TIME SERIES ANALYSIS AND THEIR DEVELOPMENT PREDICTION OF GROSS PREMIUM WRITTEN OF LIABILITY INSURANCE OF EMPLOYERS FOR WORK INJURIES AND OCCUPATIONAL DISEASES IN THE FRAME OF THE CZECH INSURANCE MARKET

Karina Mužáková

Technical University of Liberec
Faculty of Economics
Department of Insurance Management
Studentská 2, 461 17, Liberec 1, Czech Republic
karina.muzakova@tul.cz

Abstract

This paper deals with the time series analysis and their development prediction of gross premium written of liability insurance of employers for work injuries and occupational diseases in the frame of the Czech insurance market for years 2012 and 2013. The time series are defined as a sequence of data points, measured typically at successive times, spaced at time intervals. Data in this modeling are gross premium written of liability insurance of employers for work injuries and occupational diseases of members of CAP (Czech Insurance Association, within the analyzed insurance, this association covers 100% of the Czech insurance market) in years 2001 to 2011. This analysis does not include economic factors (for example: inflation, economic progress, economic recession, economic shocks).

Introduction

All employers who have at least one employee are required to be insured against liability for damage, accident at work or occupational diseases (statutory insurance). If employers arranged this insurance with the Česká pojišťovna, a.s. by 31 December 1992, they are insured with the insurance company. Other employers are insured by the insurer Kooperativa, pojišťovna, a.s.

When characterizing the Czech insurance market, several basic economic indicators appear. For example, gross premium written of life and non-life insurance, number of insurance contracts, number of employees in insurance companies and so forth. In this paper, gross premium written of liability insurance of employers for work injuries and occupational disease (in legislative conditions of the Czech Republic, it is the only statutory insurance) within years 2001 to 2011 will be analyzed and their development prediction for years 2012 and 2013 will be given. The data for this analysis used from the Czech Insurance Association (CAP) and from the Czech National Bank (CNB). The analysis is developed for the Student Project Grant Competition 2012; grant No. 3838. Time series analysis is discussed in many textbooks, see Hamilton (1994) [1]; Hindls, Hronová and Novák (2000) [2]; Chatfield (2003) [3] and Tsay (2005) [4].

In the first part of this paper, basic characteristic development of time series will be analyzed.

The second part will aim at identification of the trend by means of hypotheses tests, than an acceptable model with prediction for years 2012 and 2013 will be chosen. The estimate of trend function values will be analyzed by using the statistic program Statgraphic Centurion XVI. In the final tables *R.M.S.E.* (root mean square error), *I_adjusted*^2 (adjusted index of determination), *t-tests* (tests criterion), *P-values* (critical significance limits) and total *F-test*

will be presented. Each contribution must have a chapter called “Introduction” which is not numbered.

1 Time Series Analysis

1.1 Basic Characteristic Development of Time Series

For calculation of basic characteristic development of time series it is necessary analyze data about development of gross premium written of liability insurance of employers for work injuries and occupational disease and total gross premium written of non-life insurance and their percentage proportion (see Tab. 1 and Fig. 1). Another interesting indicator in the insurance is a loss ratio, which is the share of claims costs (total benefit) and total gross premiums written (see Tab. 2 and Fig. 2).

Tab. 1: *Development of Gross Premium Written of Insurance of Employers for Work Injuries and Occupational Disease*

