©2021 International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies

ISSN 2228-9860 eISSN 1906-9642 CODEN: ITJEA8

International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies

http://TuEngr.com

Statistics Application of the Dynamics Socio-Economic Processes: A Case of Russian Insurance Data

Alfira Kumratova[1]*, Elena Popova[1], Elena Khudyakova[2], Igor Vasilenko[1], Victor Saykinov[1]

[1] *Department of Information systems, Kuban State Agricultural University named after I.T. Trubilin, RUSSIA.*

[2] *Department of Applied Informatics, Russian State Agrarian University Moscow Agricultural Academy named after K.A. Timiryazeva, Moscow, RUSSIA.*

**Corresponding Author (Email: AKumratova@yandex.ru)*

Paper ID: 12A3I

Volume 12 Issue 3

Received 19 August 2020
Received in revised form 16 December 2020
Accepted 06 January 2021
Available online 14 January 2021

Keywords:
Statistical indicators; Pre-predictive analysis; R/S analysis; risk criteria; Hurst exponent; Insurance time series; Insurance risk.

Abstract

This article demonstrates the work of the tool for forecasting the dynamics of socio-economic time series, data from insurance companies based on the complex use of both classical and nonlinear statistics. To obtain pre-forecast information about the time series, the authors proposed an analysis of classical statistical coefficients (kurtosis, asymmetry, and variation). Thus, a multi-criteria assessment of the stability of the dynamics of time series is presented. The methods of nonlinear dynamics adapted by the authors are proposed to be used in a multi-criteria (two-criteria) mathematical model. The result of the model's operation is an assessment of the trend stability of the time series. The first criterion reflects the time series's memory depth in the form of a fuzzy set obtained based on the R/S-analysis. The second criterion is the Hurst exponent. A two-criteria approach to assessing the trend stability of time series makes it possible to differentiate them according to the trend stability indicator and select working forecast models.

Disciplinary: Information Systems, Application, and Analysis, Mathematics.

Cite This Article:

Kumratova, A., Popova, E., Khudyakova, E., Vasilenko, I., Saykinov, V. (2021). Statistics Application of the Dynamics Socio-Economic Processes: A Case of Russian Insurance Data. *International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies, 12*(3), 12A3I, 1-9. http://TUENGR.COM/V12/12A3I.pdf DOI: 10.14456/ITJEMAST.2021.51

1 Introduction

This study's relevance is due to the state's need to create favorable conditions for the dynamic development of the insurance market. Life and health insurances of citizens in modern conditions have become fairly popular services. Virtually every bank uses the services of an affiliated insurance company. The demand for insurance services naturally determines the offer, and the marketing variety of insurance products allows the client to choose a product with individual conditions.

In today's conditions, in the presence of such factors as insufficiently developed market relations in insurance services' low competence, there is a rather meager database of available statistical data. Thus, economic risk is a basic concept in the construction and dynamics of the insurance market. The total nature of risk, its constantly changing components, and, accordingly, dynamics require constant adjustment and development of risk management methods in the economy as a whole and insurance activities in particular [10]. The ambiguity of risk management problems in insurance activities led the authors to study this topical topic.

"Structured" information is information obtained about a time series through analysis, generalization, description. It allows the analyst to develop methods and substantiate approaches to obtain the most accurate predictive conclusions [1]. The article examines the values of the time series's aggregated data for the accounting of contracts for all insurance types of STERKH company. The main indicator of the activity of an insurance company is the number of insured clients. The authors investigated both the time series (TS) themselves and the aggregated TS: the general series, separately the TS of insured men and women. Note that the study of averaged (typical) values for predictive conclusions is not effective [2, 3]. Forecast information is determined, first of all, by the sequence of data, which allows you to identify the possibility of the next value in time.

The right policy of insurance organizations should become the basis for conducting numerous Russian insurers' insurance operations and financial stability. Good risk management implies their application based on reliable forecasting.

Forecasting, in turn, presupposes a scientifically sound judgment about the possible states of the economic system in the future, about alternative ways and timing of its implementation, and also involves obtaining qualitative estimates of these states using mathematical and instrumental methods of economics.

