DOI: https://doi.org/10.53052/9788367652124.13

Vasilliy KISH[1], Miroslava KRESYAK[2]

Supervisor: Vitaly GERASIMOV[3]

OPRACOWANIE PROGRAMU KOMPUTEROWEGO DO PLANOWANIA EKSPERYMENTU W OPARCIU O RÓWNANIE REGRESJI MATEMATYCZNEJ

Streszczenie: Artykuł bada aspekty teoretycznego i praktycznego planowania eksperymentu opartego na pełnym eksperymencie czynnikowym. Opracowany program pozwala obliczyć współczynniki równania regresji dla różnych ilości czynników wejściowych, a także określić adekwatność modelu zgodnie z kryterium Fishera z graficzną konstrukcją krzywej odpowiedzi w trójwymiarowej formie.

Słowa kluczowe: współczynniki, program, równanie regresji

DEVELOPMENT OF A COMPUTER PROGRAM FOR PLANNING AN EXPERIMENT BASED ON A MATHEMATICAL REGRESSION EQUATION

Summary: The article exams aspects of theoretical and practical planning of an experiment based on a full factorial experiment. The developed program allows you to calculate the coefficients of the regression equation for different amounts of input factors, as well as to determine the adequacy of the model according to the Fisher criterion with the graphical construction of the response curve in a three-dimensional form.

Keywords: factors, program, regression equation.

1. Introduction

The most important task of methods for processing information obtained during an experiment is the task of constructing a mathematical model of the phenomenon,

[1] Mukachevo State University, Department of engineering, technology and professional education, specialty, specialty: light industry, k_pd@mail.msu.edu.ua
[2] Uzhgorod National University, Physical Department, specialty: bio-engineering, kresyak.miroslava@icloud.com
[3] PhD, Docent, Mukachevo State University, Department of engineering, technology and professional education, vitgerv@gmail.com

process, or object being studied. It can be used in process analysis and object design. It is possible to obtain a well-approximating mathematical model if an active experiment is purposefully used. Another task of processing the information obtained during the experiment is the optimization problem, i.e. finding such a combination of influencing independent variables that the selected optimality indicator takes an extreme value.

Currently, experimental planning methods are embedded in specialized packages widely available on the software market, for example: StatGrapfics [1], Statistica, SPSS, SYSTAT, etc.

The use of these programmes requires licences and a certain amount of time to learn the theory and practice of working with the product. Therefore, there is a need to develop a simple and convenient program for obtaining an experiment model based on a full-factor experiment.

2. Results and discussion

For a detailed study of the object of study, its detailed model is necessary. A suitable model is the "black box", which is quite common in various fields of research [2]. Its construction is based on the principle: optimal control is possible with incomplete information.

The black box diagram is shown in Fig. 1. The object of study corresponds to a rectangle. The outputs, indicated by arrows coming out of the object, correspond to the optimization parameters. Arrows entering an object -inputs- correspond to possible ways of influencing the object. In experimental design terminology, inputs are called factors. A factor is a measurable variable that at some point takes on a certain value and corresponds to one of the possible ways of influencing the object of study [3]. The number of possible impacts on an object is fundamentally unlimited.

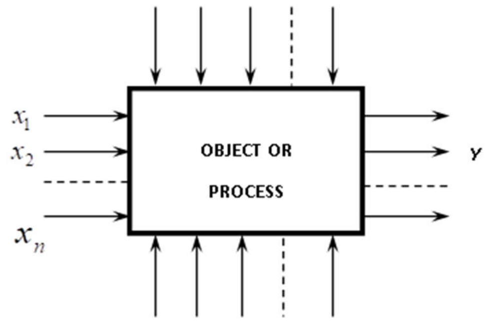

Figure 1. Model of "BLACK BOX"

The property (Y) of an object that interests us depend on several (n) independent variables (X_1, X_2, ..., X_n) and we want to find out the nature of this dependence – $Y=F(X_1, X_2, ..., X_n)$, about which we have just a general idea. The value Y is called the "response" and the dependence $Y=F(X_1, X_2, \ldots, X_n)$ itself is called the "response function". The response must be quantified. However, there may also be qualitative characteristics of Y. In this case, it is possible to use the rank approach. An example of a ranking approach is an assessment in an exam, when a complex set of information obtained about a student's knowledge is assessed with a single number.

Independent variables X1, X2, ..., Xn - otherwise factors, must also have a quantitative assessment. If qualitative factors are used, then each level should be assigned a number. It is important to select only independent variables as factors, i.e. only those that can be changed without affecting other factors. The factors must be clear.

The simplest case of linear representation of the response function is through a linear regression equation which has the following form:

$$\hat{Y} = b_0 + \sum_1^k b_i x_i + \sum_1^k b_{ij} x_i x_j \qquad (1)$$

where,

b_0 b_i and b_{ijk} - regression coefficient.