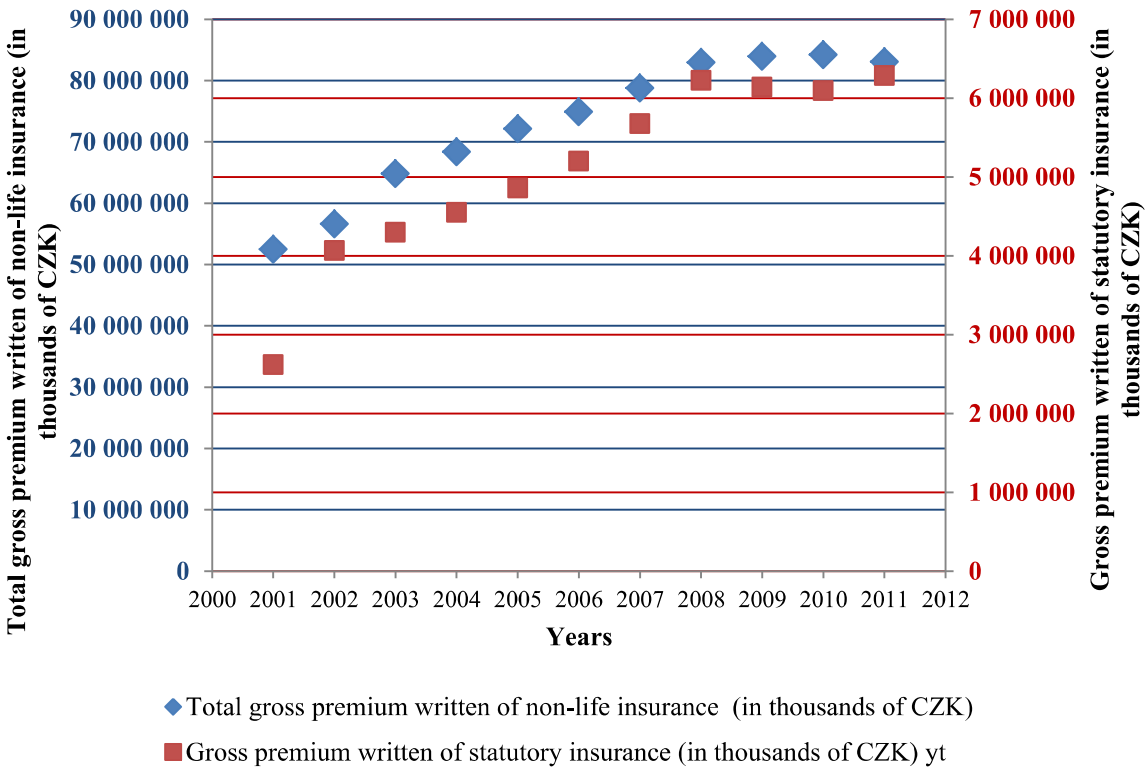
Year (t)	Total gross premium written of non-life insurance (in thousands of CZK)	Gross premium written of statutory insurance (in thousands of CZK) y_t	Percentage share of gross premium written of statutory insurance in total gross premium written (%)
2001	52 462 237	2 618 722	4.99
2002	56 624 001	4 067 310	7.18
2003	64 817 070	4 297 011	6.63
2004	68 377 194	4 550 996	6.66
2005	72 125 154	4 859 277	6.74
2006	74 889 748	5 199 571	6.94
2007	78 767 841	5 675 928	7.21
2008	82 942 866	6 222 365	7.50
2009	83 961 630	6 139 931	7.31
2010	84 231 496	6 094 095	7.23
2011	83 083 478	6 285 543	7.57

Source: Own from [5], [6]

Tab. 2: *Development of Gross Premium Written of Insurance of Employers for Work Injuries and Occupational Disease*

