE plans of type  $2^2$  and four star points are added to them and another point at the center of the plan [10-12];

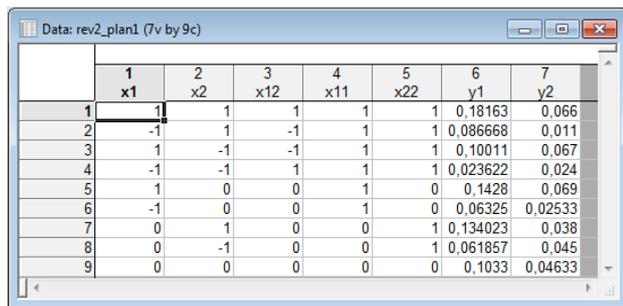
✓ *Non symmetric quasi-D-optimal plan when  $m = 2$*  – less number of tests is the main advantage compared to symmetric plans of experiment. The characteristic of those plans is their intensity when  $m = 3$ , i.e. inability to verify the adequacy of the model. Controlled tests are conducted to avoid that in one or more points, participating in the plan [9, 12].

The plans of the experiment are presented in Tables 2 to 4. Extended matrix of the experiment, corresponding to the common type of the wanted regression model of second degree could be seen in Figure1 to Figure 3:

$$y=b_0+b_1x_1+b_2x_2+b_{12}x_1x_2+b_{11}x_1^2+b_{22}x_2^2 \quad (1)$$

**Table 2.** Symmetric compositional plan of type  $B_2$  and experimental data for  $y_1$  and  $y_2$

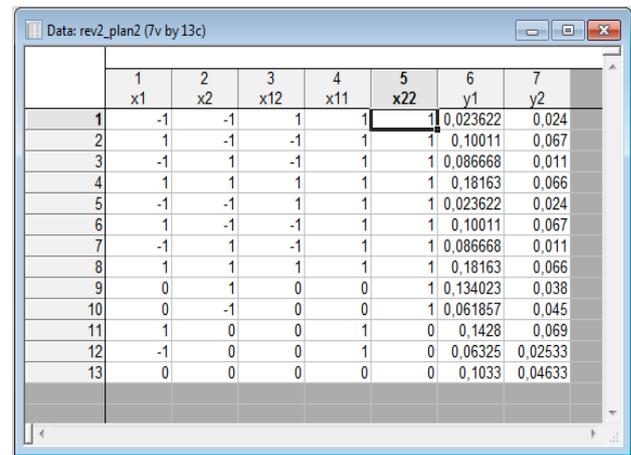
Test №	Coded variables		Natural variables		Output parameters	
	$X_1$	$X_2$	$x_1, ms$	$x_2$	$y_1, ms$	$y_2, \%$
1	+1	+1	0,8	150	0,18163	0,066
2	-1	+1	0,2	150	0,086668	0,011
3	+1	-1	0,8	50	0,10011	0,067
4	-1	-1	0,2	50	0,023622	0,024
5	+1	0	0,8	100	0,1428	0,069
6	-1	0	0,2	100	0,06325	0,02533
7	0	+1	0,5	150	0,134023	0,038
8	0	-1	0,5	50	0,061857	0,045
9	0	0	0,5	100	0,1033	0,04633



**Fig. 1.** Extended matrix of the experiment, obtained for Symmetric compositional plan of type  $B_2$ .

**Table 3.** Symmetric quasi-D-optimal plan (Pesochinski's plan) for second order polynomial in  $m = 2$  and experimental data for  $y_1$  and  $y_2$