The authors faced the task of automating the calculation of regression coefficients using a software environment with the condition of constructing a response curve in spatial form [2-5].

The latter is important under the condition of selecting the input parameters in various levels of their change and fixing their influence on the response function. Of course, this task could be solved by more difficult specialized programs, given at the beginning of the article, but this development of the authors is more flexible and does not require licenses for the software product or input of intermediate data

$$b_0 = \frac{\sum_{u=1}^{n} \bar{y}_u}{n} \quad b_i = \frac{\sum_{u=1}^{n} x_i \bar{y}_u}{n} \quad b_{ij} = \frac{\sum_{u=1}^{n} x_i x_j \bar{y}_u}{n} \quad b_{ijk} = \frac{\sum_{u=1}^{n} x_i x_j x_k \bar{y}_u}{n} \qquad (2)$$

The authors faced the task of automating the calculation of regression coefficients using a software environment with the condition of constructing a response curve in spatial form.

The latter is important under the condition of selecting the input parameters in various levels of their change and fixing their influence on the response function.

Naturally, this task could be solved by more difficult specialized programs, given at the beginning of the article, but this development of the authors is more flexible and does not require licenses for the software product or input of intermediate data.

The development of the program was carried out in the DELPHI environment, which made it possible to use the suitable potential of the Pascal program language for solving such problems [6].

The procedure for working with the program is as follows: the user enters the input parameters (X-factors) and the range of their variation. After this, on the second tab, the calculation of the parameters of the plan model in the view of a multivariate experiment is activated and the computer will show the values of the regression coefficients (Fig.2). On the third tab, the program makes a 3D representation of the recall function. This option is the most effective from the point of view of analyzing the resulting model. Naturally, the program provides the rotation of the 3D curve along various geometric axes (Fig.3).

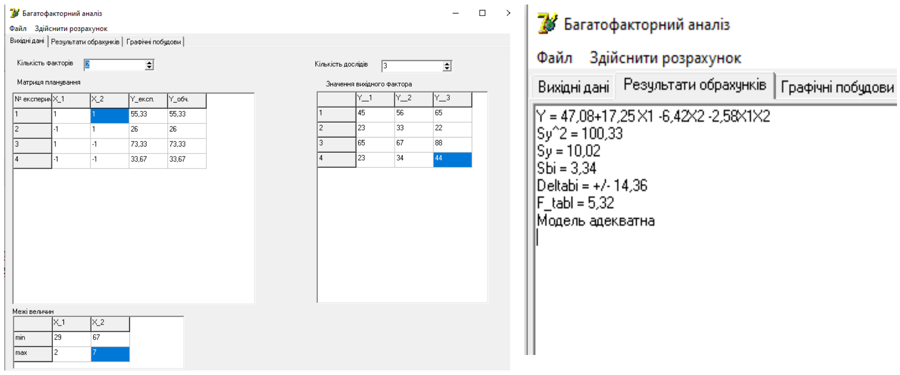

Figure 2. The FFE data entry window (left side) and the regression equation coefficients calculation output window (right side)

The algorithm of the program is based on the theory of full-factor experiment. The number of experiments is calculated according to the formula $N=2^k$, where k is the number of factors. The program provides a selection of the number of factors k=2 or k=3.

