RECENT J. (2020), 62:130-135
https://doi.org/10.31926/RECENT.2020.62.130

Study of an Industrial Process Using the Correlation Analysis

Flavius Aurelian SARBU
Transilvania University of Brasov, Romania, sflavius@unitbv.ro

Abstract

The work presents a method of analysis specific to statistical processing of a technological industrial process. On the basis of the experimental data produced and recorded from the analyzed technological process, the link between two specific variables of the processing process by volumetric electrochemical erosion has been calculated and verified. For the purpose of determining the correlation between process variables, the values obtained in the experimental determination of the basic electrochemical characteristics obtained for a 55VMoCrNi17 matrix steel. This study was conducted based on the determination of the processing parameters: current density j [A/mm^2], feed speed v_A [mm/min], or front gap s_{90} [mm]. When estimating the links between the different factors and the status variable (s), the correlation analysis shall be used.

Keywords

industrial process, correlation, technological variables, Pearson correlation coefficient, volumetric electrochemical erosion

1. Introduction

On the basis of the experimental data produced and recorded from the analyzed technological process, the link between two specific variables of the processing process by volumetric electrochemical erosion has been calculated and verified [1].

The processing technologies by volumetric electrochemical erosion are based on the laws of the anode dissolution of the material in which the processing takes place.

For the development of technologies, several steps are taken in design, starting with the material of the processed part, the type, precision, complexity and roughness of the surfaces to be obtained, but also the electrochemical erosion pick-up facility available.

Irrespective of the degree of complexity regarding the processed surfaces, knowledge of the basic electrochemical characteristics is absolutely necessary to develop the custom technology.

2. Determination of Basic Electrochemical Characteristics

Basic electrochemical characteristics are data specific to the torque of the processed material - electrolyte solution used during processing. The importance of these characteristics for the electrochemical processing is the same as the one of the of the breaking tension σ_r and flow rate σ_c for chipping processes.

The basic electrochemical characteristics of a material, to be processed by electrochemical erosion, are:

V_{ef} - specific actual volume of dissolved material, in [mm^3/A·min];
B - the coefficient of dissolution rate, in [mm/min];
U_{pol} - voltage drop on electrodes, in [V];
U_p - the voltage drop in the passive layer from the anode, in [V].

These characteristics are obtained separately for non-passive electrolytes (NaCl) and passive electrolyte (NaNO3) using a special device (supported by CIRP) [1]. The basic electrochemical characteristics are essential in modelling the processing process by volumetric electrochemical erosion, based on the determination of the processing parameters: current density j [A/mm^2], feed speed v_A [mm/min], or front gap s_{90} [mm].

3. The Correlation Analysis Specific to the Preliminary Experiment

If the dependence between the parameters of a process is studied, the following problems arise:

- determine whether the parameters initiating the process are independent or influence each other (depending on each other), and therefore determine the correlation between the parameters;
- establishing the type of connection between the parameters.

The correlation is defined as the interdependence or the link between the variables observed in statistical populations, so the correlation is the technique of analyzing the links between statistical variables. When estimating the links between the different factors and the status variable (s), the correlation analysis shall be used.

With the help of statistical criteria, factors and interactions that have significant influences are selected; these will be taken into account in the basic experiment and will be included as variables in the mathematical model of the process.

The correlation of two sizes X and Y means that both sizes are determined by the same phenomena and legities, often unknown, and not a direct dependence of sizes on each other.

At that time, the relationship between two sets of observations considered simultaneously is highlighted. These series are usually obtained by measuring two quantitative (variable) characteristics for the same studied sample.

By means of correlation analysis, it is possible to identify all the elements necessary to confirm the existence of a link between two phenomena, based on the mode of expression observation. For this purpose, measurements of variables that characterize the phenomena being tested are used.

The two-dimensional distribution obtained by processing the pairs of values determined by measuring the two variables that characterize the cause phenomenon and the effect phenomenon is required to use this method.

The size called the covariance is used to establish the correlation between the parameters of a process.

The two-dimensional distribution obtained by processing the pairs of values determined by measuring the two variables that characterize the cause phenomenon and the effect phenomenon is required to use this method.

If the parameters X and Y are considered which have certain experimental values: X (x_1, x_2, ..., x_n) and Y (y_1, y_2, ... , y_n), then the size of the covariance is [2, 3]:

$$S_{xy} = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - \bar{x}) \cdot (y_i - \bar{y}) \tag{1}$$

where: n = volume of sample (number of data pairs x_i, y_i)

x_i, y_i = data pairs;

$\bar{x}$, $\bar{y}$ = the arithmetic means of the experimental data sets.

