Proceedings in MANUFACTURING SYSTEMS
Proceedings in Manufacturing Systems, Vol. 5 (2010), No. 4 ISSN 2067-9238

REGRESSION MODELS APPLIED TO MANUFACTURING SYSTEMS

Adrian Stere PARIS[1], Constantin TÂRCOLEA[2]

Abstract: *The problems of manufacturing systems are explored by means of regressions functions. In the first part of the paper are presented some statistical models for manufacturing costs, quality loss function (shifted gamma) and reliability (vapor pressure failure function). An application concerning the manufacturing accuracy of machine-tools is developed under the term of technologic reliability.*

Key words: *regression, technologic reliability models, manufacturing application*.

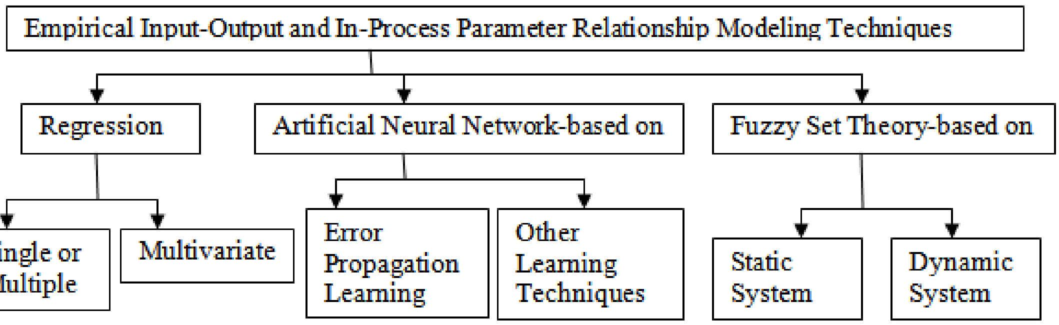

Fig. 1. Manufacturing modeling techniques [9].

¶

1. INTRODUCTION AND LITERATURE REVIEW

Statistical models are widely used in research, industrial manufacturing, reliability, security, safety, and this list is by no means exhaustive. A very important aspect of applied statistics concerns the statistical analysis and modeling of manufacturing systems [19]. Regression analysis is a statistical tool for the investigation of relationships between variables [20]. Usually, the investigator seeks to ascertain the causal effect of one variable upon another - the effect of a price increase upon demand, for example. To explore such issues, the investigator assembles data on the underlying variables of interest and employs regression to estimate the quantitative effect of the causal variables upon the variable that they influence. The investigator typically assesses the “statistical significance” of the estimated relationships, that is, the degree of confidence that the true relationship is close to the estimated relationship. Regression techniques have long been central to the field of economic statistics (“econometrics”). Increasingly, they have become important to lawyers and legal policy makers as well [18].

Regression is a conceptually simple technique for investigating functional relationship between output and input decision variables of a manufacturing process and may be useful for process data description, parameter estimation, and control [9]. Although statistical regression may work well for modelling, this technique may not describe precisely the underlying non-linear complex relationship between the decision variables and responses. A prior assumption regarding functional relationship(s) [such as linear, quadratic, higher-order-polynomial, and exponential] between output(s), and input decision variable(s), is a pre-requisite for regression equation-based modelling.

Prediction of output(s) for an unknown set of input(s) based on regression technique is valid only over the region of the regressor variable(s) contained in the observed cutting process data. It is only an aid to confirm cause-effect relationship, and does not imply a cause and effect relationship. Moreover, error components of regression equation need to be mutually independent, normally distributed, and having constant variance [8].

A large number of techniques has been developed by researchers to solve these types of parameter optimization problems, and may be classified as conventional and nonconventional optimization techniques. Fig. 1 [9] provides a general classification of different input-output and in-process parameter relationship modelling and optimization techniques in metal cutting processes, respectively. Whereas conventional techniques attempt to provide a local optimal solution, non-conventional techniques based on extrinsic model or objective function