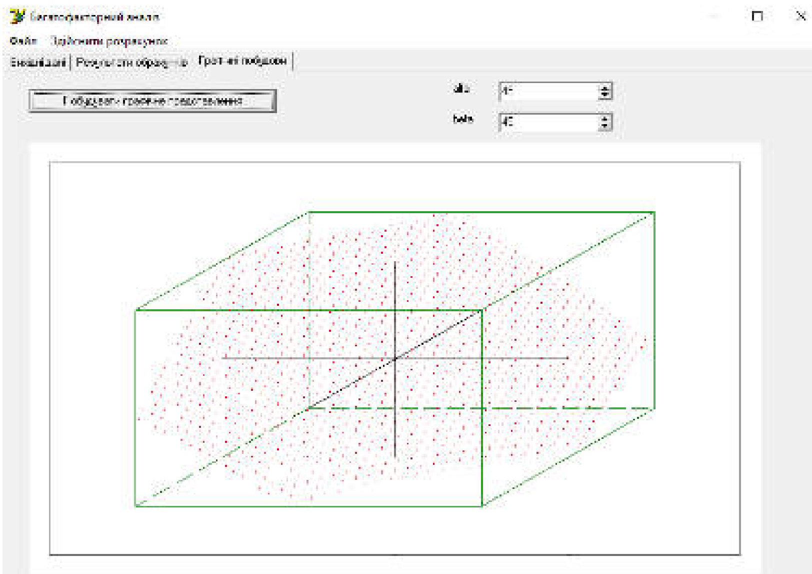

Figure 3. "Response curve" in volumetric form

Below is a part of the code for calculating the coefficients b_0 b_i and b_{ij} :

```
procedure GetKoefB_i(n,m:integer; var X:TMatr; var Y:TVect;
var bi:TVect; var b_ji:TMatr); // Rozrahunok b b123
 var i,j,l:integer;
begin
 SetLength(bi,n+1);
 bi[0]:=GetAverageValueFromVect(n,Y);
 for j:=1  to m do
```

```
begin
  bi[j]:=0;
 for i:=1  to n do
 begin
  bi[j]:=bi[j]+X[i,j]*Y[i];
 end;
 bi[j]:=bi[j]/n;
end;
SetLength(b_ji,m+1,m+1);
for i:=1  to m do
begin
  for j:=1  to i-1 do
  begin
    b_ji[j,i]:=0;
   for l:=1  to n do
   begin
    b_ji[j,i]:=b_ji[j,i]+X[l,j]*X[l,i]*Y[l];
   end;
    b_ji[j,i]:=b_ji[j,i]/n;
  end; end; end;
```

Thus, the value of the initial factor is entered in the amount of 4 or 8 (in accordance with the value of the plan matrix), with the selection of the number of repetitions from 1 to 5. This allows you to calculate the value of the variance of the original value and determine the adequacy of the resulting equation according to the Fisher test.

Testing the program with the previously obtained experimental data for k=3 (number of factors) gives us the result - a regression equation that represents the "response curve". The input factors for the given task (bonding two materials) were factor X1 - the temperature of the impact on the object of study (T, ^{0}C), X2 - the pressure on the object (F, N), X3 - time (W, min):

$$Y = 60 + 11,58X_1 + 13,5X_2 + 11,67X_3 - 1,58X_1X_2 - 4,92X_1X_3 - 1,17X_2X_3 \quad (3)$$

Testing the model by Fisher's criterion yields the following results: F calculated F = 0.04, and F tabulated F_{tab} = 4.49. Since F<Ft, the program concludes that the model is adequate.

In the case of a 2-factor experiment, the following plan matrix is obtained by the action of two factors (Table 1).

Table 1. Matrix of plan for k=2

N	X1	X2	Yexp	Ycalc
1	1	1	29,44	29,33
2	-1	1	24,33	24,33
3	1	-1	49	49
4	-1	-1	54,67	54,7

In this case, we will get the regression equation with the calculated values of the coefficients:

$$Y = 39,33 - 0,17X_1 - 12,5X_2 + 2,67X_1X_2 \quad (4)$$

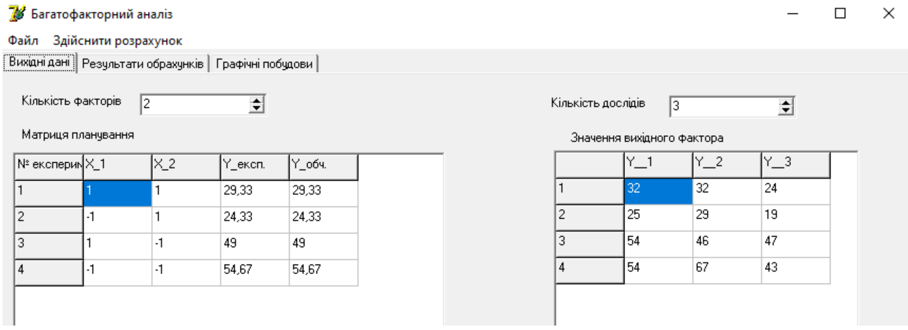

Figure 4. The FFE entry data and Y calculations for k=2

Conclusions

The developed program allows you to calculate the parameters of a full-factorial experiment during planning. The program allows you to obtain the values of regression coefficients and determine the adequacy of the model using the Fisher criterion with graphical construction of the response curve in volumetric form. The development will be useful when performing laboratory work for students, as well as for building a mathematical model of a scientific research object.

REFERENCES

1. Internet resourse. *https://www.statgraphics.com,* 01/11/2023.
2. NOTTELET B, PEKTOK E, MANDRACCHIO D, et al. Factorial design optimization and in vivofeasibility of poly("-caprolactone)-micro- and nanofiber-based small diameter vasculargrafts. J Biomed Mater Res A 89(2008), 865–875.
3. BOX, G., W. HUNTER, AND J. S. Hunter. Statistics for Experimenters. New York: Wiley, 1978.
4. BAILEY R. A., CHENG C. S., KIPNIS P.: Construction of trend-resistant factorial designs. Statistica Sinica 2(1992), 393–411.
5. BHOWMIK A., VARGHESE E., JAGGI S., VARGHESE C.: On the generation of factorial designs with minimum level changes. Communications in Statistics - Simulation and Computation 51(2020)6, 3400–3409.
6. ZAVADSKYI I. O.: Fundamentals of visual programming / I.O.Zavadskyi, R.I.Zabolotnyi. – K. : Publishing Group BHV, 2009. – 272 p. (in Ukrainian).