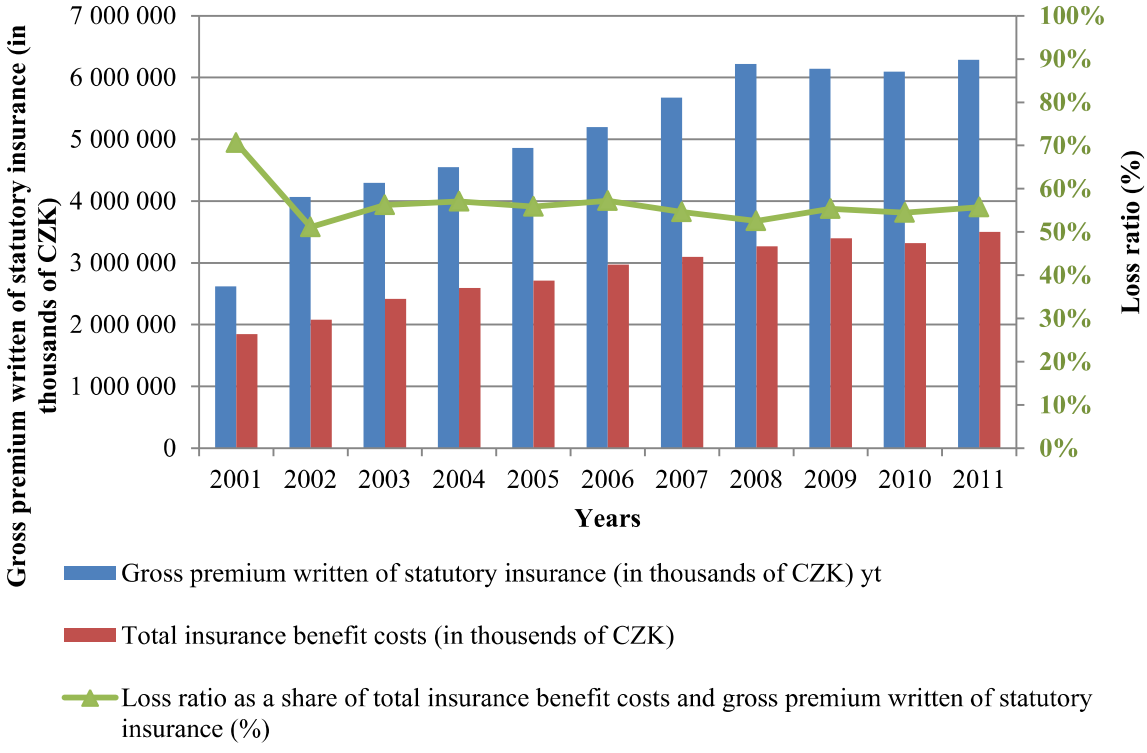
Year (t)	Gross premium written of statutory insurance (in thousands of CZK) y_t	Total insurance benefit costs (in thousands of CZK)	Loss ratio as a share of total insurance benefit costs and gross premium written of statutory insurance (%)
2001	2 618 722	1 849 074	71
2002	4 067 310	2 079 224	51
2003	4 297 011	2 415 953	56
2004	4 550 996	2 595 634	57
2005	4 859 277	2 713 900	56
2006	5 199 571	2 973 255	57
2007	5 675 928	3 098 801	55
2008	6 222 365	3 269 080	53
2009	6 139 931	3 397 763	55
2010	6 094 095	3 317 664	54
2011	6 285 543	3 502 693	56

Source: Own elaboration from [5], [6]



Source: Own elaboration from Tab. 1

Fig. 1: *Development of Gross Premium Written of Total No-life Insurance and Statutory Insurance*



Source: Own elaboration from Tab. 2

Fig. 2: *Development of Gross Premium Written of Statutory Insurance and Total Insurance Benefits Costs with Development of Loss Ratio*

The values y_t for statutory insurance are in Tab. 2. The subscript t in next equations characterizes time period (in this paper one year).

We consider other six indicators [2], [4]:

the first difference (absolute gain, ${}_1\Delta_t$), the second difference

$$ {}_2\Delta_t = {}_1\Delta_t - {}_1\Delta_{t-1} \tag{1}$$

growth coefficient

$$k_t = \frac{y_t}{y_{t-1}} \tag{2}$$

the growth rate

$$\delta_{y_t} = T_{y_t} - 100 \tag{3}$$

increase rate

$$T_{y_t} = k_t \cdot 100 \tag{4}$$

are given for statutory insurance in Tab. 3. In Fig. 3, the development of values of absolute gain is shown.

The average absolute gain (5) and the average growth coefficient (6) are of important characteristics [2], [4]

$${}_1\overline{\Delta} = \frac{\sum_{t=2}^{n} {}_1\Delta_t}{n-1} = \frac{y_n - y_1}{n-1}, \tag{5}$$

where n is the number of values (in this paper $n = 11$).

The results of average absolute gain is 366 682 100 CZK for statutory insurance.

