Original Article

CONSTRUCTION OF MULTIPLE REGRESSION EQUATIONS FOR STUDYING REGULARITIES OF CHEMICAL PROCESSES

Ph.D Q.S.ALIYEV[1,2], Ph.D R.J.HAJIYEVA[2], Ph.D E.N.AHMADOVA[2], Ph.D B.H.ASGEROVA[3], M.H.ALMAMMADOVA[3] İsmailov Alemdar Alesker [4]

Institute of Catalysis and Inorganic Chemistry named after Acad.M.Nagiyev, Azerbaijan[1,2]
Western Caspian University, Azerbaijan[2]
Azerbaijan State Oil and Industry University, Department of Computer Engineering, Azerbaijan[3] Azerbaijan State Agricultural University
Ganja, Azerbaijan[4]
chemproblem@mail.ru https://orcid.org/0000-0002-2237-3444
rena_gajieva@yahoo.com https://orcid.org/0000-0001-6507-26521
aesmiranq@gmail.com
bahar.askarova@asoiu.edu.az ORCID: https://orcid.org/0000-0002-8406-7178
almammadovamehriban@gmail.com, [4]https://orcid.org/0000-0002-6358-6171

DOI: 10.47750/pnr.2023.14.02.137

Abstract

The article presents the results of a study of some regularities of the process of catalytic oxidation in the gas phase of methanol to formic acid. For this, a multiple regression equation for the process was created, the coefficients of the equation were calculated, and their significance was determined. The adequacy of the equation was checked and a more sensitive parameter was identified. Using the Fisher criterion, the statistical significance of the equation was proved. Pair correlation indicators of the coefficients of the equation are also calculated. It was also found that the model parameters are statistically significant. It has been established that the greatest influence on the yield of the target product - formic acid is exerted by changes in the temperature of the reaction medium. This technique for constructing a multiple regression model is universal and can be applied to other chemical-technological processes. Also checked the significance of the parameters of the multiple regression equation. The statistical significance of the regression coefficients is confirmed.

1. Introduction

An analysis of patent and technical literature containing information about the named process indicates the absence of an efficient and cost-effective industrial technology for the process. Currently, there are two main methods for the industrial production of formic acid: the hydrolysis of methyl formate and the hydrolysis of formamide. The disadvantage of these liquid-phase processes is the multi-stage nature, capital and energy costs, the formation of a number of by-products, pollution of the atmosphere and wastewater. In this regard, the efforts of researchers working in this field are aimed at developing efficient highly active catalysts, conjugating technological stages, and creating a cost-effective technology for the production of formic acid (Aliyev A.M. et al., 2010a].

Mathematical modeling is a powerful tool for solving various problems that arise when optimizing chemical processes and maximizing profitability. Building a mathematical model of the process, choosing the most optimal parameters and expressing them as a mathematical function, can help increase the yield and reduce its cost. Therefore, the mathematical model should not only accurately describe the actual process, but also be simple and ensure the accuracy of calculations (Majidzade V.A. et al., 2021). Based on the study of the results of the process by methods of mathematical statistics, it is easy to determine the influence of the main parameters (concentration of the initial components, temperature, current density) on its course and the pattern of the reaction, as well as the optimal mode of its implementation. To confirm the obtained experimental results, a regression equation is constructed, the criterion of significance and adequacy is calculated (Javadova S.P. et al., 2021).

One of the main problems of industrial catalysis is the development of processes based on readily available and cheap raw materials. Due to the increase in the amount of alcohol-containing raw materials generated from food waste and in the woodworking industry, the study of the oxidation of aliphatic alcohols creates an objective prerequisite for the widespread use of these alcohols as raw materials in the industry, the acquisition of various valuable products. Heterogeneous catalytic oxidative conversion of aliphatic alcohols is considered the most promising method for obtaining aldehydes, acids and esters. It should be noted that the above substances can be used in obtaining the main product, depending on the nature of the catalyst and reaction conditions.

Despite the diversity and complexity of the tasks being solved, the methodology of reactor development of catalytic processes is based on a structural-hierarchical approach.

The essence of this approach lies in the consistent solution of problems at different levels: the kinetic equation, the mathematical model of the process occurring in the grain and the catalyst bed, the calculation of the reactor, the complete chemical-technological scheme. At this time, all the information obtained in the previous stage becomes part of the next stage.

2. Experimental part

As a result of systematic laboratory studies, an effective Pd^{+2} containing a modified zeolite catalyst was developed, with the use of which experiments were carried out on the production of formic acid by gas-phase oxidation of methanol.

This process is characterized by high yields of the target product with high selectivity, as well as ease of technological design, which creates a good prospect for its introduction into industry. For subsequent design and successful large-scale implementation of this process, it becomes necessary to develop a mathematical model and optimize the process based on the developed kinetic model and reaction mechanism. For subsequent design and successful large-scale implementation of this process, it becomes necessary to develop a mathematical model and optimize the process based on the developed kinetic model and reaction mechanism.