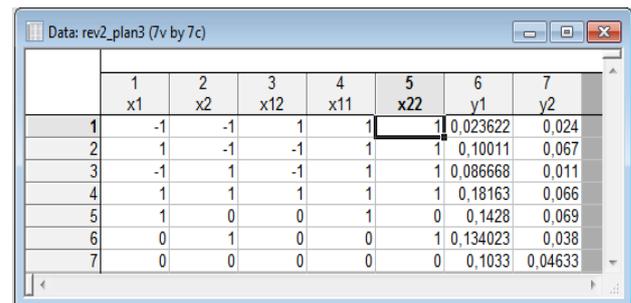
Test №	Coded variables		Natural variables		Output parameters	
	$X_1$	$X_2$	$x_1, ms$	$x_2$	$y_1, ms$	$y_2, \%$
1	-1	-1	0,2	50	0,023622	0,024
2	+1	-1	0,8	50	0,10011	0,067
3	-1	+1	0,2	150	0,086668	0,011
4	+1	+1	0,8	150	0,18163	0,066
5	-1	-1	0,2	50	0,023622	0,024
6	+1	-1	0,8	50	0,10011	0,067
7	-1	+1	0,2	150	0,086668	0,011
8	+1	+1	0,8	150	0,18163	0,066
9	0	+1	0,5	150	0,134023	0,038
10	0	-1	0,5	50	0,061857	0,045
11	+1	0	0,8	100	0,1428	0,069
12	-1	0	0,2	100	0,06325	0,02533
13	0	0	0,5	100	0,1033	0,04633



**Fig. 2.** Extended matrix of the experiment, obtained for Symmetric quasi-D-optimal plan when  $m = 2$

**Table 4.** Non symmetric quasi-D-optimal plan when  $m = 2$  and experimental data for  $y_1$  and  $y_2$

Test №	Coded variables		Natural variables		Output parameters	
	$X_1$	$X_2$	$x_1, ms$	$x_2$	$y_1, ms$	$y_2, \%$
1	-1	-1	0,2	50	0,023622	0,024
2	+1	-1	0,8	50	0,10011	0,067
3	-1	+1	0,2	150	0,086668	0,011
4	+1	+1	0,8	150	0,18163	0,066
5	+1	0	0,8	100	0,1428	0,069
6	0	+1	0,5	150	0,134023	0,038
7	0	0	0,5	100	0,1033	0,04633



**Fig. 3.** Extended matrix of the experiment, obtained for Non Symmetric quasi-D-optimal plan when  $m = 2$

3.1. Regression analysis of model (1) for parameters  $y_1$  and  $y_2$

Regression Summary for Dependent Variable: $y_1$ (rev2_plan1)						
R= ,99965451 R <sup>2</sup> = ,99930914 Adjusted R <sup>2</sup> = ,99815771						
F(5,3)=867,88 p<,00006 Std.Error of estimate: ,00206						
N=9	b*	Std.Err. of b*	b	Std.Err. of b	t(3)	p-value
Intercept			0,103148	0,001538	67,08322	0,000007
x1	0,753790	0,015175	0,041833	0,000842	49,67248	0,000018
x2	0,650878	0,015175	0,036122	0,000842	42,89090	0,000028
x12	0,067949	0,015175	0,004619	0,001031	4,47764	0,020767
x11	-0,000485	0,015175	-0,000047	0,001459	-0,03199	0,976488
x22	-0,053386	0,015175	-0,005132	0,001459	-3,51796	0,038974

a)

Regression Summary for Dependent Variable: $y_2$ (rev2_plan1)						
R= ,99864530 R <sup>2</sup> = ,99729244 Adjusted R <sup>2</sup> = ,99277985						
F(5,3)=221,00 p<,00048 Std.Error of estimate: ,00178						
N=9	b*	Std.Err. of b*	b	Std.Err. of b	t(3)	p-value
Intercept			0,046479	0,001328	35,00623	0,000051
x1	0,975402	0,030042	0,023612	0,000727	32,46800	0,000064
x2	-0,144586	0,030042	-0,003500	0,000727	-4,81279	0,017084
x12	0,101189	0,030042	0,003000	0,000891	3,36825	0,043467
x11	0,014589	0,030042	0,000612	0,001260	0,48560	0,660504
x22	-0,120524	0,030042	-0,005053	0,001260	-4,01186	0,027792

b)

Fig. 4. Results of regression analysis for parameters a)  $y_1$  и b)  $y_2$  in regard to Symmetric compositional plan of type B2

