On the application of methods of correlation-regression analysis and fuzzy logic in the analysis of industrial injury rates

B V Sevastyanov, R O Shadrin, V A Lisin

Department of Technosphere Security, Kalashnikov Izhevsk State Technical University, 7, bld. 1, Studencheskaya, Izhevsk 426069, Russia

E-mail: bvd47@mail.ru

Abstract. The paper discusses the results of applying the methods of correlation-regression analysis and fuzzy logic for analyzing the statistical information of industrial injury rates in the Udmurt Republic, and shows the stages of modeling. Correlation-regression analysis was carried out using the MS Office Excel “Data Analysis” package; processing by the Mamdani algorithm (fuzzy logic) was carried out in the Mathlab environment. For the number of workers injured with further disability per 1 working day or more and fatal outcome in the Udmurt Republic, the results of using both methods are shown. Recommendations are given on the application of a particular method in the practical activities of managing the labor protection of organizations as a whole, by type of economic activity or constituent entity of the federation. Recommendations are provided, the implementation of which will improve the accuracy and efficiency of the analysis of industrial injuries by the considered methods.

1. Introduction and Background

The Udmurt Republic is a dynamically developing region where industrial production is stabilizing and the labor market is steadily functioning. All this allows us to form a strategic policy direction in which the idea of labor safety becomes the dominant factor in social well-being.

A preliminary analysis of working conditions in the Udmurt Republic showed that the main causes of high injuries in various sectors of the economy are the low efficiency and imperfection of the labor safety management system.

The degree of favorable working conditions in the Udmurt Republic is not higher than in the surrounding regions. High rates of injuries and deaths are recorded among workers in individual sectors of the economy. The difference between individual types of economic activity is significant.

The relevance of the analysis and assessment of the industrial injuries rate in organizations is confirmed by the modernization of the regulatory framework of labor protection management in the Russian Federation [1,2]. The solutions proposed in the documents are well-known and tested methods for identifying and assessing risks that are widely used in practice [3-9]. Their use is justified for small enterprises or divisions if the case is about some employees or one profession. Analysis and assessment of industrial injuries rate of the enterprise as a whole or of a larger facility, for example, a type of economic activity, a municipality or a constituent entity of the federation is possible, but the result is uninformative, of a general nature, and does not allow managing labor protection with
maximum efficiency [10]. To solve this problem, it is proposed to use evidence-based methods of analysis and assessment of the sample frame [11].

2. Materials and Methods
Depending on the complexity of constructing an adequate mathematical model, all objects can be conditionally divided into two classes: “simple” and “complex”. Simple objects allow the construction of a completely adequate and relatively uncomplicated mathematical model that corresponds to real processes in the object and is suitable for implementation on computer technology. Complex objects, forming a much more extensive class compared to simple ones, have a number of distinctive features.

1. The number of factors influencing the processes occurring in the object is so large, and the relationships between its individual elements are so complicated that creating an adequate mathematical model is very difficult, and sometimes impossible. If such a model can be built, it turns out to be very cumbersome and unacceptable for practical use because the reaction time of the control system to the input action is unacceptably long. On the other hand, ignoring individual facts and relationships in the structure of an object in order to obtain a simpler mathematical model can lead to unjustified idealization of the object and loss of adequacy.

2. Lack of sufficiently accurate and reliable information about the nature of the functioning of the object and the processes occurring in it, or insufficient amount of this information to build an accurate and adequate model using the traditional mathematical apparatus.

3. The existence of a significant part of the information that is necessary for the mathematical description of the object, only in the form of representations of specialists with experience in working with the object in question.

4. In some cases, the conditions affecting the choice of management strategy can be expressed only qualitatively. The impossibility of presenting quantitative ratios of the very goal of management, for example, in the form of any objective function.

The official statistics of industrial injuries in the Udmurt Republic and indicators of the socio-economic development of the region were used as initial data [12-16]. Methods of analysis and assessment of indicators of industrial injuries are correlation-regression analysis and methods of fuzzy logic (Mamdani algorithm).