The Covariance (pair variability) is independent of the sample size.

In a narrow sense, the correlation is a measure of the degree of statistical link between the quantitative variables, referred to as the ‘correlation coefficient’.

The most commonly used is Pearson's coefficient of correlation “r” (linear correlation coefficient) which measures the degree of connection between the variables [2-4]:

$$r_{xy} = \frac{\frac{1}{n} \sum (x - \bar{x})(y - \bar{y})}{S_x S_y} = \frac{S_{xy}}{S_x S_y}, \tag{2}$$

where: S_x, S_y = Standard deviation (Dev. std.) of all values of x, respectively y; (S_x S_y = total covariance).

The standard deviations of the y variables are calculated with the relationships:

$$S_x = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}}, \tag{3}$$

$$S_y = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n}} . \tag{4}$$

By replacing relationships (1), (3) and (4) in relation (2) is obtained:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x - \bar{x})(y - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \cdot \sum_{i=1}^{n}(y_i - \bar{y})^2}} . \tag{5}$$

Correlation coefficient ”*r*” is a quantitative value describing the relationship between two or more variables. It varies between (–1 and 1), where the extreme values assume a perfect relationship between the variables, whereas 0 means a total lack of linear relationship (non-linear dependence).

The higher the absolute value of *r*, the stronger the correlation link: values where $r = \pm 1$ corresponding to a rigorous functional dependence between process variables [2-5].

4. Experimental Research on the Determination of Correlation between Processing Process Variables by Volumetric Electrochemical Erosion

For the purpose of determining the correlation between process variables, the values obtained in the experimental determination of the basic electrochemical characteristics obtained for a 55VMoCrNi17 matrix steel [1] and which are given in Table 1 have been taken into account.

In the analyzed case, non-passive electrolytes with a concentration of 12 % NaCl [1] were used. For each feed speed considered from the range of the $v_A \in (0.25 \ldots 2)$ [mm/min] anode current has been measured and taking into account the known front surface of the cylindrical specimen used, the current density (*j*) and the front gap s_{90} [mm] [1] can be determined by experimental calculations.

Table 1. The values of the experimental data

No. crt.	v_A [mm/min]	j [A/mm^2]	s_{90} [mm] resulted from the calculation
1	0.25	0.12	2
2	0.45	0.18	1.78
3	0.85	0.32	0.92
4	1.12	0.42	0.74
5	1.22	0.50	0.62
6	1.35	0.58	0.51
7	1.48	0.60	0.42
8	1.64	0.70	0.38
9	1.83	0.75	0.35
10	1.92	0.80	0.22
Average	1.211	0.497	0.794
Dev. std.	0.558	0.234	0.614

Table 2 shows the method of calculation to determine the value of the coefficient “*r*” Pearson according to the relationships (1) ... (5), specific to the correlation between the values of the forward speed (v_A) and current densities (*j*); it is to be identified $v_A = x$ and $j = y$.

With the data obtained in Table 2, the specific values of S_{xy}, S_x and S_y are calculated according to the relationships (1), (3) and (4), giving the values: $S_{xy} = 0.117$; $S_x = 0.529$; $S_y = 0.222$.

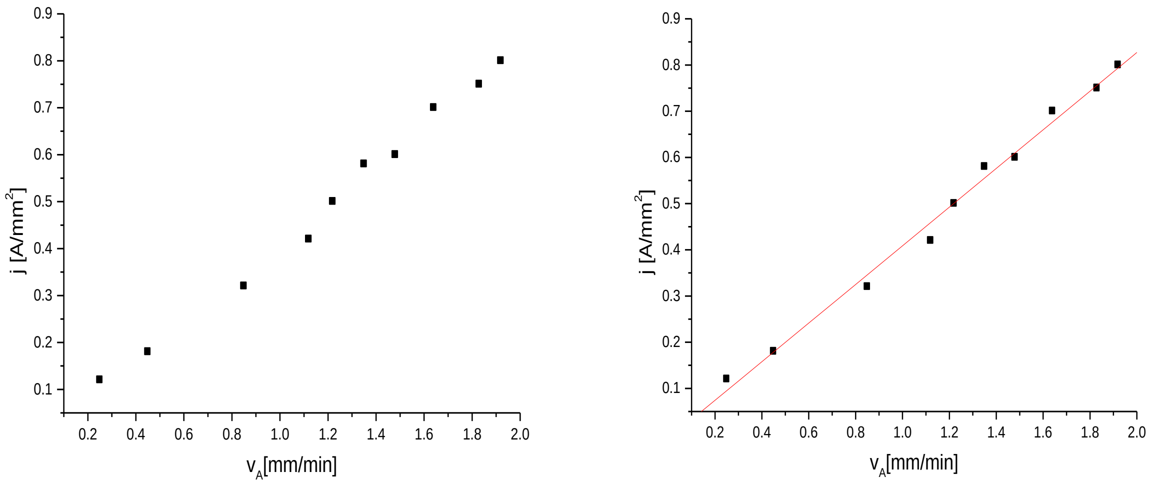
The final calculated value of the coefficient “*r*” Pearson is: $r_{xy} = 0.996$