[1] PhD, Math., Professor, University "Politehnica" of Bucharest, Department of Mathematics, Splaiul Independentei 313, Sector 6, 060042, Bucharest, ,Romania,
E-mail: constantin_tarcolea@yahoo.com
[2] PhD, Eng, Lecturer, University "Politehnica" of Bucharest, Materials Tehnology and Welding Department, Splaiul Independentei 313, Sector 6, 060042, Bucharest, Romania,
E-mail: adrianparis@amza.camis.pub.ro

developed, is only an approximation, and attempt to provide near-optimal cutting condition(s). Conventional techniques may be broadly classified into two categories: in the first category, experimental methodology (RSM) are referred to and in the second category, iterative mathematical search techniques, such as linear programming (LP), non-linear programming (NLP), and dynamic programming (DP) algorithms are included. Non-conventional meta-heuristic search-based techniques, which are sufficiently general and extensively used by researchers in recent times are based on genetic algorithm (GA), tabu search (TS), and simulated annealing (SA) [9].

2. MANUFACTURING APPLICATIONS OF REGRESSION ANALYSIS

2.1. Manufacturing costs modelling

An important application of regression analysis is concerned with the economic aspects of the manufacturing processes. In order to support the product and process development needed to meet market expectations such as greater choice of products, greater choice of manufacturing process and greater emphasis on minimizing overall life cycle costs of products, it is expected that the quantity, type, accuracy and complexity of cost information will need to be greatly increased. These changes will have a dramatic effect on the cost estimating process, the process of calculating the expected cost resources, i.e. labor, material and overhead costs that are required to accomplish a manufacturing task or to manufacture or purchase a specific product [17]. For example, Huang [6] develops cost models based on Support Vector Regression for complex products during the early design phases.

2.2. Quality loss functions

In 1986, Taguchi presented the quadratic quality loss function for reducing deviation from the target value. The objective of this quality improvement method is to minimize total losses to society. The canonical distributions have usually unbounded range, while the quality loss function can be non-zero only on bounded interval [2]. The distribution analysis of experimental data points indicated that curve profiles can be described more adequately different of the degrees of skewness and peakedness of the curves. For a sample of product, the average loss can be decomposed as a sum of the variance and bias. According to this approach, a manufacturing process must have two complementary goals: zero bias and the smallest possible variance.

Taguchi philosophy of quality control focuses on the design stage. The deviations from control target should be evaluated in terms of the loss of quality they cause. A loss occurs even if the outcome is still within pre-specified tolerance bounds. In [14], some types of models using polynomial, exponential and generalized gamma pattern were proposed. Figure 2 illustrates an example of asymmetrical gamma law.

The analysis of experimental data were performed on CurveExpert [5] software developed for curve fitting (nonlinear regression – least squares method, Levenberg-Marquardt algorithm).

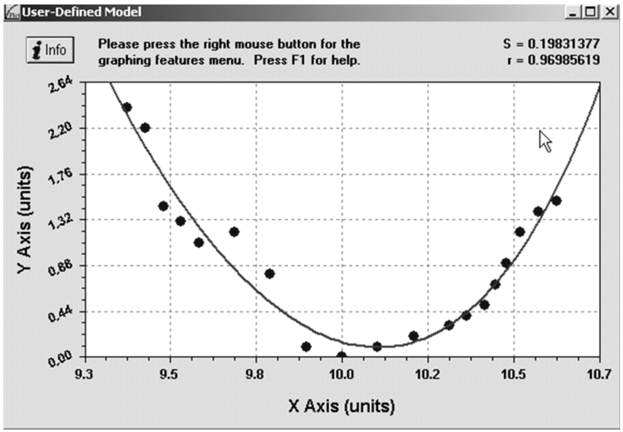

Fig. 2. Plot shifted generalized gamma model.

3. REGRESSION MODELS IN RELIABILITY OF MANUFACTURING SYSTEMS

3.1. Reliability models

The reliability community has become well experienced in fitting of survival distributions, the use of design of experiments (DOE) and the associated general linear model (linear regression and analysis of variance methods) approach to analysis [3].

A study focuses on accelerated degradation tests (ADT) data analysis [16]. Most of the ADT analyses use parametric regression models to estimate the lifetime of the product under usual use. To relax the assumptions on the form of regression functions and let data speak for themselves in searching for a suitable model for data, it propose a nonparametric regression model to analyze ADT data [16]. Nonparametric regression techniques are useful in obtaining a smooth fit to noisy data, to describe the relationship between response variables and independent variables. These smoothing techniques are powerful tools in statistical data-analysis because of the model flexibility, appealing look of the fitted curves (or surfaces) [16].