3. Experimental Section
The work consisted of two stages. The first was to process the formed sample frame by correlation-regression analysis [17, 18]. The following procedures were carried out: elimination of false data and verification of data for their randomness and independence; testing the hypothesis for the normal distribution of the sample; calculation of pair correlation coefficients and verification of their significance using Student and Fisher-Snedecor criteria; correlation-regression analysis, which consists of developing a multiple regression equation and assessing its quality using the multiple determination coefficient, verification of the significance of the coefficients of the linear multiple regression equation and the significance of the linear equation as a whole. The result of the work done are equations that allow for the short-term (2-4 years) forecasting of the industrial injuries rate. For the number of workers injured with further disability per 1 working day or more and fatal outcome (m1), the equation has the following form:

\[ m_1 = 2680.01 - 0.083S - 0.033E - 0.0028V_p \]

where S is the funds spent on labor protection measures per worker (rubles); E – the number of people trained in labor protection in organizations of the Udmurt Republic (person); Vp – gross regional product per capita (rubles). The work of the obtained model on the basis of retrospective data, comparison with actual data and calculated estimated figures of the number of workers injured with further disability per 1 working day or more and fatal outcome for 2020 – 2024 are presented in Figure 1.
The second stage of the work consisted in processing the generated sample frame using fuzzy logic methods (based on the results of a preliminary analysis of the methods, the Mamdani algorithm was chosen as the most suitable for solving the problem) [19,20]. The following procedures were carried out: formation of a base of rules for fuzzy inference; Phasing procedure – introducing fuzziness of input variables; fuzzy conclusion (finding the degree of truth for the premises of each of the rules); activation procedure (finding the truncated membership function for the output variable); accumulation procedure (combining the found truncated functions in order to obtain the final fuzzy set and the resulting membership function); defazification procedure (reduction to clarity) of output variables. The result is an integral index, which is formed from variations of indexes that specify the X and Y axes (Figures 2-5). For the number of workers injured with further disability per 1 working day or more and fatal outcome, the following variations were identified: L - m1, m2 - m1, N - m1, m3 - m1, where L is the average number of workers (person); N – the number of days of disability of workers injured with further disability per 1 working day or more and fatal outcome; m3 – the number of workers with occupational diseases (person); m2 – the number of fatal injuries (person).
Figure 3. Correlation of $m_1$ and $m_2$ indicators with integral index.

Figure 4. Correlation of $m_1$ and $m_3$ indicators with integral index.

Figure 5. Correlation of $m_1$ and $N$ with integral index.

With an increase in the number of workers and in the number of workers injured with further disability per 1 working day or more and fatal outcome, an increase in the integral index reflecting the level of injury occurs. The $m_1$ correlation is greater than the $L$ correlation. With the growth of indicators, an increase of the integral index occurs. $m_1$ shows a significant correlation with the integral
index, since with the maximum value of each of the indicators in the case of $m_2$, the integral index almost does not change, while $m_1$ increases the integral index.

With an increase in the quantitative value of the indicators, the level of the integral index increases. The $m_1$ indicator correlates with the end result more. The indicator $N$ does not correlate, while $m_1$ has a significant correlation with the integral index.

4. Results and Discussion
The application of various methods for analyzing the industrial injuries rate gives qualitatively different results from the point of view of practical activity.

Correlation-regression analysis allows to obtain a short-term forecast of the selected index of industrial injuries. The equation includes a number of independent variables, varying which it is possible to control the selected index. Working with the equation enables to predict changes in the studied index while one or more independent variables is changing. This method allows for more effective planning of activities aimed at improving labor protection in the organization as a whole or in the municipality for the near term.

Fuzzy logic enables to get an integral index that characterizes the correlation between the studied index and the second variable. Its practical application is of a deeper, more comprehensive nature, as it allows to study the nature of the object in detail. Models of this kind will be effective in the long-term planning of labor protection measures in the constituent entities of the federation or for the type of economic activity.

5. Summary and Conclusion
The use of mathematical modeling in predicting the indicators of occupational injuries makes it possible to plan measures for the protection of labor of a preventive nature. The higher the accuracy of forecasting, the more effective labor safety management will be. To improve the accuracy of modeling in the short-term planning of measures for the management of labor protection, it is necessary to take into account the human factor: the probability of an error while performing work associated with the monotony of labor or a decrease in concentration of attention; fatigue during the working day; the employee’s labor in the periods before getting sick and after recovery (end of sick leave), accompanied by a slight deterioration in well-being; psycho-emotional aspect of the labor process, including the climate within the team and personal experiences. For long-term planning of events, it is recommended to take into account the specifics of the industry (if planning is carried out for the type of economic activity) and the features of the socio-economic development of the region, industry and economy (if planning is carried out for the constituent entity of the federation).

The work was carried out at the financial support by Kalashnikov Izhevsk State Technical University within the framework of the grant no. IIIP/20-86-11.

