Научная статья
Original article
УДК 368:519.86
doi: https://doi.org/10.55186/2413046X_2026_11_5_67
edn: YQVDUI

МОДЕЛИРОВАНИЕ ТАРИФНОЙ ПОЛИТИКИ СТРАХОВЫХ КОМПАНИЙ НА ОСНОВЕ КЛАСТЕРНОГО И РЕГРЕССИОННОГО АНАЛИЗА

MODELING INSURANCE PRICING POLICIES USING CLUSTER AND REGRESSION ANALYSIS

Игонина Татьяна Романовна, к.ф-м.н., доцент кафедры Высшей математики Института искусственного интеллекта, Федеральное государственное бюджетное образовательное учреждение высшего образования «МИРЭА – Российский технологический университет», Москва
Кесельман Владимир Михайлович, к.ф-м.н., доцент кафедры Высшей математики Института искусственного интеллекта, Федеральное государственное бюджетное образовательное учреждение высшего образования «МИРЭА – Российский технологический университет», Москва
Джиоева Мария Ивановна, к.ф-м.н., доцент кафедры Высшей математики Института искусственного интеллекта, Федеральное государственное бюджетное образовательное учреждение высшего образования «МИРЭА – Российский технологический университет», Москва
Воронов Геннадий Борисович, к.т.н., доцент кафедры высшей математики Института искусственного интеллекта, Федеральное государственное бюджетное образовательное учреждение высшего образования «МИРЭА – Российский технологический университет», Москва

Igonina Tatyana Romanovna, PhD, Associate Professor of the Department of Higher Mathematics at the Institute of Artificial Intelligence, Federal State Budgetary Educational Institution of Higher Education "MIREA – Russian Technological University", Moscow

Keselman Vladimir Mikhailovich, PhD, Associate Professor of the Department of Higher Mathematics at the Institute of Artificial Intelligence, Federal State Budgetary Educational Institution of Higher Education "MIREA – Russian Technological University", Moscow

Dzhioeva Maria Ivanovna, PhD, Associate Professor of the Department of Higher Mathematics at the Institute of Artificial Intelligence, Federal State Budgetary Educational Institution of Higher Education "MIREA – Russian Technological University", Moscow

Voronov Gennady Borisovich, PhD, Associate Professor of the Department of Higher Mathematics at the Institute of Artificial Intelligence, Federal State Budgetary Educational Institution of Higher Education "MIREA – Russian Technological University", Moscow

Аннотация. Цель исследования заключается в разработке методики формирования оптимальной стоимости страховых услуг на основе методов кластерного и регрессионного анализа. Актуальность работы обусловлена необходимостью повышения точности тарифной политики страховых компаний в условиях усиления конкуренции и роста неопределенности страхового рынка. В качестве факторов, определяющих стоимость страхового полиса, использованы процент урегулированных убытков и средняя сумма страховых выплат. Методологическую основу исследования составили методы многомерной статистики, иерархический кластерный анализ и эконометрическое моделирование. На основе стандартизированных данных выполнена кластеризация страховых компаний с использованием метода дальнего соседа и метрики «блок». Для каждой выделенной группы построены линейные и нелинейные регрессионные модели, описывающие

зависимость стоимости полиса от ключевых показателей страховой деятельности. В результате анализа выявлены три устойчивых кластера страховых компаний, различающихся по структуре выплат и уровню урегулирования убытков. Установлено, что характер зависимости стоимости полиса от исследуемых факторов является неоднородным: для отдельных кластеров наилучшие результаты демонстрируют линейные модели, тогда как для других более высокой точностью обладают нелинейные зависимости. Полученные модели обеспечивают высокий уровень объясняющей и прогностической способности. Практическая значимость исследования заключается в возможности использования разработанного кластерно-регрессионного подхода для обоснования тарифной политики, прогнозирования стоимости страховых услуг и повышения эффективности управления страховым портфелем. Предложенная методика обладает универсальным характером и может быть адаптирована к различным видам страхования при наличии соответствующих статистических данных.

Abstract. The study aims to develop a methodology for determining optimal insurance service pricing based on cluster analysis and regression modeling techniques. The relevance of the research is driven by the growing need to improve pricing accuracy in insurance markets characterized by increasing competition and uncertainty. The percentage of settled claims and the average claim payment amount were selected as the primary determinants of insurance policy cost. The methodological framework combines multivariate statistical analysis, hierarchical clustering, and econometric modeling. Insurance companies were classified using standardized data, the furthest-neighbor clustering method, and the block distance metric. Linear and nonlinear regression models were subsequently developed for each cluster to describe the relationship between policy cost and key insurance performance indicators. The empirical analysis identified three stable clusters of insurance companies with distinct claim settlement and payment characteristics. The results indicate that the relationship between policy cost and explanatory

variables is heterogeneous across clusters. Linear specifications provide the best fit for certain groups, whereas nonlinear models demonstrate superior predictive performance for others. The developed models exhibit strong explanatory power and forecasting accuracy. The practical significance of the study lies in the applicability of the proposed cluster–regression framework for insurance pricing, policy cost forecasting, and portfolio management optimization. The methodology is universal in nature and can be adapted to various insurance segments provided that relevant statistical data are available.