The classical (linear) paradigm assumes that the observed evolutionary process's behavior obeys a regular law. The latter ensures the fulfillment of the principle: a small perturbation to a small extent affects the nature of the system's behavior. It is important to note that classical forecasting economic time series (including time series of insurance data) is based on econometrics' mathematical apparatus. Simultaneously, which is fundamentally important, this establishing is carried out on the assumption that the observations that make up the predicted time series are independent. By which the necessary obedience to the normal law is fulfilled. However, the latter is more the exception than the rule for socio-economic TSs, which have long-term memory. There are no complete theories of forecasting time series with memory and, thus, the need to develop adequate methods for their forecasting is undoubtedly relevant.

This study shows the presence of long-term memory in the time series of insurance data. Accordingly, the authors are developing and adapting predictive methods of nonlinear dynamics (the method of "cellular automata", the method of phase portraits) applicable to these time series.

2 Method

This article presents the results of a pre-forecast analysis of the approbation process on real-time series (daily) of personal insurance of the insurance company STERKH, Cherkessk and economic time series of regional offices of the state non-budgetary fund of the Russian Federation (Karachaevsky and Prikubansky districts of the Karachay-Cherkess Republic) for 2015-2018.

For analyzing the dynamics of time series is presented a tool which allows the complex to calculate both "risky" indicators of classical statistics (standard deviation, kurtosis, asymmetry, and variation) and such indicators as the memory depth of the time series and the Hurst exponent (indicators of the new "nonlinear") statistics). The analysis of the calculated indicators makes it possible to reveal the qualitative properties of the studied time series.

A multi-criteria (two-criteria) assessment of the stability of the dynamics of time series is proposed. The first criterion reflects the depth of memory of the time series in the form of a fuzzy number obtained based on the R/S analysis, the second criterion is the Hurst exponent. A two-criteria approach to assessing the trend stability of time series makes it possible to differentiate them according to the trend stability indicator and select them in future working forecast models.

Approbation of the developed algorithm was carried out on the triad of model socio-economic time series of insurance data: the initial time series, the series of its increments, and data aggregates.

The use of classical statistical methods to obtain pre-forecast information about TS is based on calculating such indicators as kurtosis, asymmetry, and variation. These three main coefficients give a multicriteria assessment of the stability of TS dynamics [6, 7]. Let us consider the kurtosis indicator separately. Visualization of the empirical distribution function allows us to distinguish three main areas: 1. [MX-3CKO; MX + 3CKO]; 2. (MX + 3CKO); 3. (MX + 3CKO), where MX = mathematical expectation and CKO = standard deviation. The calculated kurtosis values on intervals 2 and 3 will reveal a “heavy tail”, and the choice of the forecast model and the forecast's reliability will depend on these numerical values [4, 6]. A comparative analysis of the initial time series is obtained characteristics and predictability properties, and its other formations (increment and aggregation) are of interest for further research.

The initial time series of the number of insured persons is denoted by u_i^k where $k = 1$ - daily data values, $k = 2$ - daily data increments, $k = 3$ - aggregated weekly data, $k = 4$ - increments of aggregated weekly data and $i = 1, 2, \dots, n$ (calendar period from 03.11.2015 - 15.12.2019). Similar to the TS data on the number of insured men and women, denote v_i^k and w_i^k respectively. In this way:

u_i^1 - daily data on all insured clients;

u_i^2 - increments of daily data for all insured clients;

u_i^3 - aggregated weekly data of all insured clients;

u_i^4 - increments of the aggregated weekly baseline TS;

v_i^1 - daily data on insured men;

v_i^2 - increments of daily data on insured men;

v_i^3 - aggregated weekly data on insured men;

v_i^4 - increments of the aggregate weekly TS of insured men;

w_i^1 - daily data on insured women;

w_i^2 - increments of daily data on insured women;

w_i^3 - aggregated weekly data on insured women;

w_i^4 - increments of the aggregate weekly TS of insured women.

The mathematical expectation, variance, or standard deviation are the two main risk indicators, which are defined by the Nobel laureate G. Markowitz. Later studies confirm that the same statement applies to the asymmetry coefficients $A = \sigma^{-3} \sum_{s=1}^{n} (W_s - M)^3 P_s$ and excess $E = \sigma^{-4} \sum_{s=1}^{n} (W_s - M)^4 P_s$ where P_s is is the probability (relative frequency) of the value of a random variable equal to W_s, $1 \leq s \leq n$.