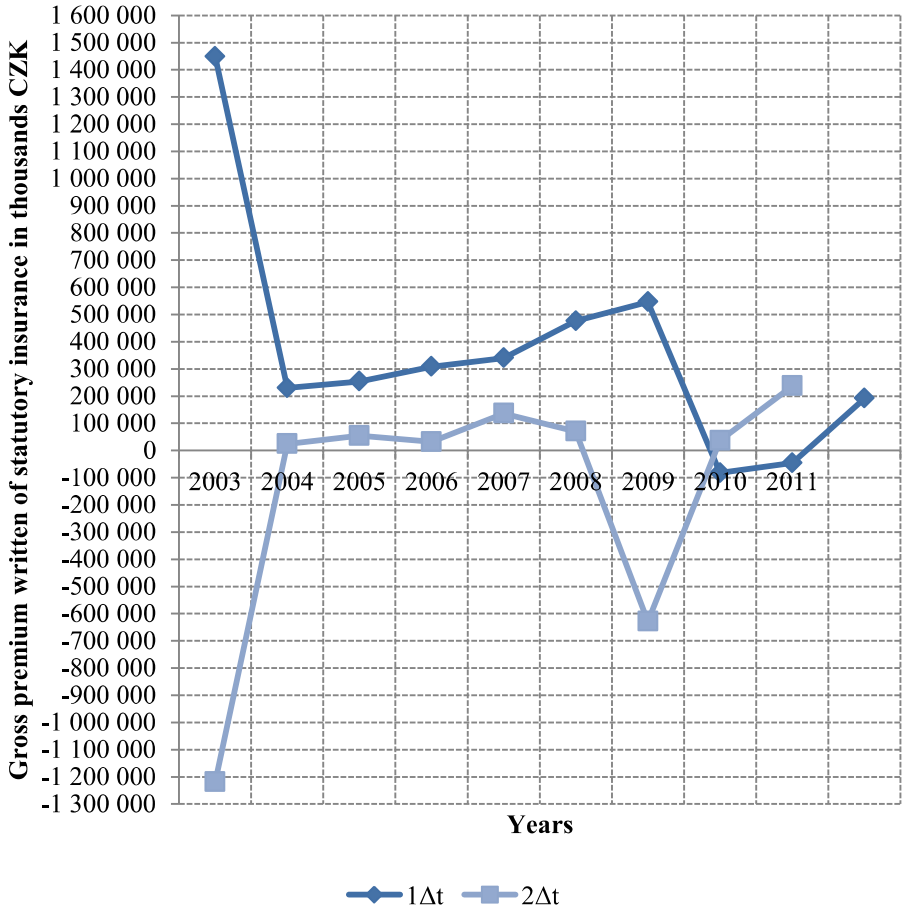
$$\overline{k} = \sqrt[n-1]{\frac{y_n}{y_1}} \tag{5}$$

The results of average growth coefficient are 1.091504 (which corresponds to 9.15%) for statutory insurance.

Tab. 3: *Elementary Characteristic Development of Gross Premium Written of Statutory Insurance*

Year (t)	Gross premium written of statutory insurance in thousands CZK (y_t)	$_1\Delta_t$	$_2\Delta_t$	k_t	T_{yt}	δ_{yt}
2001	2 618 722	×	×	×	×	×
2002	4 067 310	1 448 588	×	1.553166	155.3166	55.32
2003	4 297 011	229 701	-1 218 887	1.056475	105.6475	5.65
2004	4 550 996	253 985	24 284	1.059107	105.9107	5.91
2005	4 859 277	308 281	54 296	1.067739	106.7739	6.77
2006	5 199 571	340 294	32 013	1.070030	107.0030	7.00
2007	5 675 928	476 357	136 063	1.091615	109.1615	9.16
2008	6 222 365	546 437	70 080	1.096273	109.6273	9.63
2009	6 139 931	-82 434	-628 871	0.986752	98.67520	-1.32
2010	6 094 095	-45 836	36 598	0.992535	99.25348	-0.75
2011	6 285 543	191 448	237 284	1.031415	103.1415	3.14

Source: Own elaboration from Tab. 1



Source: Own elaboration from Tab. 3

Fig. 3: *Development of First and Second Difference of Gross Premium Written of Statutory Insurance*

1.2 Identification of the Trend

The trend identification was analyzed by the program Statgraphics Centurion.

The results of tests of individual trend functions parameters can be find in Tab. 4.