Regression Summary for Dependent Variable: $y_1$ (rev2_plan2)						
R= ,99975415 R <sup>2</sup> = ,99950836 Adjusted R <sup>2</sup> = ,99915719						
F(5,7)=2846,2 p<,00000 Std.Error of estimate: ,00148						
N=13	b*	Std.Err. of b*	b	Std.Err. of b	t(7)	p-value
Intercept			0,103139	0,001076	95,83966	0,000000
x1	0,756956	0,008381	0,042245	0,000468	90,32235	0,000000
x2	0,647382	0,008381	0,036130	0,000468	77,24768	0,000000
x12	0,074019	0,008381	0,004619	0,000523	8,83214	0,000048
x11	-0,000286	0,008456	-0,000033	0,000982	-0,03383	0,973956
x22	-0,044056	0,008456	-0,005118	0,000982	-5,20995	0,001239

a)

Regression Summary for Dependent Variable: $y_2$ (rev2_plan2)						
R= ,99904950 R <sup>2</sup> = ,99809989 Adjusted R <sup>2</sup> = ,99674268						
F(5,7)=735,40 p<,00000 Std.Error of estimate: ,00128						
N=13	b*	Std.Err. of b*	b	Std.Err. of b	t(7)	p-value
Intercept			0,046488	0,000929	50,02945	0,000000
x1	0,977774	0,016476	0,023967	0,000404	59,34698	0,000000
x2	-0,142788	0,016476	-0,003500	0,000404	-8,66669	0,000054
x12	0,109469	0,016476	0,003000	0,000452	6,64433	0,000292
x11	0,011730	0,016624	0,000599	0,000848	0,70561	0,503229
x22	-0,099293	0,016624	-0,005066	0,000848	-5,97289	0,000557

b)

Fig. 5. Results of regression analysis for parameters a)  $y_1$  and b)  $y_2$  in regard to Symmetric quasi-D-optimal plan when  $m = 2$

Regression Summary for Dependent Variable: $y_1$ (rev2_plan3)						
R= ,99989132 R <sup>2</sup> = ,99978266 Adjusted R <sup>2</sup> = ,99869594						
F(5,1)=920,00 p<,02503 Std.Error of estimate: ,00180						
N=7	b*	Std.Err. of b*	b	Std.Err. of b	t(1)	p-value
Intercept			0,102370	0,001543	66,35868	0,009593
x1	0,768874	0,015695	0,042630	0,000870	48,98952	0,012993
x2	0,647654	0,015695	0,035909	0,000870	41,26587	0,015424
x12	0,075593	0,014743	0,004619	0,000901	5,12753	0,122618
x11	-0,012417	0,016716	-0,001269	0,001709	-0,74277	0,593291
x22	-0,032527	0,016716	-0,003325	0,001709	-1,94581	0,302220

a)

Regression Summary for Dependent Variable: $y_2$ (rev2_plan3)						
R= ,99976967 R <sup>2</sup> = ,99953939 Adjusted R <sup>2</sup> = ,99723636						
F(5,1)=434,01 p<,03643 Std.Error of estimate: ,00120						
N=7	b*	Std.Err. of b*	b	Std.Err. of b	t(1)	p-value
Intercept			0,045709	0,001030	44,36164	0,014348
x1	0,957016	0,022848	0,024345	0,000581	41,88660	0,015196
x2	-0,143695	0,022848	-0,003655	0,000581	-6,28924	0,100383
x12	0,107023	0,021462	0,003000	0,000602	4,98667	0,125993
x11	-0,009210	0,024335	-0,000432	0,001141	-0,37846	0,769669
x22	-0,073168	0,024335	-0,003432	0,001141	-3,00667	0,204409

b)

Fig.6. Results of regression analysis for parameters a)  $y_1$  and b)  $y_2$  in regard to Non symmetric quasi-D-optimal plan when  $m = 2$

The results of regression analysis for the relevant plans of experiment are presented on Figure 4 to Figure 6. High factors of definition R<sup>2</sup> for parameters  $y_1$  and  $y_2$ , closed to 1, could be seen. The lowest values R<sup>2</sup>=0.99930914 and R<sup>2</sup>=0.997292 have been received in symmetric compositional plan of type B<sub>2</sub>, while the highest in non-symmetric quasi-D-optimal plan, respectively R<sup>2</sup>=0.99978266 and R<sup>2</sup>=0.99953939. The best indicators in regard to the criteria of Fisher F(5;7) = 2846.2 and F(5;718) = 735.4 and their relevant probabilities p < 0.0000 << 0.05, where 0.05 is the accepted level of importance  $\alpha$ , for parameters  $y_1$  and  $y_2$  are determined in symmetric quasi-D-optimal plan. For comparison the worst indicators are achieved in non-symmetric quasi-D-optimal plan, respectively p < 0.02503 < 0.05 and p < 0.032643 < 0.05.