Table 1 shows the values of risk indicators for each separately investigated TS: initial TS; the daily number of male (female) clients; rows of increments.

Table 1: Risk' statistic indicators of signed personal insurance contracts for everyday time series.

Time series name	TS Basic	Main TS increments	TS Men	TS increments Men	TS Women	TS increments Women
	TS designation					
	u_i^1	u_i^2	v_i^1	v_i^2	w_i^1	w_i^2
MX	13.8	-0.0013	3.7	0.01	10.1	0.001
DX	124.7	65.5	11.44	15.56	69.01	61.91
CKO	11.16	8.09	3.38	3.94	8.3	7.86
V	0.81	0.61	0.91	0.3	0.82	0.6
A	0.72	-0.13	1	-0.007	0.76	-0.05
E	3.05	3.82	4.02	4.32	3.11	3.96
E for X < (MX+3CKO)*	0	0.51	0	0.54	0	0.45
E [MX-3CKO; MX+3CKO]**	2.39	2.85	2.14	2.62	2.43	2.97
E for X > (MX+3CKO)***	0.66	0.45	1.87	1.16	0.67	0.54

MX - mathematical expectation

DX - variance

CKO - standard deviation

V - coefficient of variation

A - coefficient of asymmetry

E - coefficient of kurtosis

*E for X <(MX + 3CKO) - the proportion of the kurtosis coefficient that is less than the range of values: the sum of the mathematical expectation and three times the standard deviation;

**E [MX-3CKO; MX + 3CKO] - the proportion of the kurtosis coefficient values falling within the given interval;

***E for X> (MX + 3CKO) - the proportion of the kurtosis coefficient that is greater than the range of values: the sum of the mathematical expectation and three times the standard deviation.

A detailed description of the sequential R/S analysis algorithm is presented in [1, 2, 4, 13].

Based on Table 1, the following conclusions can be drawn:

- on average, women are insured more than men and account for 73.2% of the total number of insured clients of the company;

- for all studied daily time series, the coefficient of kurtosis is $E>3$, which indicates the presence of a "heavy tail" and, thus, one can assume the most probable expectation of an event with a large number of insurance contracts;

- for all series of increments, negative values of the asymmetry coefficient were obtained, which in turn means the presence of large values of increments (the difference in the number of contracts from the previous time period to this one);

- for the coefficient of variation in a pairwise comparison of statistical indicators, we have the following: the series's volatility decreases for several increments. This makes it possible for the researcher to work with him to obtain pre-forecast information.

Table 2 presented the results of calculating statistical indicators for aggregated weekly TS.

Table 2: Risk' statistic indicators for aggregated weekly time series and incremental time series.

Time series name	TS "Basic"	Main TS increments	TS "Men"	TS increments "Men"	TS "Women"	TS increments "Women"
	TS designation					
	u_i^3	u_i^4	v_i^3	v_i^4	w_i^3	w_i^4
MX	94.68	-0.14	25.37	-0.09	69.59	-0.058
DX	2834.9	3284.05	240.35	270.28	1518.94	1787.58
CKO	53.2	57.3	15.5	16.44	38.9	42.27
V	0.56	0.38	0.61	0.18	0.56	0.72
A	0.54	-0.18	0.67	-0.24	0.5	-0.13
E	3.3	2.65	3.39	2.92	3.15	2.66
E for X < (MX+3CKO)	0	00	0	0	0	0
E [MX-3CKO; MX+3CKO]	2.23	2.65	2.41	2.92	2.17	2.66
E for X > (MX+3CKO)	1.07	0	0.98	0	0.98	0

Analysis of the calculated indicators in Table 2 allows us to draw the following conclusions:

- in contrast to the time series of daily data of insured persons, the aggregated weekly TSs do not have “heavy tails” in the series of increments, which indicates a smoothing of the data;

- the value of the coefficient of variation does not exceed 72%;

- the range of kurtosis coefficient values is the interval (2.65; 3.4), which determines that the aggregated weekly TSs belong to the normal distribution law.