The reliability function is a corner-stone of reliability theory. In [12 and 13] a wide range of possible continuous failure functions are presented, such a Hoerl model, sinus model and vapour pressure model, the last one being plotted in Fig. 3: the experimental data were processed by software Curve Expert [5].

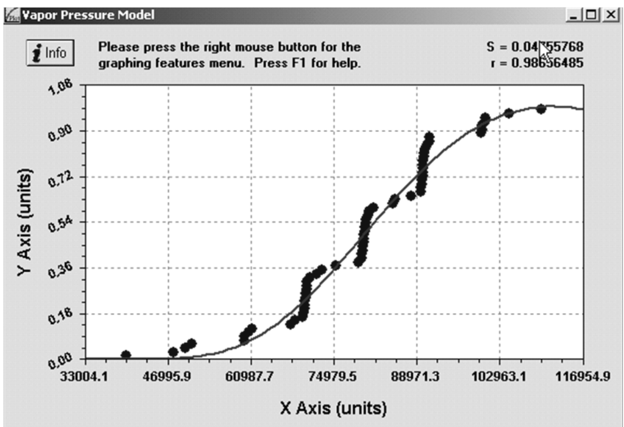

Fig. 3. Vapor pressure failure function:

$$\beta_1 y(x) = e^{141.4 - \frac{1.25*10^6}{x} - 11.2\ln x}$$

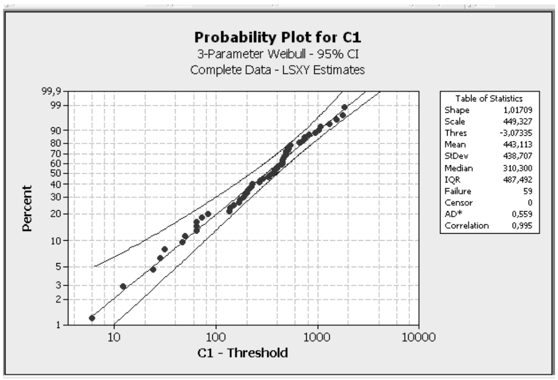

Fig. 4. Weibull failure function.

Anyway, for reliability studies a frequent solution remains Weibull function (Fig. 4) [11]; here, the failure function was plotted with software Minitab 15.

3.2. Technological reliability models

To emphasize the problem of accuracy for manufacturing equipments it was introduced and developed the term of technological reliability [10 and 15].

The technological reliability at t moment can be quantitatively defined as the probability of a manufacturing equipment (namely a machine-tool) to maintain her working accuracy limits by the time t. This means to check the machine-tool accuracy at different time moments and establish the corresponding function of technological reliability [7, 10, and 15]. It follows a short description of the experimental researches of the authors in the field of technological reliability of a family of lathes, summarized in the Table 1 [7]. The results were processed using some applications of multivariate data analysis, especially correlation theory and regression analysis.

Let two random variables *X* and *Y* at which is the question of determining their level of stochastic linear relation. In general, the indicator used, *r* is called selection correlation coefficient and estimates the correlation coefficient. It is important to study the dependence between deviations of working accuracy, namely deviations from cylindricity of the work piece in a normal lathe, and time. In this respect, measurements were performed on 47 pieces of standard form, processed on the same type of normal lathes, recording the actual working time for each machine, and the time since they were put into service until the considered experiment. Were obtained as 47 pairs of values consisting of a series of values of cylindricity deviations of the work piece, in millimetres, y_i, and correspondently a second string x_i of the values expressed in actual hours of operation until time of measurement [7]. From a statistical viewpoint, now it should be determined whether the two random variables present a linear correlation, which it is translated into the technical developments, while the deviations of accuracy work to be linear.

Next will be calculated the correlation coefficient:

$$r = \frac{\sum_{i=1}^{n} x_i y_i - n\,\bar{x}\,\bar{y}}{\sqrt{\sum_{i=1}^{n} x_i^2 - n(\bar{x})^2}\sqrt{\sum_{i=1}^{n} y_i^2 - n(\bar{y})^2}} = 0.444 . \qquad (1)$$

Check if the value of the correlation coefficient characterizing the linear dependence measure has significant value for a confidence level chosen a priori. Calculation of the empirical value:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = 3.32 \qquad (2)$$

and comparison with the theoretical value of t given in tables. Because:

$$t^* = 3.32 > 2.01 = t_{45}^{0.975},$$

that means there is a significant correlation between two variables.