According to this method, a series of zeolite catalysts modified with palladium cations was prepared on the basis of synthetic zeolites NaX, NaY. The process of methyl alcohol oxidation on the synthesized catalyst was studied at a temperature of 85–105°C and a reaction volume rate of 900–3200 h^{-1}.

The adequacy of the model was confirmed by statistical analysis and differentiating experiments on a laboratory setup. The volume of the reaction zone of the reactor used in the laboratory was 6 cm^3. To evaluate (approbate) the resulting model, further experiments were carried out on an automated experimental device. The volume of the reactor in this setup is 250 cm^3.

Below are the experimental data carried out on a laboratory device. The experiments were carried out at atmospheric pressure in the stationary region of catalyst activity. The following technological parameters were taken as controlled parameters, the values of which vary within specified intervals $85 \leq T \leq 125^0C$, $900 \leq V \leq 3200\ h^{-1}$, $0{,}13 \leq P_{0_2} \leq 0{,}6$ atm, $0{,}09 \leq P_{CH_3OH} \leq 0{,}43$atm (where T – temperature, V - space velocity, P_{0_2}- partial pressure of oxygen, P_{CH_3OH} - partial pressure of methanol).

The results of the experiments are shown in table 1.

Table 1: Experimental data of the process for obtaining formic acid by gas-phase oxidation of methanol

№	T, ^{0}C	P_{CH_3OH}, atm	P_{0_2}, atm	V,h^{-1}	Y, Yield of formic acid, %
1	70	0.07	0.1	850	23.5
2	70	0.07	0.1	3250	20
3	70	0.07	0.7	850	23.8
4	70	0.07	0.7	3250	19.3
5	70	0.47	0.1	850	24.5
6	70	0.47	0.1	3250	20
7	7	0.47	0.7	850	28
8	70	0.47	0.7	3250	23
9	130	0.07	0.1	850	29.5
10	130	0.07	0.1	3250	23
11	130	0.07	0.7	850	31.5
12	130	0.07	0.7	3250	26.8
13	130	0.47	0.1	850	32.6
14	130	0.47	0.1	3250	26.3
15	130	0.47	0.7	850	36
16	130	0.47	0.7	3250	30.5

3. Results and discussion

At present, in practice the task of improving the performance of chemical reactors is not solved by increasing their size, and by developing new methods of process, implementation of optimal control. Process optimization is one of the most important stages of mathematical modeling. Development of methods for optimizing catalytic processes devoted considerable number of papers (Aliyev A.M. et al., 2010b).

The problem of optimization of the catalytic process is solved in two stages. The first stage of optimization of the catalytic process was based on a kinetic model. It is this step allows you to find the limiting performance of the process in view of its physical and chemical laws. The second stage - process optimization is the selection of the optimal design and operational parameters of the reactor: geometric shapes, sizes, components, flow rate of the reaction components, temperature, pressure, concentration, etc (Lacks D. J., 2003).

Here carried out some calculations on the calculation of the geometric dimensions of the reactor and the Reynolds number for this reaction. We used experimental data obtained on an automated chemical process plant related with the computer (Nauman E. B., 2002).

The optimal organization of research with the introduction of automation and electronic computing machines facilitates and accelerates the development of new chemical technological processes. One of these directions is related to automated processing of measurement data on the basis of personal computers (PC) and direct digital control of the experimental facility with PC teams in real time. For the other direction, the application of the method of active experience planning using optimal composition plans is typical. On the basis of these methods, a number of software and software complexes for computing research have been created and found their application.

Programs on optimal methods of planning for the operation of an automated facility have been developed. They provide the following basic operations: formation of the initial matrix of the planning of experiments, its expert evaluation and sending information about the modes of operation to the display screen or directly to the control system; acceptance of quantitative and qualitative indicators corresponding to established practices about the studied process and their processing; building the equation of the regression model of the process and checking its adequacy with experimental data; choosing the optimization criteria and finding the optimal region of the process flow.

The need for the numerical solution of problems of chemical kinetics for reactions with a large number of steps is caused by modern requirements of the industry. Solution of tasks on improvement of oil and gas processing, improvement of chemical reactors are actual.

For each of these tasks should be carried out mathematical modeling and before changing production processes, it is necessary to improve chemical schemes at various stages of production. However, analytically solve such problems practically do not work out because of the huge size of the systems of ordinary nonlinear differential equations corresponding to schemes of chemical reactions. The software package OptimMe was used for calculations (Manafov M.R. et al., 2015). The optimal parameters of the process were determined by the method of statistical planning of experiment. Catalyst activity rating was carried out on percentage of acetic acid in products of model reaction. It is obtained regression model of the process and the optimal regime of flow of this process is defined.

