Using machine-learning methods to analyze economic loss function of quality management processes

V A Dzedik¹, P A Lontsikh²

¹ Volgograd State University, 100 Universitetskii Ave., Volgograd, 400062
² Irkutsk National Research Technical University, 83 Lermontova St., 664074

E-mail: vdz@mail.ru

Abstract. During analysis of quality management systems, their economic component is often analyzed insufficiently. To overcome this issue, it is necessary to withdraw the concept of economic loss functions from tolerance thinking and address it. Input data about economic losses in processes have a complex form, thus, using standard tools to solve this problem is complicated. Use of machine learning techniques allows one to obtain precise models of the economic loss function based on even the most complex input data. Results of such analysis contain data about the true efficiency of a process and can be used to make investment decisions.

1. Introduction
The main drawback of the vast majority of the modern quality management systems [1,2] is the fact that regardless of their stated purpose of increasing an enterprise’s competitiveness, they either completely disregard or insufficiently consider the economic efficiency of processes [3], above all the product manufacturing processes.

Such quality management systems are based on the so-called "tolerance thinking" [4], i.e. on making decisions about the appropriateness of characteristics of a process or machine when they are located within the tolerance limit regardless of its magnitude. Under such approach, if the value of a considered parameter is within the criteria (for example, the parameter is below the upper tolerance limit of a criterion and above its lower tolerance limit), then it is always considered equally "fit" regardless of its actual value within these limits, and equally "unfit" if its outside these limits.

This paper presents an approach to solving this issue by developing the concept of loss function proposed by G. Taguchi through applying the methods of machine learning to design advanced tools for analyzing quality management processes.

2. The concept of economic loss function
Genichi Taguchi suggested losses appear in case of any, even the least possible deviation from the ideal condition of a parameter, and that the loss value can be determined by the following formula:

$$L(y) = k(y - m)^2; [5]$$

where \((L)y\) is the value of loss in the point \(y\), \(m\) is the point corresponding to the minimum loss (in the general case, \(L(m)=0\)), and \(k\) is a certain constant that reflects the specifics of this process.

Because in statistics the term "loss function" means losses that appear in case of making a wrong decision as the result of analyzing the observed data, the term "economic loss functions" is proposed to
avoid ambiguity. So, now the economic loss function is a function that determines the dependence between the value of deviation from the goal of a process or a product's parameter and the value of the losses related to this deviation. The goal of a process or a product's parameter is the value, under which the considered object is in its ideal condition form the consumer's point of view.

This way, the loss function (or economic loss function) was presented by G. Taguchi as a quadratic function.

However, practical application of this approach leads us to the following adjustments of G. Taguchi's model:

- because of the various technological processes and phenomena that occur in different halves of the tolerance zone, a loss function can be asymmetrical.
- the economic loss function can be asymmetrical due to asymmetry in the process;
- the function may not necessarily be quadratic, but can, for example, be exponential because losses can increase significantly on the boundaries of tolerance zones due to the following factors: significant increase of losses to sorting and selective assembly, significant increase of losses to reprocessing as a result of possible errors in the manufacturer's measurement systems on the boundaries of tolerance zones, significant increase of losses to fines as a result of possible errors in the consumer's measurement systems on the boundaries of tolerance zones;
- in the case of analysis of such simple products as semi-finished goods that in the future will be reprocessed many times by a producer organization's technical processes, losses can be connected only with machine wear and consumption of main and auxiliary materials of the subsequent operations, and a loss function may be linear with a break in the vicinity of the process's target value (in the middle of the tolerance zone);
- as a result of discrete (stepwise) occurrence of technological factors such as forced change of gear/feed speeds, engagement of an additional operator, enabling additional heating, etc., the function can have a piecewise form, for example piecewise linear.

On the basis of these suggestions, it was concluded that a loss function absolutely does not have to be quadratic, it can have any form: quadratic, exponential, piecewise, etc. It can be asymmetrical with respect to the optimum point and can behave in completely different ways in the vicinity of the tolerance zone boundaries both before and after crossing them, as shown in Figure 1.

![Figure 1. Process model and economic loss functions.](image_url)

To obtain a mathematical model of an economic loss function, it is necessary to carry out the following sequence of actions:

- Determine the list of economic loss items. These items are divided into two classes: additional costs and lost profits. They will be substantially different for different
organizations depending on the structure of their output streams and how products, services, and the information component are located in this stream. Examples of additional costs are costs of additional product processing, sorting, repeated provision of services, information recovery, costs of additional servicing and repair of primary and auxiliary equipment, tools, devices, expendable supplies, supplementary labor compensation and related charges, etc. Examples of lost profits are decrease of the useful production volume due to diversion of production facilities to reprocessing suboptimal products and additional provision of suboptimal services, a decrease of the useful production volume as a result of a decrease in its rhythm due to unscheduled servicing and repair of equipment, replacement of tools, devices, expendable supplies, etc. The most important items of lost profits are losses because of customer dissatisfaction.