Ключевые слова: страхование, страховой полис, ценообразование, кластерный анализ, регрессионное моделирование, эконометрика, тарифная политика, прогнозирование

Keywords: insurance, insurance policy, pricing, cluster analysis, regression modeling, econometrics, tariff policy, forecasting

Introduction

The contemporary insurance market is characterized by intense competition, ongoing digital transformation, and increasing requirements for pricing accuracy. Since insurance companies generate the majority of their revenues through the provision of insurance services, optimal pricing has become a critical determinant of financial stability and competitive positioning [1, 4, 14]. An inappropriate tariff rate may result either in the loss of customers and market share or in reduced profitability and increased exposure to financial losses. This challenge is particularly pronounced in the motor insurance sector, where policy pricing is strongly influenced by macroeconomic conditions, regional disparities, policyholder behavior, and claims experience [2, 9].

In the Russian compulsory motor third-party liability insurance (OSAGO) market, the determination of adequate tariff rates has both commercial and socioeconomic significance. Insurance premiums are influenced by numerous factors, including claim frequency, fluctuations in automobile spare-part prices, regional adjustment coefficients, and inflation expectations. According to data

published by the Russian Association of Motor Insurers (RAMI), the average OSAGO claim payment in the Republic of Tatarstan exceeded RUB 111,000 in 2024, reflecting a nationwide trend toward increasing insurance liabilities [3]. In addition to the growth in nominal claim payments, the structure of losses has become increasingly complex, characterized by a rising proportion of high-severity claims, increasing repair costs, and a growing number of fraudulent claims [7, 8]. These developments necessitate the continuous improvement of claim forecasting methodologies and tariff adjustment mechanisms in response to changing economic conditions.

Considerable attention has been devoted in the international literature to the modeling of insurance risks and policy pricing. Generalized Linear Models (GLMs) remain an industry standard due to their interpretability and robustness; however, their predictive performance is constrained by assumptions of linearity and factor independence [2, 6]. In recent years, machine learning and advanced statistical modeling techniques have gained prominence because of their ability to capture nonlinear relationships, feature interactions, and high-dimensional data structures. Particularly notable among these approaches are decision trees, gradient boosting algorithms, random forests, and regularized regression techniques such as Lasso, Ridge, and Elastic Net [5, 6, 10, 12]. These methods have demonstrated strong performance in predicting both claim frequency and claim severity, especially within the motor insurance sector, where extensive datasets and numerous risk factors are available.

Growing interest has also been observed in modeling frameworks capable of capturing specific characteristics of insurance data, particularly zero inflation, where a substantial proportion of policies generate no claims, and the asymmetric distribution of claim severity [13]. The development of Tweedie models, zero-inflated Poisson models, and hurdle models has significantly improved the accuracy of claim distribution modeling and insurance pricing forecasts [13, 15]. Another important research direction involves the utilization of telematics data that

reflect actual driving behavior. Such information enables the transition from traditional average-based rating factors to personalized risk assessment models, thereby supporting more individualized insurance pricing strategies [3, 12].

In the context of insurance sector digitalization, data mining and clustering techniques have experienced rapid development. Cluster analysis enables the identification of homogeneous groups of policyholders or insurance companies based on a combination of characteristics, including claim frequency, average claim amount, geographical attributes, and customer profiles. Such segmentation creates opportunities for the development of differentiated tariff structures and a more equitable distribution of insurance risk burdens [1, 10]. Recent studies indicate that cluster-based segmentation improves tariff stability, reduces discrepancies between expected and actual claim payments, and facilitates the identification of companies with inefficient portfolio structures [10, 11].

Despite the substantial progress achieved in insurance modeling methodologies, a considerable proportion of existing research focuses on Western insurance markets, where extensive telematics and behavioral datasets are available [3, 5, 12]. In contrast, the Russian OSAGO market is characterized by limited access to detailed claim-level information, increasing the importance of approaches based on aggregated statistical indicators. Under such conditions, econometric methods become particularly valuable, as they enable the identification of pricing patterns and latent relationships between indicators of an insurer’s financial stability and service pricing using a relatively small set of key variables, such as the percentage of settled claims and the average claim payment amount [8, 9, 14].