References
[1] GOST 12.0.230.4-2018 Occupational safety standards system (SSBT). OSH management systems. Methods for identifying hazards at various stages of performing works approved by the order of Federal Agency on Technical Regulating and Metrology dated September 7, 2018 No. 577-st “On the implementation of the interstate standard”
[2] GOST 12.0.230.5-2018 Occupational safety standards system (SSBT). OSH management systems. Methods for identifying hazards at various stages of performing works approved by the order of Federal Agency on Technical Regulating and Metrology dated September 7, 2018 No. 578-st “On the implementation of the interstate standard”
[3] Krupic F, Svantesson E, Seffo N, Westin O, Senorski E H 2020 Use of the World Health Organization Checklist–Swedish Health Care Professionals’ Experience: A Mixed-Method Study J of PeriAnesthesia Nursing. http://www.sciencedirect.com/science/article/pii/S1089947219303740
[4] Gul M, Guven B, Guneri A F 2018 A new Fine-Kinney-based risk assessment framework using FAHP-FVIKOR incorporation J of Loss Prevention in the Process Ind 53 3–16. http://www.sciencedirect.com/science/article/pii/S0950423017307489

[5] Kokangül A, Polat U, Dağşuyu C 2017 A new approximation for risk assessment using the AHP and Fine Kinney methodologies Safety Sc 91 24–32. http://www.sciencedirect.com/science/article/pii/S0925753516301503

[6] Wang W, Liu X, Qin Y 2018 A fuzzy Fine-Kinney-based risk evaluation approach with extended MULTIMOORA method based on Choquet integral Comp & Ind Eng 125 111–123. http://www.sciencedirect.com/science/article/pii/S0360833218303917

[7] Seligmann B J, Zhao J, Marmara S G, Corbett T C, Small M, Hassall M, Boadle J T 2019 Comparing capability of scenario hazard identification methods by the PIC (Plant-People-Procedure Interaction Contribution) network metric Safety Sc 119 116–129. DOI:10.1016/j.ssci.2018.10.019

[8] Chartres N, Bero L A, Norris S L 2018 A review of methods used for hazard identification and risk assessment of environmental hazards Environment Int 123 231–239 DOI:10.1016/j.envint.2018.11.060

[9] Németh E, Seligmann B J, Hockings K, Oakley J, O’Brien C, Hangos K M, Cameron I T 2011 Generating cause-implication graphs for process systems via blended hazard identification methods Comp Aided Chemical Eng 29 1070–1074. DOI:10.1016/B978-0-444-53711-9.50214-5

[10] Andreasen P, Rasmussen B 1990 Comparison of methods of hazard identification at plant level J of Loss Prevention in the Process Ind 3(4) 339–344. DOI:10.1016/0950-4230(90)80002-R

[11] Pasman X H J, Rogers W J, Mannan M S 2018 How can we improve process hazard identification? What can accident investigation methods contribute and what other recent developments? A brief historical survey and a sketch of how to advance J of Loss Prevention in the Process Ind 55 80–106 DOI:10.1016/j.jlp.2018.05.018

[12] Official statistics. Territorial body of the Federal State Statistics Service for the Udmurt Republic. http://udmstat.gks.ru

[13] Development of a strategy for socio-economic development of the Udmurt Republic for the period until 2025 dated September 29, 2009. Government of the Udmurt Republic. http://www.udmurt.ru/

[14] Regions of Russia. Socio-economic indicators. Federal State Statistics Service of the Russian Federation. http://www.gks.ru

[15] Labor and employment in Russia. Federal State Statistics Service of the Russian Federation. http://www.gks.ru

[16] The report "The state of conditions and labor protection in the Udmurt Republic and measures for their improvement". Ministry of Labor of the Udmurt Republic. http://mintrud.udmurt.ru/

[17] Elvik R, Goel R 2019 Safety-in-numbers: An updated meta-analysis of estimates Accident Analysis & Prevention 129 136–147. DOI:10.1016/j.aap.2019.05.019

[18] Singh A, Misra S C 2020 A Dominance based Rough Set analysis for investigating employee perception of safety at workplace and safety compliance Safety Sc 127. DOI:10.1016/j.ssci.2020.104702

[19] Ilbahar E, Karaşan A, Cebi S, Kahraman C 2018 A novel approach to risk assessment for occupational health and safety using Pythagorean fuzzy AHP & fuzzy inference system Safety Sc 103 124–136. DOI:10.1016/j.ssci.2017.10.025

[20] Gallab M, Bouloiz H, Alaoui Y L, Tiouat M 2019 Risk Assessment of Maintenance activities using Fuzzy Logic Procedia Comp Sc 148 226–235. http://www.sciencedirect.com/science/article/pii/S1877050919300687