r_{xy} has very high values (close to 1) which denotes a very strong correlation link between feed speed (v_A) and current density (*j*), which indicates that the variables are related.

Table 2. Method of calculation to determine the value of the coefficient "*r*" Pearson specific to the dependence between v_A and j

No. crt.	$x = v_A$ [mm/min]	$y = j$ [A/mm^2]	$(x-\bar{x})$	$(y-\bar{y})$	$(x-\bar{x})(y-\bar{y})$
1	0.25	0.12	-0.961	-0.377	0.362
2	0.45	0.18	-0.761	-0.317	0.241
3	0.85	0.32	-0.361	-0.177	0.064
4	1.12	0.42	-0.091	-0.077	0,007
5	1.22	0.50	0.009	0.003	0.000
6	1.35	0.58	0.139	0.083	0.012
7	1.48	0.60	0.269	0.103	0.028
8	1.64	0.70	0.429	0.203	0.087
9	1.83	0.75	0.619	0.253	0.157
10	1.92	0.80	0.709	0.303	0.215
Average	1.211	0.497	$\sum (x-\bar{x})(y-\bar{y})$		1.172
Dev std.	0.558	0.234	$\frac{1}{n}\cdot\sum (x-\bar{x})(y-\bar{y})$		0.117

The graphical representation of the correlation between the values of feed speed (v_A) and current densities (j) are shown in Figure 1.



a) b)

Fig. 1. Perfectly positive correlation between the values of the feed speed (v_A) and the current density (j): a) representation of the scatter; b) linear representation

Looking at Figure 1, a and b, it is noted that the shape of the perfectly positive (direct) correlation, a linear relationship between the values of the feed speed (v_A) and the current density (j) is confirmed.

Table 3 shows how to calculate the value of the coefficient "*r*" Pearson specific to the correlation between according to the relationships (1) ... (5), specific to the correlation between the values of the feed speed (v_A) and the front air (s_{90}); it is to be identified $v_A = x$ and $s_{90} = y$.

With the data obtained in Table 3, the specific values of S_{xy}, S_x and S_y shall be calculated according to the relationships (1), (3) and (4), giving the values: S_{xy} = -0.295; S_x = 0.529; S_y = 0.583.

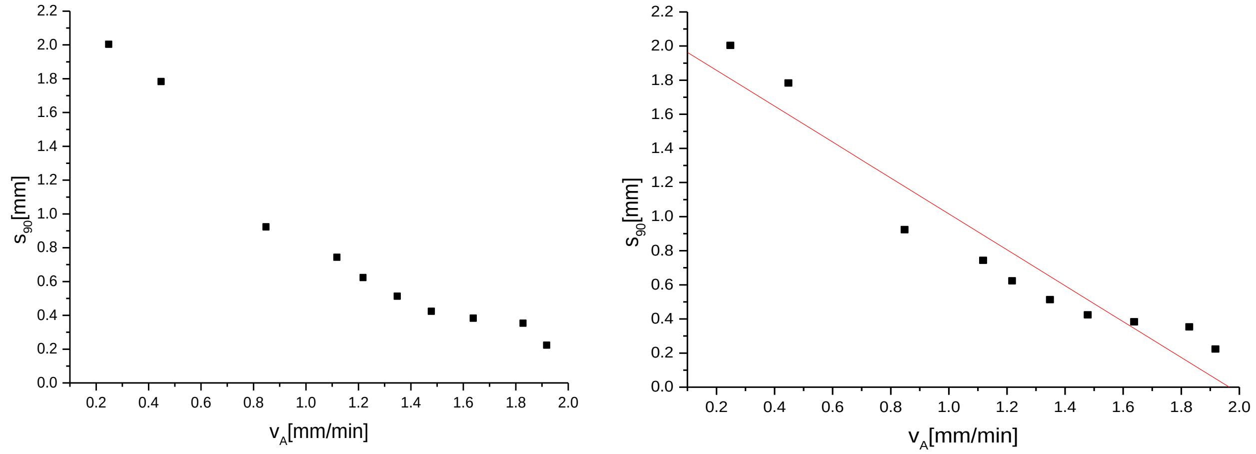
The final calculated value of the coefficient "*r*" Pearson is: $r_{xy} = -0.957$