Beyond their practical relevance for tariff development, studies in this area also possess significant theoretical value. Examination of the relationship between average policy cost and claim payment structure contributes to a deeper understanding of market equilibrium mechanisms and competitive dynamics within the insurance industry [4, 14]. Modeling the dependence of insurance policy

pricing on a combination of macroeconomic and microeconomic factors supports the development of rational risk management strategies and insurance portfolio optimization frameworks. Furthermore, contemporary research increasingly emphasizes the importance of ethical and regulatory considerations, including requirements for fairness and non-discrimination in algorithmic pricing systems [15]. These considerations further highlight the relevance of studies aimed at improving the transparency and economic justification of insurance tariffs, particularly in compulsory mass-insurance segments, where issues of social equity and financial sustainability are closely interconnected.

Consequently, current industry trends—including the technological transformation of insurance markets, intensifying competition, expanding data availability, and increasing complexity of claim structures—create a growing demand for comprehensive investigations of insurance pricing mechanisms. The optimization of insurance service pricing has evolved beyond a purely profitability-oriented objective and has become an important instrument for ensuring the sustainable development of the insurance sector. Within this context, particular attention should be given to approaches that integrate quantitative analysis, statistical modeling, and grouping techniques, thereby enabling the identification of stable patterns in policy price formation based on the key operational characteristics of insurance companies [1, 2, 6, 10].

1. Methods.

The study employed multivariate statistical techniques and econometric modeling methods for data analysis. At the first stage, cluster analysis was applied to classify insurance companies according to the similarity of two key indicators: the percentage of settled claims and the average claim payment amount. A hierarchical clustering algorithm based on the furthest-neighbor linkage criterion and the block distance metric was selected, enabling the sequential formation of clusters without requiring a predefined number of groups. Prior to the analysis, all variables were standardized using z-score normalization to eliminate the influence

of differences in measurement scales. Based on the examination of the dendrogram and within-cluster variance, three clusters were identified, differing in terms of the relationship between claim frequency and claim severity. This approach made it possible to reveal the internal structure of the sample and to identify typical patterns of insurance company behavior in the OSAGO market.

At the second stage of the study, regression models were developed to describe the relationship between the average insurance policy cost and the selected explanatory variables within each cluster. Both linear and nonlinear regression specifications were estimated for every group. Model selection was based on the coefficient of determination (R^2), the standard approximation error, and visual inspection of residual distributions. In cases where linear specifications did not provide satisfactory accuracy, exponential and cubic functional forms were employed to better capture the observed empirical relationships. This approach improved forecasting performance and enabled a more adequate representation of the nonlinear effects associated with variations in risk-related parameters and their impact on insurance policy pricing.

2. Analysis of statistical indicators of insurance companies

The dataset consists of 20 insurance companies (Table 1). For each company i, the average insurance policy cost, denoted as y_i (response variable), and two explanatory variables are available: x_{i1}, representing the percentage of settled claims, and x_{i2}, representing the average claim payment amount.

Table 1 - Initial Dataset

Company	i	x_{i1}	x_{i2}	y_i	Company	i	x_{i1}	x_{i2}	y_i
Rosgosstrakh	1	4,59	24,8	1,7	Rossiya	11	7,99	13,2	1,8
RESO-Garantia	2	6,11	24,2	3,2	Renaissance Insurance	12	4,9	33,8	3,4
Ingosstrakh	3	6,12	28,1	3,4	Moscow Insurance Company	13	6,39	25,7	2,7
ROSNO	4	7,41	25,4	3,3	GUTA Insurance	14	4,67	28,8	2,4
Uralsib	5	5,41	25,5	2,2	ORANTA	15	9,88	11,5	2,1

Spasskie Vorota	6	5,34	20,4	1,9	General Insurance Alliance	16	11,37	12,7	3,2
Russkiy Mir	7	11,92	10,7	2,2	Avikos-AFES	17	10,71	12,6	2,2
Standard-Reserve	8	6,57	26,6	2,8	Rosstrakh	18	4,2	26,7	1,8
MAX Group	9	5,51	25,8	2,3	Energogarant	19	2,91	26,6	1,4
Soglasie	10	9,25	16,8	2,3	ASKO	20	5,03	24,5	2

To identify patterns in the formation of OSAGO policy prices across different insurance companies, a classification procedure was conducted using cluster analysis techniques [16] implemented in the SPSS statistical software environment. A hierarchical clustering approach was employed because the optimal number of clusters was not known a priori, with the range of plausible solutions varying from two to four clusters. The clustering procedure was performed using the response variable and the two explanatory variables. The furthest-neighbor method was selected as the agglomeration criterion, while the block distance metric was used to measure dissimilarity between observations. To eliminate the influence of differences in measurement scales, all variables were standardized using z-score normalization.