Tab. 4: *Linear, Quadratic and Exponential Trend (Gross Premium Written of Statutory Insurance)*

Trend	Linear trend	Quadratic trend	Exponential trend
Trend function	$T_t = \beta_0 + \beta_1 t$	$T_t = \beta_0 + \beta_1 t + \beta_2 t^2$	$T_t = e^{(\beta_0 + \beta_1 t)}$
Trend function forecast	$\hat{T}_t$ = 3107200000 + 330781000t	$\hat{T}_t$ = 2324830000 + 691872000t – 30090900t^2	$\hat{T}_t = e^{(21.8924+0.0717134t)}$
R.M.S.E.	375 252 000	247 599 000	490 683 000
$I^2_{adjusted}$ **(%)**	89.4149	95.3917	79.7798
H_0 :	$\beta_0 = 0$	$\beta_0 = 0$	$\beta_0 = 0$
H_1 :	$\beta_0 \neq 0$	$\beta_0 \neq 0$	$\beta_0 \neq 0$
$\hat{\beta}_0$	3 107 200 000	2 324 830 000	21.8924
Tests criterion *t-test*	12.8045	8.54986	286.287
P-value	0.000000 < 0,05	0.000027 < 0,05	0,000000 < 0,05
Test conclusion	Disapprove H_0, prove H_1.	Disapprove H_0, prove H_1.	Disapprove H_0, prove H_1.
H_0 :	$\beta_1 = 0$	$\beta_1 = 0$	$\beta_1 = 0$
H_1 :	$\beta_1 \neq 0$	$\beta_1 \neq 0$	$\beta_1 \neq 0$
$\hat{\beta}_1$	330 781 000	691 872 000	0.0717134
Tests criterion *t-test*	9.24514	6.64332	6.36046
P-value	**0,000007** < 0,05	**0,000162** < 0,05	**0,000131** < 0,05
Test conclusion	Disapprove H_0, prove H_1.	Disapprove H_0, prove H_1.	Disapprove H_0, prove H_1.
H_0 :		$\beta_2 = 0$	
H_1 :		$\beta_2 \neq 0$	
$\hat{\beta}_2$		30 090 900	
Tests criterion *t-test*		-3.55984	
P-value		0.007402 < 0,05	
Test conclusion		Disapprove H_0, prove H_1.	
H_0 :	The linear trend is not acceptable model.	The quadratic trend is not acceptable model.	The exponential trend is not acceptable model.
H_1 :	Non H_0	Non H_0	Non H_0
F-test	85.47	104.5	40.46
P-value	0.0000 < 0.05	0.0000 < 0.05	0.0001 < 0.05
Test conclusion	Disapprove H_0, prove H_1.	Disapprove H_0, prove H_1.	Disapprove H_0, prove H_1.

Source: Own elaboration

The roots mean squared error "*R.M.S.E.*" (7) is calculated as:

$$R.M.S.E. = \sqrt{\frac{1}{n}\sum\left(y_i - T_i\right)^2} \tag{7}$$

The adjusted index of determination (8) is defined by:

$$I^2_{adjusted} = 1 - \frac{(n-1)\left[\sum\left(y_i - Y_i\right)^2\right]}{(n-p)\left[\sum\left(Y_i - \bar{y}\right)^2 + \sum\left(y_i - Y_i\right)^2\right]} = 1 - \frac{(n-1)S_R}{(n-p)\left(S_T + S_R\right)} = 1 - \frac{(n-1)S_R}{(n-p)S_y} \tag{8}$$

where S_T is theoretical sum of squares and S_R is residual sum of squares. Test criterion by the prove of hypothesis H_0 has distribution F by $(p-1)$ and $(n-p)$ degrees of freedom.

After estimating the parameters of individual trends it is necessary, by interpolation criteria and testing hypotheses, to choose the appropriate model. The character of actual trend values y_t is also examined, these differences are called residues. Accuracy of the estimate is measured by the average residual characteristics [4] of *M.A.E., M.A.P.E., M.E., M.P.E.*

The mean absolute error "*M.A.E.*" (9) is calculated as:

$$M.A.E. = \frac{\sum\left|y_t - \hat{T}_t\right|}{n} \tag{9}$$

The mean absolute percentage error "*M.A.P.E.*" (10) is calculated as:

$$M.A.P.E. = \sum\left(\frac{\left|y_t - \hat{T}_t\right|}{y_t}\right) \cdot \frac{100}{n} \tag{10}$$

The mean error "*M.E.*" (11) is calculated as:

$$M.E. = \frac{\sum(y_t - \hat{T}_t)}{n} \tag{11}$$

The mean percentage error "*M.P.E.*" (12) is calculated as:

$$M.P.E. = \sum\left(\frac{y_t - \hat{T}_t}{y_t}\right) \cdot \frac{100}{n} \tag{12}$$