Taking into account the presence of "heavy tails" in the time series of increments, the authors propose to turn to the methods of nonlinear dynamics [8, 9, 11] for further research and selection of predictive models that have proven themselves quite well both when working with small samples and with large data [2, 4, 5, 14]

3 Result and Discussion

The analysis of the time series of the insured clients of the company, separately insured women and insured men, as well as the time series of increments and aggregated data, carried out with the help of the presented tool, made it possible to identify such characteristics for each series as the presence (absence) of memory, submission (disobedience) to the normal law distribution, presence (absence) of a "heavy tail". Figure 1 shows the development "R/S-analysis" [3], Figure 2 shows the results of the R/S-Analysis tool.

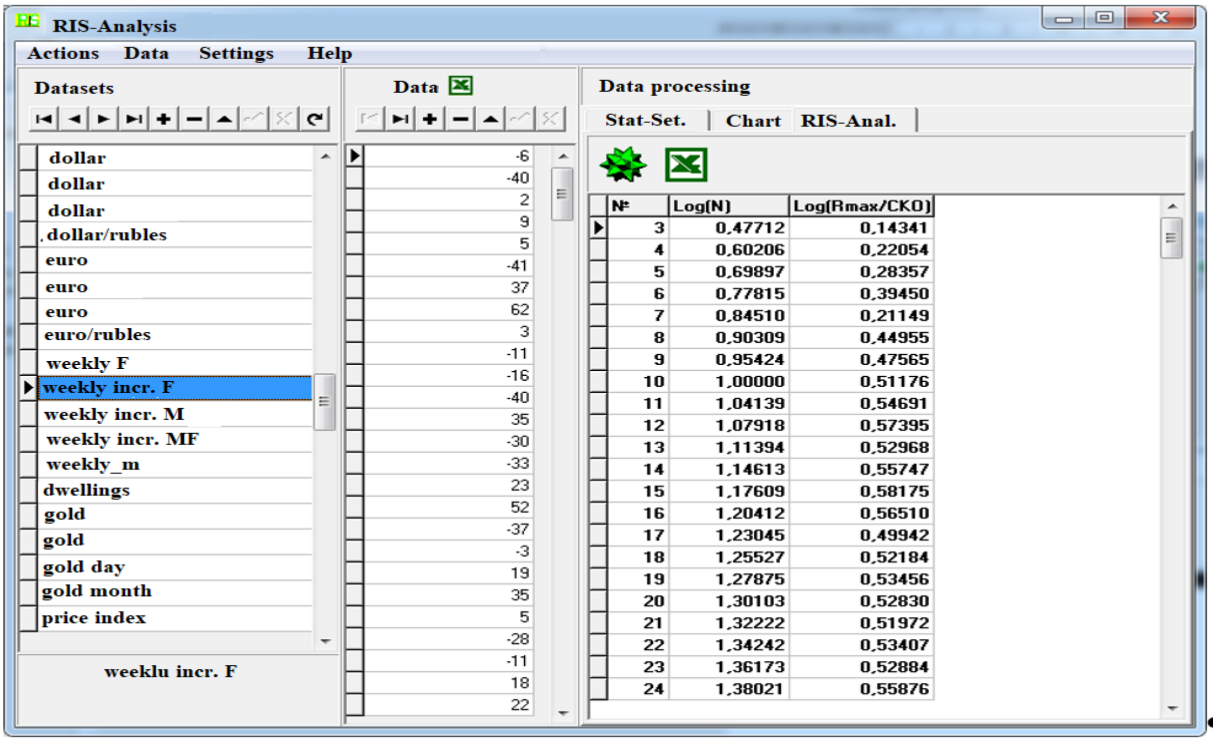
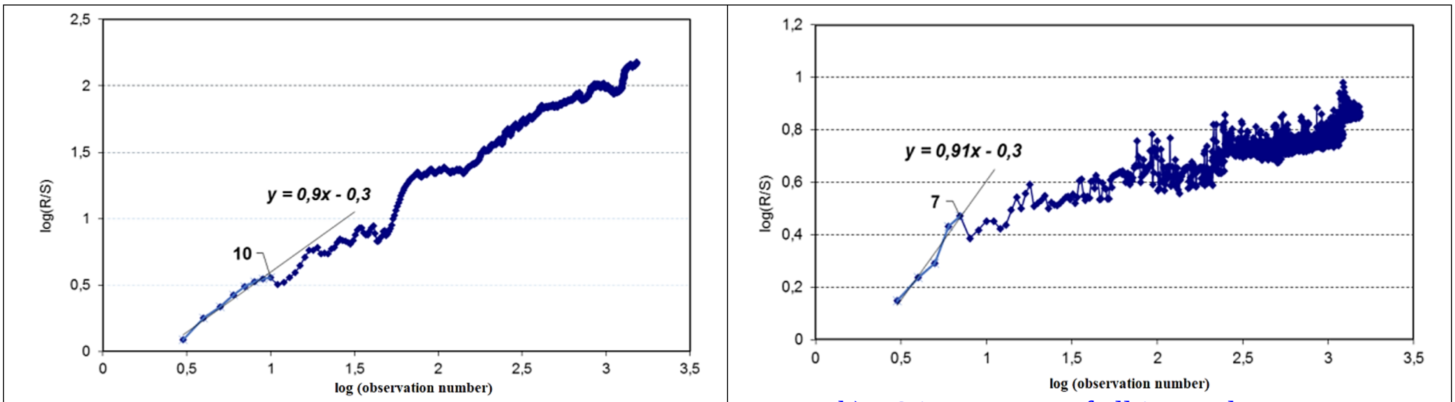