As a rule, the same method was used for most tasks without properly accuracy evaluation of the obtained results and the possibility of achieving the same goals using different methods of experimental design. In addition, possibilities of the computer for automation of procedure of a choice of the experimentations methods accepted for a concrete case, formations of an initial matrix of planning, processing of results and other possible operations with orientation to work of the researcher who doesn't have special preparation and knowledge of mathematics and programming aren't fully used. The best solution of the specified problem is creation of dialogue systems with the package of applied programs (PAP), during the work with which the researcher is exempted from studying of programming languages (Manafov M. R., 2015).

Such system not only offers the user a set of software modules for the solution a wide range of tasks, but also helps him make the right choice of a method for solution of tasks, an assessment of the accuracy of the results obtained, to control and prevent errors, etc.

Presented software package was used to develop and optimize and the search process of active catalysts. For this process have been studied metal-zeolite catalyst properties of the catalysts prepared by ion exchange of the natural zeolite – klinoptilite.
Using the well-known technique, we compiled regression equations for this process.

The multiple regression equation is presented as: $Y = f(\beta, X) + \varepsilon$, where $X = X(X_1, X_2, \dots, X_m)$ - is the vector of independent variables; β - vector of parameters (to be determined); ε - deviation; Y - is the dependent variable.

The theoretical linear multiple regression equation is: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m + \varepsilon$, where β_0 - is a free term that determines the value of Y, m the case when all explanatory variables X_j are equal to 0.

The empirical equation of multiple regression of the process of catalytic gas-phase oxidation of methanol can be represented as: $Y = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_m X_m + e$, where b_0, b_1, ..., b_m - estimates of theoretical values β_0,

$\beta_1, \beta_2, \ldots, \beta_m$ of regression coefficients (empirical coefficients of the regression equation); e - deviation estimate ε.

The least squares method (LSM) was used to estimate the parameters of the multiple regression equation. According to this method, the vector S is obtained from the expression: $S = (X^T X)^{-1} X^T Y$.

The resulting regression equation has the form

$$Y=15.5893+0.1127X_1+7.3438X_2+4.0625X_3-0.00211X_4$$

Then we carried out a statistical analysis of the obtained regression equation: checking the significance of the equation and its coefficients, studying the absolute and relative approximation errors.

From the conducted studies, it can be concluded that the mathematical model of the function is similar to the object. On the basis of the function Y, the optimization operation is carried out and the technological conditions are determined, which ensure the minimal execution of the function Y.

For the unbiased estimate of the variance, calculations were made, the results of which are shown in Table 2.

Table 2: Calculation results of the unbiased estimate of the variance

Y	$Y(x)$	$\varepsilon=Y-Y(x)$	ε^2	$(Y-Y_{av})^2$	$\|\varepsilon : Y\|$
23.5	22.606	0.894	0.799	6.989	0.038
20	17.544	2.456	6.033	37.746	0.123
23.8	25.044	-1.244	1.547	5.493	0.0523
19.3	19.981	-0.681	0.464	46.837	0.0353
24.5	25.544	-1.044	1.089	2.702	0.0426
20	20.481	-0.481	0.232	37.746	0.0241
28	27.981	0.0187	0.000352	3.446	0.00067
23	22.919	0.0812	0.0066	9.883	0.00353
29.5	29.369	0.131	0.0172	11.264	0.00445
23	24.306	-1.306	1.706	9.883	0.0568
31.5	31.806	-0.306	0.0938	28.689	0.00972
26.8	26.744	0.0562	0.00316	0.431	0.0021
32.6	32.306	0.294	0.0863	41.683	0.00901
26.3	27.244	-0.944	0.891	0.0244	0.00359
36	34.744	1.256	1.578	97.146	0.0349
30.5	29.681	0.819	0.67	18.977	0.0268
			15.217	358.939	0.499

Here $\varepsilon = Y - Y(x) = Y - X*S$ - is the unbiased error or the absolute approximation error.

The value of the average approximation error is

$$A = \frac{\sum|\varepsilon Y|}{n} \times 100\% = \frac{0.499}{16} \times 100\% = 3.12\%.$$

The variance estimate is equal to:

$$S_e^2 = (Y - Y(X))^T (Y - Y(X)) = 15.217$$

The inconsistent variance estimate is equal to:

$$s^2 = \frac{1}{n-m-1} \times s_e^2 = \frac{1}{16-4-1} \times 15.217 = 1.3834.$$

The multiple correlation coefficient is:

$$R = \sqrt{1 - \frac{0.0424}{1}} = 0.9786$$

This means that the relationship between trait Y (output of formic acid) and factors X_i is strong.