The regression line equation in this case is:

$$y = 0.000017x + 0.01320. \qquad (3)$$

Next it is calculated the correlation ratio, which also expresses the measure of dependency between the two variables, but on the other hand comparing its value with the correlation coefficient value; it can determine whether the dependence is linear or nonlinear. Since the ratio of correlation $\eta_{y/x}$ has the formula:

$$\eta_{y/x}^2 = \frac{\sum_{j=1}^{n} m_j (\bar{y}/x_j - \bar{y})^2}{(n-1)S_y^2}, \qquad (4)$$

resulting $\eta_{y/x} = 0{,}484$ and thus verify the relation:

$$0 < |r| = 0.444 < 0.484 = \eta_{y/x} . \qquad (\%)$$

The close value of $\eta_{y/x}$ to *r* indicates that the approximate of dependency is a linear function properly.

A new possibility to check the quality of regression is using ANalysis Of VAriance (ANOVA). Also in this case the use of F test is critical. A very accessible way to apply computer aided ANOVA is with MS Excel, from the ad-inn Analysis Tool Pack.

3.2. Logistic regression application

Logistic regression analysis (LRA) extends the techniques of multiple regression analysis to research situations in which the outcome variable is categorical. The main goal is on situations in which the outcome variable is dichotomous, although extension of the techniques of LRA to outcomes with three or more categories [4]. Interesting developments in the reliability field are the asset health reliability estimation based on condition data [22], the prognostic algorithm for machine performance assessment and its application [21], etc. An example is illustrated in Fig. 5 [22], which uses bearing failure simulation data and experiment run-to-failure data. One-dimensional feature, namely kurtosis, is calculated initially. This feature can be used to represent the information of bearing from normal to failure condition. Failure degradation is calculated using the LR method for case of simulated data and case of experiment data. The results are regarded as target vectors of failure probability. RVM is used for training the run-to-failure kurtosis data and

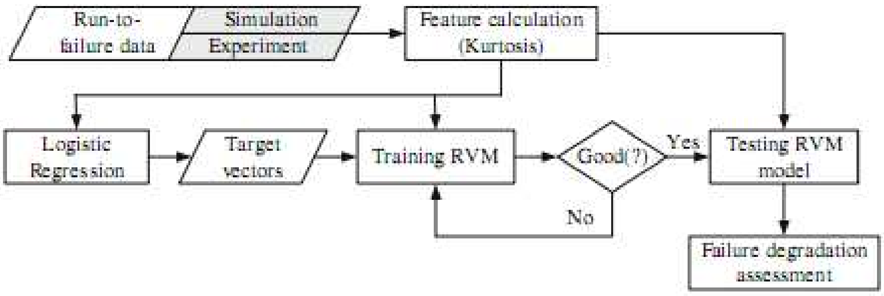

Fig. 5. Schematic diagram of machine degradation assessment model [22].

Table 1

Results of accuracy experiments on a lathes family

Test	r	$\eta_{y/x}$	F^*	$b_{y/x}$	t^*	S_x	S_y
A_3	+0.8993	0.65		0,000001	1.0403	2233.56	0.112
$A_{7(a)}$	–0.058375	0.452386	0.8603	–0.000001	0.2799	1486.79	0.0349
$A_{7(b)}$	0.1414	0.7399	3.73	0.000006	0.66327	1476.7	0.0631
$A_{10(a)}$	0.62124	0.70866	0.817	0.000012	3.387	2265.69	0.04467
$A_{10(b)}$	0.350859	0.539095	0.78718	0.000005	1.823	2284.49	0.0327
$B_{1(a)}$	0.1578	0.40419	0.93778	0.000001	1.0107	1783.83	0.0113
$B_{1(b)}$	0.444	0.484	0.31623	0.000017	3.32	1900	0.073
B_2	0.04455	0.4927	1.8547	0.000001	0.2887	1596.996	0.0207

target vectors of failure probability and then predict the individual unit of machine component. Prediction model is obtained using training process, which produces weight and bias that are saved in the model. To evaluate the training performance, root mean square error (RMSE) and correlation (R) are utilized [22].