In view of the results it can be considered that the most appropriate regression models on the parameters of tele traffic model have been obtained in symmetric quasi -D-optimal plan. According to the criteria of Fisher and their respective probabilities for the indicated plan of experiment the checked zero hypothesis H<sub>0</sub> is rejected as false, i.e. model (1) can be considered as adequate. In  $\alpha=0.05$  the factors of regression  $b_0=0.103148$ ,  $b_1=0.041833$ ,  $b_2=0.036122$ ,  $b_{12}=0.004619$  and  $b_{22}=-0.005132$  for parameters  $y_1$  and  $b_0=0.046488$ ,  $b_1=0.023967$ ,  $b_2=-0.003500$ ,  $b_{12}=0.003000$  and  $b_{22}=-0.005066$  for  $y_2$  have been determined as important.

3.2. Determination of the influence of the controllable factors on the parameters  $y_1$  and  $y_2$  for selected symmetric quasi -d- optimal plan

Regression Summary for Dependent Variable: $y_1$ (rev2_plan2)						
R= ,64888194 R <sup>2</sup> = ,42104777 Adjusted R <sup>2</sup> = ,30525733						
F(2,10)=3,6363 p<,06504 Std.Error of estimate: ,04246						
N=13	b*	Std.Err. of b*	b	Std.Err. of b	t(10)	p-value
Intercept			0,103117	0,024517	4,205941	0,001812
x2	0,647382	0,240614	0,036130	0,013428	2,690539	0,022679
x22	-0,044094	0,240614	-0,005123	0,027954	-0,183256	0,858259

a)

Regression Summary for Dependent Variable: $y_1$ (rev2_plan2)						
R= ,75698064 R <sup>2</sup> = ,57301970 Adjusted R <sup>2</sup> = ,48762364						
F(2,10)=6,7101 p<,01419 Std.Error of estimate: ,03647						
N=13	b*	Std.Err. of b*	b	Std.Err. of b	t(10)	p-value
Intercept			0,099727	0,021055	4,736561	0,000796
x1	0,756956	0,206635	0,042245	0,011532	3,663249	0,004366
x11	-0,006160	0,206635	-0,000716	0,024006	-0,029812	0,976804

b)

Fig.7. Exclusion of the factors a) x1 and b) x2 for parameter  $y_1$

Regression Summary for Dependent Variable: y2 (rev2_plan2)						
R= ,17303034 R <sup>2</sup> = ,02993950 Adjusted R <sup>2</sup> = -----						
F(2,10)=,15432 p<,85900 Std.Error of estimate: ,02414						
N=13	b*	Std.Err. of b*	b	Std.Err. of b	t(10)	p-value
Intercept			0,046887	0,013938	3,363840	0,007196
x2	-0,142788	0,311458	-0,003500	0,007634	-0,458451	0,656430
x2 <sup>2</sup>	-0,097729	0,311458	-0,004987	0,015892	-0,313780	0,760136

a)

Regression Summary for Dependent Variable: y2 (rev2_plan2)						
R= ,97777506 R <sup>2</sup> = ,95604407 Adjusted R <sup>2</sup> = ,94725288						
F(2,10)=108,75 p<,00000 Std.Error of estimate: ,00514						
N=13	b*	Std.Err. of b*	b	Std.Err. of b	t(10)	p-value
Intercept			0,043110	0,002967	14,52963	0,000000
x1	0,977774	0,066299	0,023967	0,001625	14,74788	0,000000
x1 <sup>2</sup>	-0,001509	0,066299	-0,000077	0,003383	-0,02276	0,982289

b)

Fig. 8. Exclusion of the factors a) x1 and b) x2 for parameter y2

The assessment of the level of influence of a control factor is based on the coefficient of determination R<sup>2</sup> as with the exclusion of the strongest factor, the coefficient will be with the lowest value. The process of exclusion of a concrete factor is realized through ignoring of those variables, in which this factor is included. The results after continuous exclusion of factors x1 and x2 for parameters y1 and y2 are presented in Figure 7 and Figure 8. Concerning the average downtime in the system W the lower coefficient of determination R<sup>2</sup>=0.42104777 in comparison with R<sup>2</sup>=0.57301970 shows, that the influence of the intensity of incoming calls λ<sub>1</sub> exceed that of time t<sub>d</sub>. The things related to the probability of losses B stay differently, where the influence of λ<sub>1</sub> considerably exceed t<sub>d</sub>, confirmed by the minimum value of R<sup>2</sup>=0.02993950.