Also checked the significance of the parameters of the multiple regression equation. The statistical significance of the regression coefficients b_0 , b_1 , b_2 , b_3, b_4 is confirmed.

Next, the paired correlation coefficients were calculated using the formula: The statistical significance of the equation was checked using the coefficient of determination and Fisher's test. It was also found that the model parameters are statistically significant.

$$r_{xy} = \frac{\bar{xy} - \bar{x}\bar{y}}{s(x)s(y)}$$

$$r_{yx_1} = \frac{2715.813 - 100 \times 26.144}{30 \times 4.736} = 0.714, r_{yx_2} = 0.31, r_{yx_3} = 0.257, r_{yx_4} = -0.534.$$

$$r_{x_1x_2} = 0,\ r_{x_1x_3} = 0, r_{x_1x_4} = 0, r_{x_2x_3} = 0, r_{x_2x_4} = 0, r_{x_3x_4} = 0.$$

,

Fisher's F-test was used to assess the significance of the regression equation.

From the Fisher-Snedokkor distribution table, the critical value of the F-criterion (F_{cr}) was found at a significance level of = 0.05 and two numbers of degrees of freedom $k_1=m$ and $k_2=n-m-1$. The Fisher criterion was calculated by the formula

$$R^2 = 1 - \frac{s_e^2}{\sum(y_i - \bar{y})^2} = 1 - \frac{15.217}{358.94} = 0.9576,$$

$$F = \frac{R^2}{1 - R^2} \times \frac{n - m - 1}{m} = \frac{0.9576}{1 - 0.9576} \times \frac{16 - 4 - 1}{4} = 62.118.$$

Table value at degrees of freedom $k_1 = 4$ and $k_2 = n\text{-}m\text{-}1 = 16 - 4 - 1 = 11$, $F_{cr}(4;11) = 3.36$.

Since the actual value of $F > F_{cr}$, the coefficient of determination is statistically significant and the regression equation is statistically reliable (i.e., the coefficients bi are jointly significant).

4. Conclusions

As a result of the calculations, a multiple regression equation was obtained for the process of catalytic gas-phase oxidation of methanol to formic acid. Statistical analysis of the resulting regression equation was carried out: checking the significance of the equation and its coefficients, studying absolute and relative approximation errors. Based on the maximum coefficient r_{yx_1}=0.714, we conclude that changes in the temperature of the reaction medium have the greatest impact on the yield of formic acid. The statistical significance of the equation was verified using the coefficient of determination and Fisher's test. It was also found that the model parameters are statistically significant.

References

1. Aliyev A.M., Medzhidova S.M., Shakhtakhtinsky T.N., Huseynov K.A. 2010a. Vapor-phase oxidation of methyl alcohol to formic acid on a modified zeolite catalyst // Journal "Chemistry and Chemical Technology", Ivanov. Volume 53, issue 6, p. 95-101.

2. Aliyev A.M., Mammadov E.M., Aliyev Q.S., Majidova S.M., Huseynov K.A. 2010b. // XIX International Conference on Chemical Reactors Chemreactor-19, Vienna, Austria,. PP-II-1, p. 336-337.

3. Javadova S.P., Mejidzade V.A., Aliev G.S., Aliev A.Sh., Tagiev D.B. 2021. Mathematical modeling of the process of electrochemical deposition of the bismuth-selenium system // Chemical Problems, 1 (19), p. 47-55

4. Lacks D. J. 2003. Real-Time Optimization in Nonlinear Chemical Processes: Need For Global Optimizer / Daniel J. Lacks // AIChE J.Vol. 49, № 11.

5. Majidzade V.A., Aliyev G.S., Aliyev A.Sh., Huseynova R.H., Mammadova Z.M.. 2021. Mathematical modeling and optimization of the electrodeposition process of antimony-selenium system. Azerbaijan Chemical Journal, No. 1, pp.30-36.

6. Majidzade V.A.; Aliyev G.S.; Javadova S.P.; Aliev A.Sh., Dadashova S.D., Tagiev D.B. 2022. “Mathematical modeling and optimization of the formation of functional thin films.” Journal of Mathematical Modeling, Vol 34, No. 6, pp.111-119.

7. Manafov M. R., Mammadov E. M., Aliyev G. S. 2015. Application of Software Package “OptimMe” for the Study of the Process of Partial Oxidation of Propane. American Journal of Chemistry and Application. Vol. 2, No. 4, pp. 47-51.

8. Manafov M. R. 2015. Development of a Software Application for Solving of Problems of Chemical Kinetics and its Implementation in a C #. International Journal of Engineering and Applied Sciences. Vol. 2, No. 10, pp. 33-37.

9. Nauman E. B. Chemical reactor design, optimization, and scaleup. 2002. — N.Y.: Rensselaer Polytechnic Institute Troy, 618 p.