4. CONCLUSIONS

As conclusion for technological reliability, although the operating time within the required accuracy depends on the condition of the machine accuracy (i.e. how well the machine-tool was originally set) and the speed of evolution of each parameter accuracy (generally affected by wear of bearings and machine-tool slideways), however, assuming that the machine-tools were properly adjusted, a decisive test of accuracy will be considered the one with the highest speed of development. In this case, approximating regression functions by straight regression, the rate of evolution is proportional to the slope of the straight line.

Therefore, taking into account the correlation values obtained previously (Table 1) it results the following observations:

- Of all the tests of geometric precision for no-load work of the lathe, test A10 (a) (parallelism of the main spindle axis and longitudinal movement of the lathe carriage in the horizontal plane) has the highest correlation coefficient (r = 0.62124). Otherwise (especially for tests A3 and A7 (b)) is a fairly close correlation, but nonlinear. This results in values of t^*, lower then theoretical values, for a confidence level of 0.99.
- For test $A_{10(a)}$, whereas the value of $\eta_{y/x}$ is sufficiently close to r, it will be approximate with a linear correlation. This is confirmed by the level of significance given by the value of $t^* = 3.387 > 2.49 = t_{tab}$ for a confidence level of 0.99, and the value of $F^* = 0.817 < 2.55 = F$.
- Of all tests made on workpiece (in cutting), test B1 (b) has a significant linear correlation as:

$$t^* = 3.32 > 2.5236 = t_{tab}$$

and

$$F^* = 0.93778 < 2.37 = F\,. \quad (11)$$

- Of all the tests considered so, remains in question tests $A_{10(a)}$ and $B_{1(b)}$ (part cylindrical shape in longitudinal section).
- Since at the test $B_{1(b)}$ the highest slope was found ($b_{y/x}$ = 0.00017) of all other tests, this indicates the most pronounced evolving trend and so it will be choosen as a criterion for calculating the technological reliability.

In the general case, if the coefficients $b_{y/x}$ would be comparable, however, would prefer as criterion for calculating the technological reliability the test performed on the probe, as this points out the effect, because of cutting deviations must be sought together with the actual causal evidence geometric accuracy made at no load.

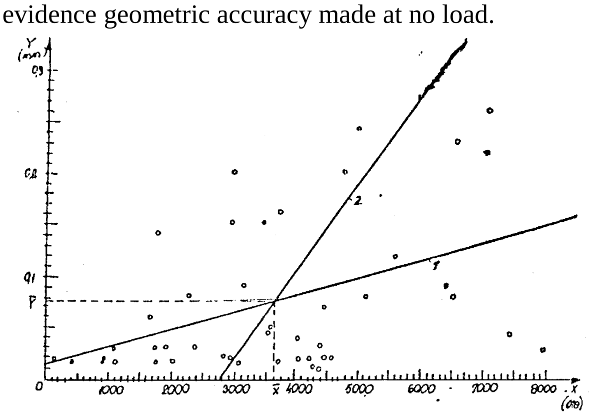

Fig. 6. Regression plot for technological reliability.

Figure 6 presents regressions (*y/x* – line 1 and *x/y* –line 2) in the accuracy test B1 (b). Higher values of r and $\eta_{y/x}$, obtained in test $A_{10(a)}$ versus $B_{1(b)}$ shows a closer correlation (linear or nonlinear) between the obtained deviations and time, but what is of interest is the speed of development, proportional with $b_{y/x}$.

REFERENCES

[1] W. Caesarendra, A. Widodo, B. Yang, *Application of relevance vector machine and logistic regression for machine degradation assessment*, Mechanical Systems and Signal Processing, No. 24 (2010), pp. 1161–1171.

[2] M. Demetrescu, A.S. Paris, C. Târcolea, *Uni-and multivariate loss functions and the Taguchi theory*, BSG Proceedings 17, (DGDS-2009) October 8–11, 2009, Bucharest, Geometry Balkan Press, 2010, pp. 79–83.

[3] T. Haifley, T., *Integrated Reliability Workshop Final Report*, IEEE International, 21–24 Oct. 2002, pp. 184–187.

[4] D.W. Hosmer, S. Lemeshow, *Applied logistic regression*, sec. ed., John Wiley Ed., New Yor, 2004.