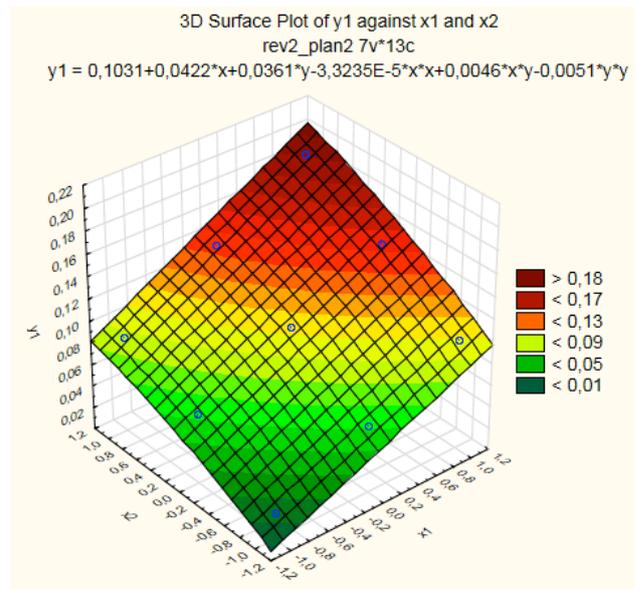
### 3.3. Analytical and graphical representation of a model for parameters y1 and y2 for selected symmetric quasi-d-optimal plan

The regression models of the parameters of tele traffic model of voice services – downtime in the system W and probability of losses B in analytical form, according to the importance of the coefficient of regression b<sub>i</sub>, can be expressed in final type, as follows:

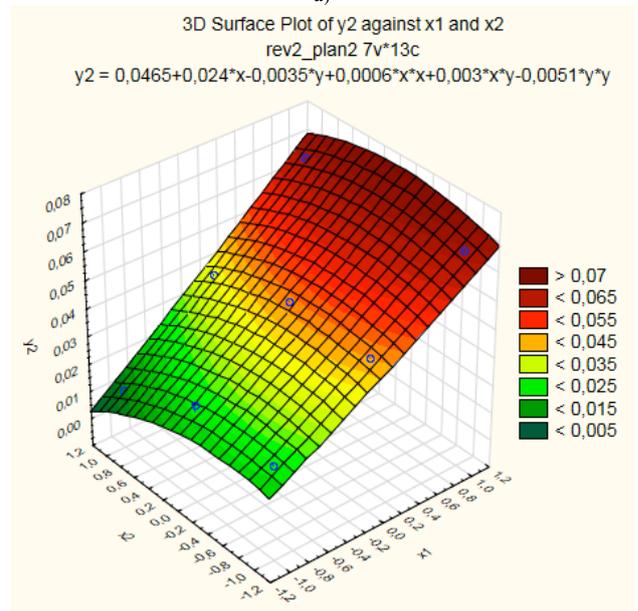
$$y_1 = 0.103139 + 0.042245x_1 - 0.036130x_2 + 0.004619x_1x_2 - 0.005118x_2^2 \quad (2)$$

$$y_2 = 0.046488 + 0.023967x_1 - 0.003500x_2 + 0.003000x_1x_2 - 0.005066x_2^2 \quad (3)$$

In three dimensional space the models describe surfaces  $y_1 = f(x_1, x_2)$  and  $y_2 = f(x_1, x_2)$  presented in graph in Figure 9. The lines of the same response  $y_1 = f(x_1, x_2)$  and  $y_2 = f(x_1, x_2)$ , shown in Figure 10, represent the intersections of those surfaces with equal parallel to plane  $x_1Ox_2$ . The specified figures show the areas of change of factors x<sub>1</sub> and x<sub>2</sub>, in which parameters y<sub>1</sub> and y<sub>2</sub> have the highest values, respectively in upper levels of x<sub>1</sub> and x<sub>2</sub> for parameter y<sub>1</sub> and upper levels of x<sub>1</sub> and all levels of variation of x<sub>2</sub> for parameter y<sub>2</sub>.