Analysis of the dendrogram (Figure 1) indicated that the highest degree of within-group homogeneity was achieved when the insurance companies were partitioned into three clusters. The first cluster consisted of eight companies: Rosgosstrakh, Uralsib, Spasskie Vorota, MAX Group, GUTA Insurance, Rosstrakh, Energogarant, and ASKO. The second cluster included six companies: RESO-Garantia, Ingosstrakh, ROSNO, Standard-Reserve, Renaissance Insurance, and Moscow Insurance Company. The remaining six companies—Russkiy Mir, Soglasie, Rossiya, ORANTA, General Insurance Alliance, and Avikos-AFES—formed the third cluster.

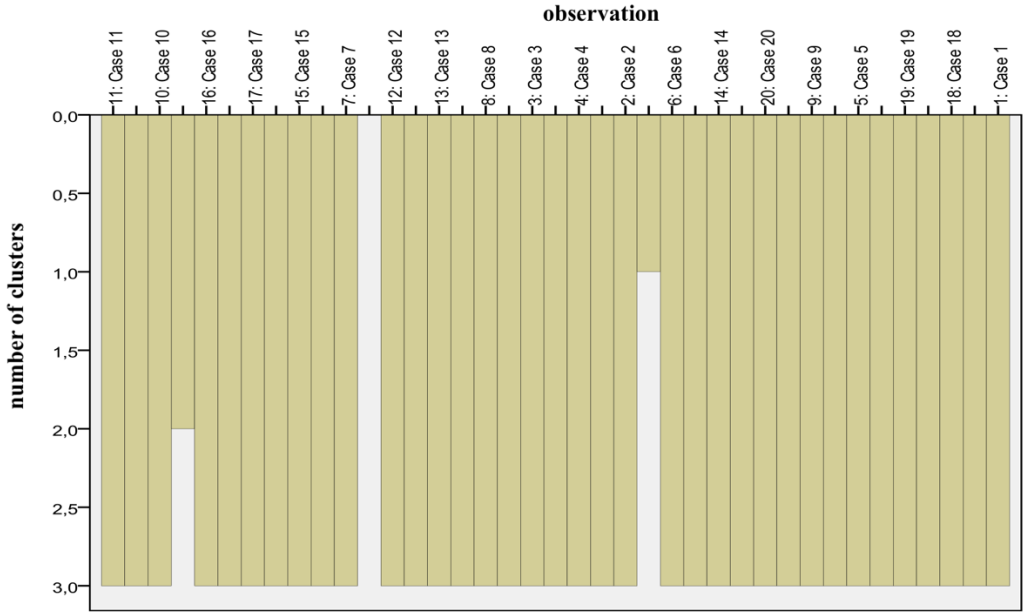

Figure 1 - Dendrogram of Insurance Company Clustering

Following cluster identification, a comparative analysis of the resulting groups was conducted. Table 2 summarizes the descriptive statistics of the three variables under consideration, including their mean, minimum, and maximum values for each cluster.

Table 2 - Descriptive Statistics of the Variables by Cluster

number of clusters	$x_{i1\text{cp}}$	x_{i1min}	x_{i1max}	$x_{i2\text{cp}}$	x_{i2min}	x_{i2max}	$y_{i2\text{cp}}$	y_{i2min}	y_{i2max}
1	4,71	2,91	5,51	25,39	20,4	28,8	1,96	1,4	2,4
2	6,25	4,9	7,41	27,3	24,2	33,8	3,13	2,7	3,4
3	10,19	7,99	11,92	12,92	10,7	16,8	2,3	1,8	3,2

Analysis of the data presented in Table 2 indicates that the first cluster is characterized by the lowest percentage of settled claims and a moderate level of average claim payments. Accordingly, this group of companies exhibits the lowest average OSAGO policy cost among all three clusters. The second cluster is distinguished by an intermediate level of claim settlement combined with the highest average claim payment amount. These characteristics result in the highest average OSAGO policy cost within the sample. The third cluster demonstrates the

highest claim settlement rate while maintaining the lowest average claim payment amount. The average policy cost for this group occupies an intermediate position relative to the other clusters.

Table 2 also provides interval estimates for each variable across all clusters, allowing a more detailed assessment of the variation range of the analyzed indicators. Consequently, given known values of the explanatory variables x_1 (percentage of settled claims) and x_2 (average claim payment amount), a particular insurance company can be assigned to one of the three identified clusters.

For each cluster, linear and nonlinear regression models were developed using the SPSS statistical software package in order to quantitatively describe the relationship between the average OSAGO policy cost and the two selected explanatory variables: the percentage of settled claims (x_1) and the average claim payment amount (x_2).