[5] `http://www.curveexpert.net/` accessed: 2006-05-23.

[6] G. Huang, *Cost Modeling Based on Support Vector Regression for Complex Products During the Early Design Phases*, PhD Thesis (2007), Virginia Polytechnic Institute and State University, Blacksburg, Virginia, available at: `http://scholar.lib.vt. edu/theses/ available/etd-08292007/`, *accessed*: 2010-03-18.

[7] A. Ionescu, C. Târcolea, A.S. Paris, *Metodă de stabilire a criteriului de calcul al fiabilităţii tehnologice la maşinile-unelte* (Method to prove the calculus criterion of the technological reliability of the machine-tools), Construcţia de maşini, Vol. 23, No. 9, 1976, pp. 485–487.

[8] D.C. Montgomery, A.E., Peck, *Introduction to Linear Regression Analysis*, John Wiley & Sons, London, 1992.

[9] I. Mukherjee, P.K. Ray, *A review of optimization techniques in metal cutting processes*, Computers & Industrial Engineering, No. 50, 2006, pp. 15–34.

[10] A. Oprean, A. Dorin, D. Drimer, A.S. Paris, A. Ionescu, *Fiabilitatea maşinilor-unelte* (Reliability of machine-tools), Edit. Tehnică, Bucharest, (1979).

[11] A.S. Paris, *Contribuţii la studiul fiabilităţii şi disponibilităţii sistemelor tehnologice de fabricaţie* (Contributions to the study of the reliability and the availability of technological manufacturing systems), PhD Thesis, University “Politehnica” of Bucharest, 2009.

[12] A.S. Paris,*Models in mechanical reliability data*, 3rd Symposyum „Durability and reliability of mechanical systems”, University “C. Brancuşi” Tg. Jiu, Fiability and Durability, No. 1(5)/2010, Edit. “Academica Brâncuşi”, Târgu Jiu, pp. 81–86.

[13] Paris, A., S., Târcolea, C. (2010) *New stochastic models applied to manufacturing systems*. Proceedings in Manufacturing Systems, vol5, nr.1, p 23-26. Editura Academiei Române, ISSN 2067-9238

[14] A.S. Paris, C. Târcolea, *Loss functions used in the quality theory*, 3rd Symposyum „Durability and reliability of mechanical systems” University “C. Brancuşi” Tg. Jiu, Fiability and Durability, No. 1(5)/2010, Edit. “Academica Brâncuşi”, Târgu Jiu, pp.

[15] A.S. Pronikov, *Technologhiceskaia nadiojnosti stankov* (The technological reliability of machine-tools), Edit. Maşinostroenie, Moscow, 1971.

[16] J-J.H. Shiau, H-H. Lin, *Analyzing Accelerated Degradation Data By Nonparametric Regression*, IEEE Transactions on reliability, Vol. 48, No. 2, 1999 June, pp. 149–165.

[17] D.J. Stockton, Q. Wang, *Developing cost models by advanced modeling technology*, Proceedings of the I MECH E part B: Journal of engineering manufacture, 218 (2) (2004), pp. 213–224.

[18] A.O. Sykes, *An Introduction to Regression Analysis* available at: *`http://www.gudono.com/metostat/ ebook%20regresi.pdf`*, accessed: 2010-02-13.

[19] C. Târcolea, A.S. Paris, A. Târcolea, *Statistical Models Applied to Manufacturing Systems*, BSG Proceedings 12 the 3-rd international colloquium of mathematics in engineering and numerical physics (MENP-3), October 7–9, 2004, Bucharest, Geometry Balkan Press, 2005, pp. 259–264.

[20] C. Târcolea, A.S. Paris, *Discriminant Analysis and Applied Regression*, The International Conference of Differential Geometry and Dynamical Systems (DGDS-2010), 25–28 August 2010, University “Politehnica” of Bucharest.

[21] J. Yan, M. Koc, J. Lee, *A prognostic algorithm for machine performance assessment and its application*, (2004), Production Planning & Control, Vol. 15, No. 8, December, pp. 796–801.

[22] S. Zhang, M. Lin, S. Yong, M. Joseph, *Asset health reliability estimation based on condition data*, In Proceedings World Congress on Engineering Asset Management, 2007, pp. 2195–2204, Harrogate, England.