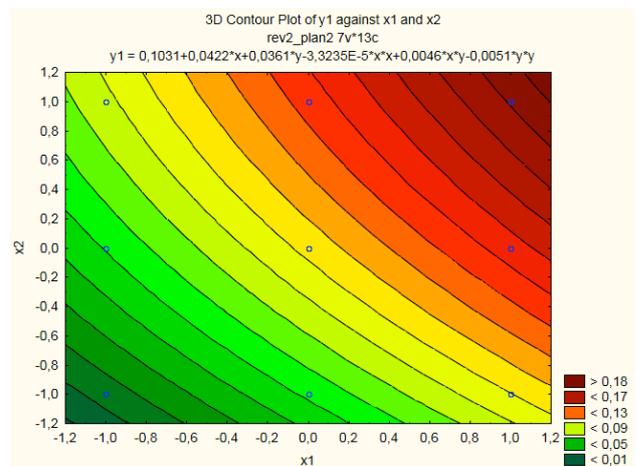


a)



б)

Fig. 9. Presentation of the surfaces of response a) y1= f(x1,x2) and б) y2= f(x1,x2)



a)

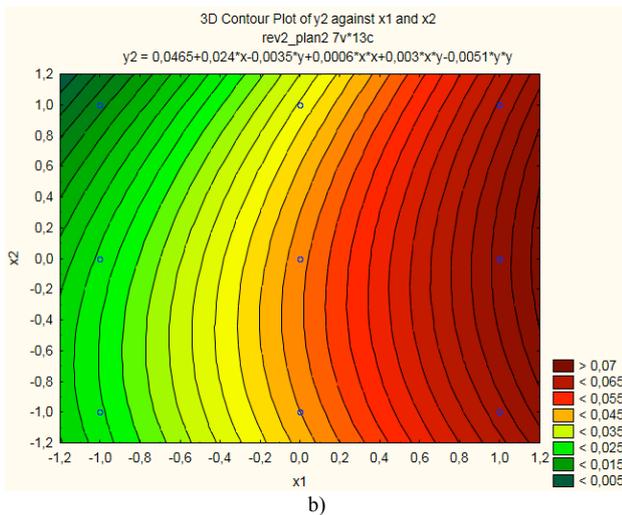


Fig. 10. Presentation of the lines of equal response a)  $f(x_1, x_2) = y_1 = \text{const.}$  and b)  $f(x_1, x_2) = y_2 = \text{const.}$

**3.4. Analyses of the residue of models for parameters  $y_1$  and  $y_2$  for the selected symmetric quasi-d-optimum plan**

A procedure for regression diagnostic through implementation of analysis of residue has been executed to check the relevance of the found models (2) and (3). The procedure comes to the design of normal probabilistic graphs depicting the values of the residues (Residuals)  $\epsilon_j$  when  $j=1,2,\dots,N$  compared to the values of the quantity  $t_{\alpha_j}$ , named  $\alpha_j$  quantiles of random quantity  $T$ , which distribution is the normalized random distribution. The residues are the differences between the experimental and the calculated values of the parameters  $y_1$  and  $y_2$ .

Based on the rule as far as plotted points  $(\epsilon_j, t_{\alpha_j})$  are placed in a straight line at an angle of  $45^\circ$ , so they will be normally distributed. For the presented on Figure 11 normal probability graphs of models (2) and (3) the location of  $\epsilon_j$  is closed to the straight lines. There are great similarities between the specified graphic terms, therefore regression models (2) and (3) can be defined as adequate.