Figure 1: R/S-analysis program interface (fragment)



a) Insured women

b) TS increments of all insured persons

Figure 2: R/S-trajectories of model time series.

So, according to the two-criteria assessment of trend stability, the most trend-stable are the time series with the greatest depth of memory and the highest value of the Hurst exponent: daily data on insured men and aggregated weekly data on insured men.

The regression obtained during the construction of the R/S-trajectory trend provides an additional opportunity to obtain pre-forecast information from studying the initial process in dynamics. Table 3 shows the results of the sequential R/S-analysis algorithm [2, 3].

Table 3: Memory Depth (R/S Breakout Points).

Time series name	daily data						aggregated weekly data					
	u_i^1	u_i^2	v_i^1	v_i^2	w_i^1	w_i^2	u_i^3	u_i^4	v_i^3	v_i^4	w_i^3	w_i^4
Breakpoint of memory depth	10	7	11	5	10	7	12	6	12	4	6	6
Hurst exponent value (H)	0.78	0.3	0.81	0.28	0.75	0.25	0.77	0.35	0.82	0.35	0.74	0.35

Based on Table 3, the following conclusions can be drawn:

- for the series of increments of the investigated TS, the memory depth's value varies in the range from 4 to 7; for the initial TS in the range from 6 to 12. This means that the series of increments are less trend-stable [2, 6];

- the value of the Hurst exponent for the series of increments refers to the "pink noise" for the original TS - "black" noise [1].

The proposed and tested methods are presented in the form of a multi-criteria (two-criteria) mathematical model for assessing the trend stability of insurance time series. As the first criterion, the authors proposed an indicator reflecting the depth of memory of a time series in the form of a fuzzy set, obtained based on an R/S analysis. The second criterion is the Hurst exponent. A two-criteria approach to assessing the trend stability of time series makes it possible to differentiate them by the trend stability indicator and select working forecast models. The selected pre-forecast characteristics for each time series will allow in the future to use adequate predictive models of nonlinear dynamics methods for planning insurance companies' activities.

4 Conclusion

Using the sequential R/S analysis algorithm's mechanism, the synergistic effect obtained from the study of complex socio-economic processes in the context of the triad results is important: the initial time series, a number of its increments, and aggregated data.

The economic and mathematical model improved risk management quality in insurance activities and revealed risk factors for personal and social insurance. The provisions developed in the article are the basis for further research, development, and adaptation of economic and mathematical predictive models that are useful in planning the activities of an insurance company and, as a consequence, for developers of information and analytical systems to support management decisions [12, 13].

5 Availability of Data And Material

Data can be made available by contacting the corresponding authors.

6 Acknowledgement

The reported study was funded by RFBR according to the research project 19-010-00415 A.

7 References

[1] Peters, E. (2000). *Chaos and order in the capital markets*. M. Mir, 333.

[2] Perepelitsa, V. A., Popova, E. V., Komissarova K. A. (2007). *Modeling the activity of insurance companies by methods of nonlinear dynamics: monograph*. Scientific Ed., Krasnodar: KubGAU, 201.