1. Linear Model

For the first cluster, the linear regression model demonstrated a high degree of agreement with the observed data. The coefficient of determination reached $R^2 = 0.902$, while the relative model error did not exceed 0.3%. These results indicate that the linear regression equation provides an adequate statistical representation of the relationship under investigation:

$$y = 0.395x_1 + 0.086x_2 - 2.084$$

where y denotes the average OSAGO policy cost.

The high value of R^2 indicates that the linear model explains more than 90% of the variation in the dependent variable, confirming the statistical significance of the selected explanatory factors.

2. Nonlinear Model

To obtain a more accurate representation of the relationship between the variables, the curve-fitting procedure was applied to identify the most appropriate functional form linking the dependent and explanatory variables.

Dependence on x_1.

The coefficient of determination ($R^2 = 0.64$) and a model error of 1.7% indicate a pronounced exponential relationship. The corresponding model can be expressed as

$$y = a\exp(x_1) + b$$

This result suggests that an increase in the percentage of settled claims exerts a nonlinear effect on the average insurance policy cost.

Dependence on x_2.

Based on the obtained value of $R^2 = 0.129$, a cubic specification was selected:

$$y = ax_2^3 + bx_2^2 + cx_2 + d$$

Although the coefficient of determination is relatively low, the cubic form allows nonlinear effects associated with variations in the average claim payment amount to be incorporated into the model.

3. Combined Nonlinear Model

Taking both explanatory variables into account, a combined nonlinear regression model was constructed to describe the average policy cost within the first cluster:

$$y = 0.04\exp(x_1) + 0.07x_2^3 - 0.52x_2^2 + 12.513x_2 - 98.801$$

The coefficient of determination for this specification is $R^2 = 0.968$, which is close to unity and exceeds the value obtained for the linear model. This finding indicates the superior adequacy of the nonlinear specification, which more accurately reproduces the observed data and captures the complex relationships among the variables.

4. Methodological Significance

The application of the combined regression model provides several methodological advantages:

1. Quantitative forecasting of OSAGO policy costs for insurance companies belonging to the first cluster based on known values of x_1 and x_2.

2. Identification of the dominant functional form associated with each explanatory factor and assessment of its influence on the dependent variable.

3. Objective comparison of linear and nonlinear model performance through the coefficient of determination and model error measures, thereby facilitating the evaluation of model adequacy.

Regression Analysis for the Second and Third Clusters

Second Cluster

For the second cluster, the linear regression model demonstrated limited explanatory power:

$$y = 0.05x_1 + 0.049x_2 + 1.475$$

The coefficient of determination was $R^2 = 0.204$, while the relative model error reached 71%, indicating a poor fit to the observed data.

In contrast, the nonlinear regression model provided a substantially more accurate description of the observations:

$$y = 1.05x_1^3 - 19.443x_1^2 + 118.586x_1 + 0.69x_2 - 236.654$$

with a coefficient of determination of $R^2 = 0.926$.

Therefore, the nonlinear regression specification should be preferred for forecasting the average OSAGO policy cost within the second cluster. The results demonstrate that, for this group of insurance companies, the nonlinear model substantially outperforms the linear alternative in terms of explanatory and predictive accuracy.

Third Cluster:

For the third cluster, the following linear regression model was obtained:

$$y = 0.275x_1 + 0.109x_2 - 1.912$$

with a coefficient of determination of $R^2 = 0.51$ and a relative error of 34.2%.

The nonlinear model is expressed as

$$y = -0.065x_2^2 + 1.953x_2 - 12.272$$

with a coefficient of determination of $R^2 = 0.498$, which is slightly lower than that obtained for the linear specification. Based on these indicators, the linear model provides a more adequate representation of the observed data.

However, given the relatively limited explanatory power of both models, forecasting the policy cost for companies belonging to the third cluster is more

appropriately performed using a nonlinear regression equation estimated from the complete dataset prior to cluster partitioning:

$$y = 0.01x_1^3 - 0.035x_1^2 + 0.783x_1 + 0.131x_2 - 4.354$$

The application of this model provides a more accurate description of the relationship between the average policy cost and the explanatory variables, namely the percentage of settled claims (x_1) and the average claim payment amount (x_2).

The proposed cluster analysis and regression modeling framework can be applied universally to the valuation of virtually any insurance product. Implementation of the approach requires only information on the number of insurance contracts, the number of insured events, and the corresponding claim payments. Based on these indicators, insurance companies can be classified according to key performance measures such as the percentage of settled claims and the average claim payment amount. In the present study, three clusters characterized by different levels of within-group homogeneity were identified.