- Extraction of economic loss data by items from accounting, managerial, and other forms of records. Added cost and lost profits data are extracted from accounting and managerial records, quality assurance outcomes, technological operation records, and reports on servicing and repair, and consumption and wear of materials, tools, and devices.
- An exact sum of economic loss is calculated for each point on the X axis (i.e. points of a particular deviation from the optimum). To sum up, it can be said that for each point the added costs and lost profits by items for each point are determined and aggregated.
- As a result, the authors obtain a set of points in the space inside and in the vicinity of the tolerance zone, which characterize economic loss.
- For the obtained set, it is necessary to determine the mathematical formula of the function, whose graph will approximate the economic loss function. The obtained formula can be used to forecast economic loss volume in particular batches of products of the same technological process.

Significant difficulty arises at this stage because, as was proven above, these points can be located in the most complex way and it is impossible to determine the economic loss function with sufficient accuracy using standard methods.

3. **Use of machine learning methods to optimize economic loss function simulation**

Machine learning methods can be used to solve this problem [6]. The term machine learning is connected with automatic identification of significant regularities in data. During the past several decades, it became a common tool for solving virtually any task that requires extraction of information from large sets of data. People are surrounded by the machine learning technology: search engines learn how to achieve better results (in advertising), spam filters learn how to filter our messages, while credit card operations are safeguarded by software that identifies indications of fraud. Digital cameras learn to recognize faces and smart personal assistant applications in smartphones learn to recognize voice commands. Cars are equipped with collision prevention systems designed using machine learning algorithms. Machine learning is also broadly used in such areas of science as bioinformatics, medicine, and astronomy [7].

To solve this problem, the author designed a software product, whose algorithm is presented in figure 2.
Figure 2. Algorithm of the software product for analyzing the function of economic loss in processes using machine learning methods.

To analyze and simulate the economic loss function, let us use entities based on the following classes:

- linear regression;
- random forest regression;
- nearest neighbors regression [8];
- support vector machine regression [9];
- neural network [10].

Results of approximation of factual data by different machine learning methods are shown in Figure 3.
Figure 3a clearly shows the superiority of the random forest method over other methods. However, comparison of this method with the neural network does not produce such an obvious conclusion. A number of experiments show the variability of results of these methods, which is natural due to their specifics, has different values. This way, in isolated cases, the neural network is capable of producing results with very high-quality up to the coefficient of determination of 0.95; however, at the same time it is characterized by a very high variance of results. The random forest method does not demonstrate such high upper limit of results but is significantly more stable. In the end, the expected value of the coefficient of determination of the random forest method is significantly higher and its variance is significantly lower than that of the neural network. Furthermore, the random forest method requires several orders of magnitude less computing power than a neutral network. Nevertheless, the choice of a simulation method is always left to the user.

Later on, a trained entity is used to process arrays of data of results of instrument control of product batches. As a result of this processing, the user receives data about the total economic loss in the product batch, its expected value, and standard deviation. On the basis of this information, it becomes possible to make justified decisions concerning the development of processes, in particular, their modernization, investments, etc. Moreover, use of this methodology allows one to lay the foundation for integration of quality management systems and financial management systems of an enterprise.

4. Conclusion
As a result of studying the concept of loss functions proposed by G. Taguchi, a number of adjustments for it were formulated and tested. A software product for analyzing economic loss functions by machine learning methods developed in the process of this research allows one to remove all identified limitations, in particular, entities based on a number of machine learning classes are capable of building a regression with respect to virtually any layout of initial points that characterize the magnitude of economic loss. A comparative analysis of these methods from the perspective of determination and stability of results was carried out. The obtained methodology allows one to obtain objective data about the amounts of economic loss, their expected values, and standard deviation in product batches produced by processes, which in turn allows one to make justified decisions about their development (including investment) and lay the groundwork for integration of quality management systems and financial management systems of an enterprise.

References
[1] ISO 9001:2015, Quality management systems – Requirements
[2] Ezrahovich A Y, Vladimirsev A V, Livshitz I I, Lontsikh P A, Karaseva V A 2017 Risk-based
thinking of ISO 9001:2015 - The new methods, approaches and tools of risk management, 2017, Proceedings of the 2017 International Conference "Quality Management, Transport and Information Security, Information Technologies", IT and QM and IS 2017 8085872, pp. 506-511

[3] Potyomkina M V, Kibirev Y V, Lontsih P A, Kunakov E P, Livshitz 2017 Machinery manufacturing and construction organisations processes and safety efficiency evaluation based on lean manufacturing approach 2017. Proceedings of the 2017 International Conference "Quality Management, Transport and Information Security, Information Technologies", IT and QM and IS 2017, 8085937, pp. 761-765

[4] 2006 Statistical management of processes. SPC. Reference guide. (Nizhny Novgorod: OOO SMTs "Prioritet")

[5] Taguchi Gen’ichi 2016 Taguchi on robust technology development: bringing quality engineering upstream (ASME Press. New York)

[6] Mayer-Schönberger V 2014 Big data. A Revolution That Will Transform How We Live, Work, and Think (Moscow: Mann)

[7] Shai Shalev-Shwartz and Shai Ben-David. Understanding Machine Learning: From Theory to Alghorythms. Cambridge University Press. New York)

[8] Beygelzimer A, Kakade S, and Langford J 2006 Cover trees for nearest neighbor, International Conference on Machine Learning

[9] Wahba G 1997 Support vector machines, reproducing kernel Hilbert spaces and the randomized GACV, Tech. Report 984 (Department of Statistics, University of Wisconsin, Madison)

[10] Haykin S 1998 Neural networks : A comprehensive foundation (Macmillan, New York)