[3] Yangishieva, A. M. (2003). *Alpha-fractal. Certificate of registration of the computer program RU* 2003611093, 08.05.2003. Application No. 2003610565.

[4] Yangishieva, A. M. (2005). *Modeling economic risks by methods of nonlinear dynamics*. Based on the materials of the Karachay-Cherkess Republic, Abstract of the dissertation of a candidate for economics. Sciences SSU, Stavropol, 24.

[5] Krichevsky, M. L. (2005). *Intellectual methods in management*. SPb. Peter, 305.

[6] Kumratova, A. M. Popova, E. V., Kurnosova, N. S., Popova M. I. (2015). Economic risk reduction based on pre-forecast analysis. *Modern Economy: Problem and Solution*, 3(63), 18-28.

[7] Kumratova, A. M. Popova, E. V., Tretyakov N. V. (2014). Methods of multicriteria optimization and classical statistics for the assessment of extreme risk values. *Bulletin of the Kuban State University. Natural Sciences*, 1, 55-60.

[8] Kumratova, A. M., Popova, E. V., Piterskaya, L. Y. (2019). Application of nonlinear dynamics methods for predictive testing the economic time series data. *Indo American Journal of Pharmaceutical Science*, 6(3), 5598-5602.

[9] Kumratova, A. A., Popova, E., Costa, Luis de Sousa. (2019). Hybrid approach of fractal and lin-guistic forecasting of winter wheat yields in southern Russia. *Indo American journal of pharmaceutical science*, 6(3), 5299-5303.

[10] Kumratova, A., Latysheva, L., Piterskaya, L., Shamrina, S., Kumratova, A., Popova, E. (2018). Systemic Inflationary Risk As A Factor To Investment Climate Formation. *Research Journal of Pharmaceutical, Biological & Chemical Science*, 9(6), 776-779.

[11] Kumratova, A., Bogoviz, A., Savinskaya, D., Kumratova, A. (2018). Expected Scenarios of Development of Information Economy in the Global Economic System. *Models of modern information economy: conceptual contradictions & practical examples*, 303-312.

[12] Kumratova, A., Popova. E., Costa. L., Zamotajlova, D. (2019). Methods of nonlinear dynamics as a hybrid tool for predic-tive analysis and research of risk-extreme levels. *International Journal of Hybrid Intelligent System*, 15(4), 221-241.

[13] Kumratova, A., Popova, E., Costa, L. de S. (2019). Hybrid instrumental means of predic-tive analysis of the dynamics of natural and economic processes. *Advances in Intelligent Systems & Computing*, 923, 31-39.

[14] Kumratova, A. Popova, E., Temirova, L., Shaposhnikova, O. (2019). Forecasting development of econom-ic processes using adapted nonlinear dynamics methods. *International Journal of Engineering & Advanced Technology*, 9(1), 3082-3085.

Dr.Alfira Kumratova is an Associate Professor of the Department of Information Systems of the Kuban State Agrarian University named after I.T. Trubilin. She is a Candidate of Economic Sciences.

Dr.Elena Popova is Professor and Head of the department of Information Systems of the Kuban State Agrarian University named after I.T. Trubilina. She holds a Doctor of Economics. She is a Candidate of Physical and Mathematical Sciences.

Dr.Elena Khudyakova is Professor and Head of the Department of Applied Informatics of Russian State Agrarian University - Moscow Timiryazev Agricultural Academy. She holds a Doctor of Economics. She is a Agricultural Cybernetic Economist. Her research interests are Agricultural Consumer Cooperation; Modeling of Business Processes at Agricultural Enterprises; Development and Development of Mathematical Models and Management Methods in the Agro-industrial Complex.

Dr.Igor Vasilenko is an Associate Professor of the Department of Computer Technologies and Systems, Candidate of Agricultural Sciences. He has an education in Fruit and Vegetable Growing and Viticulture. His research interests are Algorithms and Data Structures; Computing Systems, Networks and Telecommunications; Microelectronics and Circuitry.

Dr.Victor Saykinov is an Assistant of the Department of Information Systems. He has an education in Applied Informatics. His research interests are Mathematical and Instrumental Support of Workstations; Operating Systems, Environments and Shells; Cloud Computing Technologies.