For each cluster, the mean, minimum, and maximum values of the analyzed variables were calculated, providing an objective basis for distinguishing the characteristics of the identified groups. For the first cluster, the linear regression model demonstrated a high degree of agreement with the observed data; however, the combined nonlinear specification provided a more accurate representation of the underlying relationships, as evidenced by its higher coefficient of determination. For the second cluster, the linear model proved inadequate, whereas the nonlinear specification accurately described the observed patterns and can therefore be employed for policy cost forecasting. In the third cluster, the linear model provided a better fit than the locally estimated nonlinear alternative; nevertheless, improved predictive performance can be achieved through the use of the nonlinear equation estimated from the complete dataset prior to cluster segmentation.

Consequently, the proposed methodology enables adaptive forecasting of insurance policy costs according to the cluster membership of a particular

insurance company while accounting for the specific influence of key explanatory factors. The combined application of linear and nonlinear regression models facilitates the selection of the most appropriate specification for describing the dependence of average policy cost on the percentage of settled claims and the average claim payment amount. This feature makes the proposed methodology both universal and practically applicable to the analysis and pricing of a wide range of insurance products.

Conclusion

The results of the study confirm that insurance service pricing is determined by a combination of factors related to claim experience, payment structures, and the characteristics of an insurer’s portfolio. The application of cluster analysis and regression modeling made it possible to identify stable patterns in the distribution of insurance companies according to key performance indicators and to establish typical relationships that define the optimal range of tariff values.

The three identified clusters of insurance companies differ in terms of the relationship between the percentage of settled claims and the average claim payment amount. The first cluster is characterized by a low level of claims and moderate claim payments, corresponding to the lowest average policy cost. The second cluster exhibits the highest average claim payments and an intermediate claim settlement rate, resulting in the highest tariff levels. The third cluster occupies an intermediate position, combining a high claim settlement rate with relatively low claim severity, which leads to a moderate average policy cost. This distribution reflects natural market mechanisms in which the level of risk determines both tariff structures and insurance company strategies.

A comparison of linear and nonlinear regression models revealed that the relationship between insurance policy cost and the analyzed factors is heterogeneous in nature. For certain clusters, linear specifications provide an adequate description of the observed relationships, with policy costs changing proportionally to payment levels. For other clusters, exponential and cubic models

produce superior results by accounting for nonlinear effects associated with changes in risk parameters. These findings demonstrate that insurance pricing processes cannot be reduced to simple functional relationships and require an adaptive modeling framework capable of capturing the specific characteristics of individual market segments and insurance portfolios.

The practical significance of the proposed methodology lies in the possibility of its direct application by insurance companies. Given information on the percentage of settled claims and the average claim payment amount, an insurer can determine cluster membership and estimate the expected policy cost. Such an approach facilitates tariff optimization, improves the balance between risk and profitability, and enhances the transparency and predictability of financial outcomes. The methodology may be implemented as an instrument of internal analytical control or as a component of strategic portfolio management systems.

The findings support the effectiveness of cluster–regression techniques for the analysis and forecasting of insurance service pricing. These methods combine statistical rigor with practical applicability, which is particularly important under conditions of limited data availability and high market volatility. Future research may extend the proposed framework by incorporating additional demographic, macroeconomic, regional, and behavioral factors, as well as by adapting the methodology to other insurance segments.

More broadly, the results demonstrate that the application of quantitative analytical methods contributes to improving the economic validity of insurance tariffs, reducing market inefficiencies, and strengthening the financial stability of insurance institutions. The systematic use of such tools can promote a balance between the commercial objectives of insurers and the social function of insurance, thereby creating the foundation for a more transparent, rational, and sustainable development of the insurance industry.

Список источников

1. Jamotton C., Hainaut D., Hames T. Insurance Analytics with Clustering Techniques. Risks. 2024; 12(9): 141. DOI: 10.3390/risks12090141.

2. Wilson A.A., Nehme A., Dhyani A., Mahbub K. A Comparison of Generalised Linear Modelling with Machine Learning Approaches for Predicting Loss Cost in Motor Insurance. Risks. 2024; 12(4): 62. DOI: 10.3390/risks12040062.

3. Boylan J., Meyer D., Chen W.S. A Systematic Review of the Use of In-Vehicle Telematics in Monitoring Driving Behaviours. Accident Analysis & Prevention. 2024; 199: 107519. DOI: 10.1016/j.aap.2024.107519.

4. Hodula M., Janků J., Časta M., Kučera A. On the Macrofinancial Determinants of Life and Non-Life Insurance Premiums. The Geneva Papers on Risk and Insurance. 2021. DOI: 10.1057/s41288-021-00249-z.

5. Meng S., Gao Y., Huang Y. Actuarial Intelligence in Auto Insurance: Claim Frequency Modeling with Driving Behavior Features and Improved Boosted Trees. Insurance: Mathematics and Economics. 2022; 106: 115–127. DOI: 10.1016/j.insmatheco.2022.06.001.