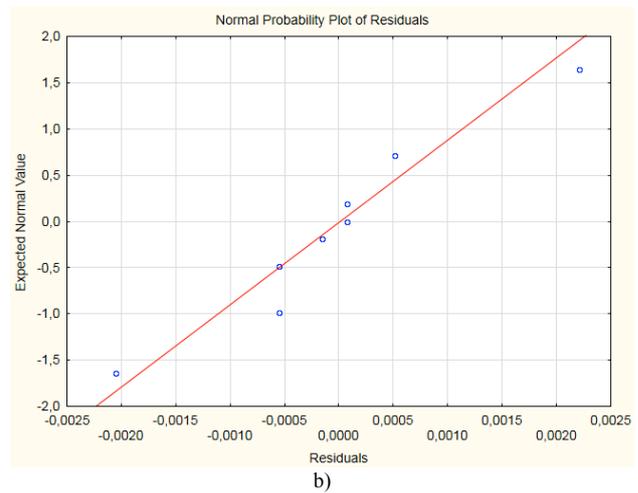
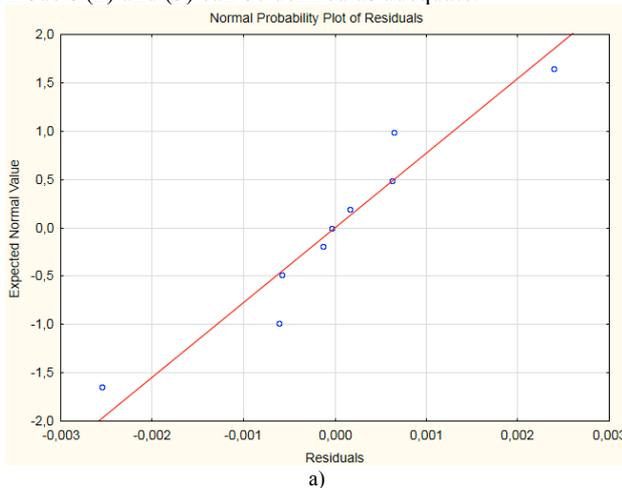


Fig. 11. Normal probable graphs of residue for a) model of  $y_1$  and b) model of  $y_2$

**4. Conclusions**

Symmetric compositional plans of experiments of type  $B_m$ , symmetric quasi-D optimum and non-symmetric quasi-D-optimum have been selected and implemented during simulation of the teletraffic model of voice services On-off+H/M/1/k with tail FIFO. Regression analysis of simulation results has been made for each plan. Quantitative assessment of the received regression indicators has been done and according to this a symmetric quasi-D-optimum plan of the experiment has been selected. Regression models describing completely experimental data have been found. It was found that the intensity of receiving the call  $\lambda_1$  and the time  $t_d$  influence in fairly equal degree to the average time of the stay in the system  $W$ , while the change of probability of losses  $B$  substantially is due to the influence of the first to that of the second factor.

**References**

1. Mitkov A. Theory experiment. Ed. "Dunav press" Ruse, 2011. (Bulgarian)
2. Adler, JP Introduction to experimental design. Moscow, Metallurgy 1969. (Russian)
3. Nalimov, V., N.Chernova. Statistical methods for planning a pilot experiment. Nauka, Moscow 1965. (Russian)
4. Fisher, R.A. The Place of the Design of Scientific Inference. – Sankhaya, ser.A, 27, 1965.
5. Telford J.K., A Brief Introduction to Design of Experiments. Johns Hopkins APL Technical Digest, vol. 27, pp. 224-232, 2007.
6. Gentle J.E., Theory of Statistics. pp.1-917, 2013.
7. Yan X., Gang Su X., Theory and Computing: Linear Regression Analysis. pp.1-348, 2009.

8. Sen A., Srivastava M., Regression Analysis: Theory, Methods and Applications. Springer Texts and Statistics, 1997.
9. Eriksson L., Johansson E., Kettaneh-Wold N., Wikström C., Wold S., Design of Experiments: Principles and Application. Umetrics Academy, pp.1-425, 2008.
10. Bojanov E. I. Vuchkov. Statistical modeling and optimization of multifactor objects. S., Technique, 1983. (Bulgarian)
11. Grobar, M.I., K.A. Krasnyanskaya. Application of mathematical statistics in pedagogical research. Non-parametric methods. M. Pedagogy 1977. (Russian)
12. Low L., D-optimum and quasi-D-optimum second order designs on a cube. Biometrika, **62**, 1975, pp. 335-340.