r_{xy} has high values which denote a strong correlation link between feed speed (v_A) and current density (j), which indicates that the variables are related.

The graphical representation of the correlation between the values of feed speed (v_A) and the front gap (s_{90}); they are shown in Figure 2.

Table 3. Method of calculation to determine the value of the coefficient "*r*" Pearson specific to dependence between v_A and s_{90}

No. crt.	$x = v_A$ [mm/min]	s_{90} [mm] determined from calculation	$(x-\bar{x})$	$(y-\bar{y})$	$(x-\bar{x})(y-\bar{y})$
1	0.25	2	-0.961	1.206	-1.444
2	0.45	1.78	-0.761	0.986	-0.976
3	0.85	0.92	-0.361	0.123	-0.153
4	1.12	0.74	-0.091	-0.054	-0.022
5	1.22	0.62	0.009	-0.174	0.001
6	1.35	0.51	0.139	-0.284	0.002
7	1.48	0.42	0.269	-0.374	-0.021
8	1.64	0.38	0.429	-0.414	-0.050
9	1.83	0.35	0.619	-0.444	-0.091
10	1.92	0.22	0.709	-0.574	-0.196
Average	1.211	0.794	$\sum (x-\bar{x})(y-\bar{y})$		-2.950
Dev. std.	0.558	0.614	$1/n\sum (x-\bar{x})(y-\bar{y})$		-0.295



a) b)

Fig. 2. Perfectly positive correlation between the values of the feed speed (v_A) and the front air (s_{90}) a) representation of the scatter; b) linear representation

Looking at Figure 2, a and b, it is noted that the shape of the perfectly negative (inverse) correlation, linear relationship, between the values of the forward speed (v_A) and those of the front air (s_{90}) is confirmed.

5. Conclusions

The following conclusions can be drawn from the analysis of the results obtained:

- The processing technologies by volumetric electrochemical erosion are based on the laws of the anode dissolution of the material in which the processing takes place.
- When estimating the links between the different factors and the status variable (*s*), the correlation analysis shall be used.
- Note that the forward speed (v_A) and current density (*j*) values have a perfectly positive (direct) correlation, the relationship being linear.
- Note that the values of feed speed (v_A) and front air (s_{90}) are perfectly negative (inverse), the relationship being linear; note the reduction in the size of the front air s_{90} as the processing speed v_A increases.

References

1. Marinescu N.I. – coordonator (2006): *Tratat de tehnologii neconvenţionale. Vol. IV – Prelucrarea prin eroziune electrochimică* (*Treaty of unconventional technologies. Vol. IV - Electrochemical erosion processing*). Printech Bucureşti, ISBN 978-973-718-613-3, Bucureşti, Romania (in Romanian)
2. Leoveanu I.S. (2006): *Optimizarea proceselor tehnologice. Vol 1. Aplicații generale (Optimization of technological processes. Vol 1. General applications)*. Lux Libris, ISBN 973-9458-54-8, p. 61-62, Brasov, Romania (in Romanian)
3. Hauke J., Kossowski T. (2011): *Comparison of Values of Pearson's and Spearman's Correlation Coefficients On the Same Sets of Data*. Quaestiones Geographicae, ISSN 2081-6383, Vol. 30, no 2, pp. 87-93, DOI: 10.2478/v10117-011-0021-1, https://www.researchgate.net/publication/227640806_Comparison_of_Values_of_Pearson%27s_and_Spearman%27s_Correlation_Coefficients_on_the_Same_Sets_of_Data
4. Taloi D. (1987): *Optimizarea proceselor tehnologice (Optimization of technological processes).* Editura Academiei Române, p. 101-104, Bucharest, Romania (in Romanian)
5. Haining R. (1991): *Bivariate Correlation with Spatial Data*. Geographical Analysis, Vol. 23, No. 3, p. 210-227, https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1538-4632.1991.tb00235.x