6. Henckaerts R., Côté M.-P., Antonio K., Verbelen R. Boosting Insights in Insurance Tariff Plans with Tree-Based Machine Learning Methods. North American Actuarial Journal. 2021; 25(2): 255–285. DOI: 10.1080/10920277.2020.1745656.

7. Zinyoro T., Aziakpono M.J. Performance Determinants of Non-Life Insurance Firms: A Systematic Review of the Literature. Cogent Business & Management. 2024; 11(1): 2345045. DOI: 10.1080/23311975.2024.2345045.

8. Bryzgalov D.V., Gryzenkova Yu.V., Tsyganov A.A. Consumer Limitations on the Digitalization of the Insurance Market and Ways to Overcome Them. Studies on Russian Economic Development. 2022; 33(5): 539–546. DOI: 10.1134/S1075700722050057.

9. Tulenty D.S. К вопросу об адекватности цен и маржинальности в ОСАГО. Имущественные отношения в Российской Федерации. 2020;(10): 40–51. DOI: 10.24411/2072-4098-2020-11002.

10. Clemente C., Guerreiro G.R., Bravo J.M. Modelling Motor Insurance Claim Frequency and Severity Using Gradient Boosting. Risks. 2023; 11(9): 163. DOI: 10.3390/risks11090163.

11. Poufinas T., Gogas P., Papadimitriou T., Zaganidis E. Machine Learning in Forecasting Motor Insurance Claims. Risks. 2023; 11(9): 164. DOI: 10.3390/risks11090164.

12. Usman F., Chan J.S.K., Makov U.E., Wang Y., Dong A.X.D. Claim Prediction and Premium Pricing for Telematics Auto Insurance Data Using Poisson Regression with Lasso Regularisation. Risks. 2024; 12(9): 137. DOI: 10.3390/risks12090137.

13. Deng M. Inference for the Parameters of a Zero-Inflated Poisson Model: Bayesian and Frequentist Approaches. Risks. 2024; 12(7): 104. DOI: 10.3390/risks12070104.

14. Siopi E., Poufinas T. Impact of Internal and External Factors on the Profitability and Financial Strength of Insurance Groups. International Advances in Economic Research. 2023; 29(3): 129–149. DOI: 10.1007/s11294-023-09873-y.

15. Xin X. Antidiscrimination Insurance Pricing: Regulations, Fairness and Practical Approaches. North American Actuarial Journal. 2024. DOI: 10.1080/10920277.2023.2190528.

16. Мхитарян В.С., Архипова М.Ю., Сиротин В.П. Эконометрика М.: ЕАОИ, 2008

17. Математическое моделирование аномальных доходностей криптовалют под воздействием инфоповодов / А. А. Сидоров, Т. А. Бурцева, Е. С. Дарда, О. Р. Параскевопуло // Московский экономический журнал. – 2026. – Т. 11, № 2. – С. 204-241. – DOI 10.55186/2413046X_2026_11_2_27. – EDN ZESDHX.

18. Сидоров, А. А. Сравнительный анализ двух комбинаторных методов вычисления математического ожидания наилучшего ранга в задаче о размещении / А. А. Сидоров, Р. У. Астафьев // Оптические технологии, материалы и системы (Оптотех - 2025) : Международная научно-техническая

конференция, Москва, 08–12 декабря 2025 года. – Москва: МИРЭА - Российский технологический университет, 2026. – С. 1125-1131. – EDN ACIABJ.

19. Астафьев, Р. У. Имитационное моделирование сложных иерархических систем / Р. У. Астафьев // НАУКА XXI ВЕКА: ВЫЗОВЫ, СТАНОВЛЕНИЕ, развитие: сборник статей XXIV Международной научно-практической конференции, Петрозаводск, 24 ноября 2025 года. – Петрозаводск: Международный центр научного партнерства «Новая Наука» (ИП Ивановская И.И.), 2025. – С. 287-291. – EDN QTROAE.

References

1. Jamotton C., Hainaut D., Hames T. Insurance Analytics with Clustering Techniques. Risks. 2024; 12(9): 141. DOI: 10.3390/risks12090141.

2. Wilson A.A., Nehme A., Dhyani A., Mahbub K. A Comparison of Generalised Linear Modelling with Machine Learning Approaches for Predicting Loss Cost in Motor Insurance. Risks. 2024; 12(4): 62. DOI: 10.3390/risks12040062.

3. Boylan J., Meyer D., Chen W.S. A Systematic Review of the Use of In-Vehicle Telematics in Monitoring Driving Behaviours. Accident Analysis & Prevention. 2024; 199: 107519. DOI: 10.1016/j.aap.2024.107519.

4. Hodula M., Janků J., Časta M., Kučera A. On the Macrofinancial Determinants of Life and Non-Life Insurance Premiums. The Geneva Papers on Risk and Insurance. 2021. DOI: 10.1057/s41288-021-00249-z.

5. Meng S., Gao Y., Huang Y. Actuarial Intelligence in Auto Insurance: Claim Frequency Modeling with Driving Behavior Features and Improved Boosted Trees. Insurance: Mathematics and Economics. 2022; 106: 115–127. DOI: 10.1016/j.insmatheco.2022.06.001.

6. Henckaerts R., Côté M.-P., Antonio K., Verbelen R. Boosting Insights in Insurance Tariff Plans with Tree-Based Machine Learning Methods. North American Actuarial Journal. 2021; 25(2): 255–285. DOI: 10.1080/10920277.2020.1745656.

7. Zinyoro T., Aziakpono M.J. Performance Determinants of Non-Life Insurance Firms: A Systematic Review of the Literature. Cogent Business & Management. 2024; 11(1): 2345045. DOI: 10.1080/23311975.2024.2345045.
8. Bryzgalov D.V., Gryzenkova Yu.V., Tsyganov A.A. Consumer Limitations on the Digitalization of the Insurance Market and Ways to Overcome Them. Studies on Russian Economic Development. 2022; 33(5): 539–546. DOI: 10.1134/S1075700722050057.
9. Tulenty D.S. К вопросу об адекватности цен и маржинальности в ОСАГО. Имущественные отношения в Российской Федерации. 2020;(10): 40–51. DOI: 10.24411/2072-4098-2020-11002.
10. Clemente C., Guerreiro G.R., Bravo J.M. Modelling Motor Insurance Claim Frequency and Severity Using Gradient Boosting. Risks. 2023; 11(9): 163. DOI: 10.3390/risks11090163.
11. Poufinas T., Gogas P., Papadimitriou T., Zaganidis E. Machine Learning in Forecasting Motor Insurance Claims. Risks. 2023; 11(9): 164. DOI: 10.3390/risks11090164.
12. Usman F., Chan J.S.K., Makov U.E., Wang Y., Dong A.X.D. Claim Prediction and Premium Pricing for Telematics Auto Insurance Data Using Poisson Regression with Lasso Regularisation. Risks. 2024; 12(9): 137. DOI: 10.3390/risks12090137.
13. Deng M. Inference for the Parameters of a Zero-Inflated Poisson Model: Bayesian and Frequentist Approaches. Risks. 2024; 12(7): 104. DOI: 10.3390/risks12070104.
14. Siopi E., Poufinas T. Impact of Internal and External Factors on the Profitability and Financial Strength of Insurance Groups. International Advances in Economic Research. 2023; 29(3): 129–149. DOI: 10.1007/s11294-023-09873-y.
15. Xin X. Antidiscrimination Insurance Pricing: Regulations, Fairness and Practical Approaches. North American Actuarial Journal. 2024. DOI: 10.1080/10920277.2023.2190528.

16. Mxitaryan V.S., Arxipova M.Yu., Sirotin V.P. E`konometrika M.: EAOI, 2008
17. Matematicheskoe modelirovanie anomal`ny`x doxodnostej kriptovalyut pod vozdejstviem infopovodov / A. A. Sidorov, T. A. Burceva, E. S. Darda, O. R. Paraskevopulo // Moskovskij e`konomicheskij zhurnal. – 2026. – T. 11, № 2. – S. 204-241. – DOI 10.55186/2413046X_2026_11_2_27. – EDN ZESDHX.
18. Sidorov, A. A. Sravnitel`ny`j analiz dvux kombinatorny`x metodov vy`chisleniya matematicheskogo ozhidaniya nailuchshego ranga v zadache o razmeshhenii / A. A. Sidorov, R. U. Astaf`ev // Opticheskie texnologii, materialy` i sistemy` (Optotex - 2025) : Mezhdunarodnaya nauchno-texnicheskaya konferenciya, Moskva, 08–12 dekabrya 2025 goda. – Moskva: MIRE`A - Rossijskij texnologicheskij universitet, 2026. – S. 1125-1131. – EDN ACIABJ.
19. Astaf`ev, R. U. Imitacionnoe modelirovanie slozhny`x ierarxicheskix sistem / R. U. Astaf`ev // NAUKA XXI VEKA: VY`ZOVY`, STANOVLENIE, razvitie: sbornik statej XXIV Mezhdunarodnoj nauchno-prakticheskoj konferencii, Petrozavodsk, 24 noyabrya 2025 goda. – Petrozavodsk: Mezhdunarodny`j centr nauchnogo partnerstva «Novaya Nauka» (IP Ivanovskaya I.I.), 2025. – S. 287-291. – EDN QTROAE.

© Игонина Т.Р., Кесельман В.М., Джиоева М.И., Воронов Г.Б.,2026.