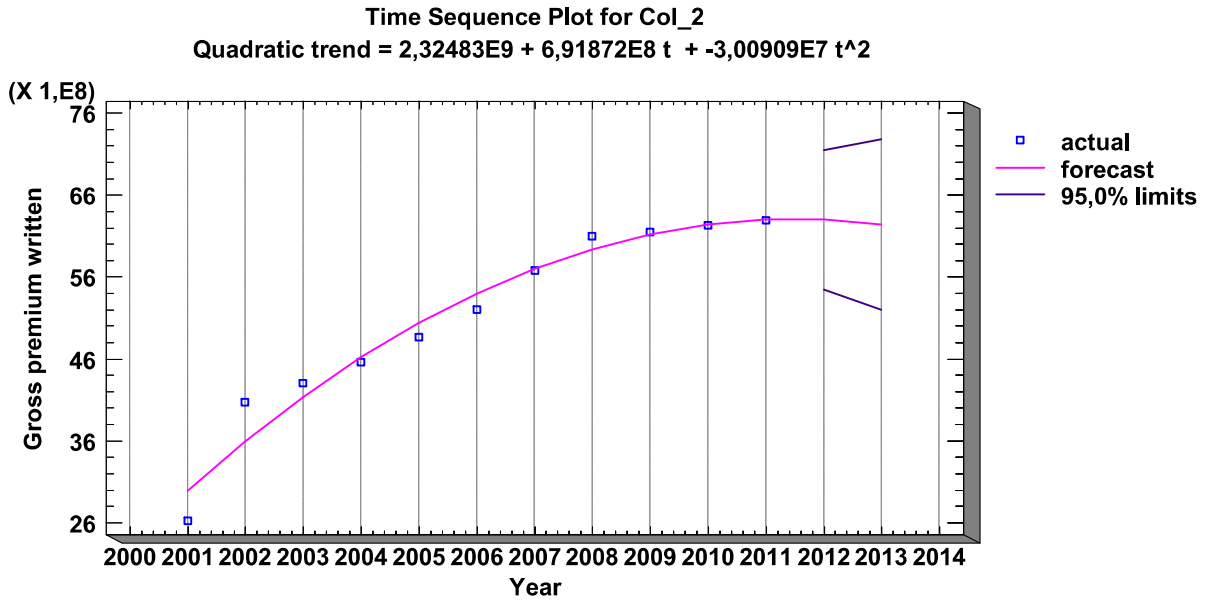
The results of average characteristics of the residues of *M.A.E.*, *M.A.P.E.*, *M.E.*, and *M.P.E.* are presented in Tab. 5.

Tab. 5: *Average Characteristics of the Residues of M.A.E., M.A.P.E., M.E., M.P.E.*

Residues	Linear trend	Quadratic trend	Exponential trend
M.A.E.	267 664 000	**151 306 000**	379 312 000
M.A.P.E.	6.48551	**3.81792**	8.6224
M.E.	-0.000000520186	-0.000000520186	**0.000000136702**
M.P.E.	-1.09686	**-0.414535**	-0.606808

Source: Own elaboration

The quadratic trend is available according to the results of *RMSE*, *M.A.E.*, *M.A.P.E.*, *M.E.*, *M.P.E.*, adjusted index of determination, *t-tests*, *P-values* and total *F-test* [2], [3], [4]. For the forecast of this model, see Fig. 4.



Source: Own elaboration

Fig. 4: *Time Series Equalization by Quadratic Trend and Forecast of Development by Next Two Years*

According to the results of statistic program Statgraphics Centurion, the predicted values of gross premium written of statutory insurance with 95% confidence level will be in the interval of 5 446 170 000 CZK – 7 142 250 000 CZK (with point prediction of 6 294 210 000 CZK) in the year 2012 and in the interval of 6 233 810 000 CZK – 7 272 450 000 CZK (with point prediction of 6 233 810 000 CZK) in the year 2013.

Conclusion

The development of gross premium written of statutory insurance can be modelled with quadratic trend. The development point prediction with 95% confidence level for next two years is 6 294 210 000 CZK for year 2012 and 6 233 810 000 CZK for year 2013.

This analysis doesn't include economic factors (for example: inflation, economic progress, economic recession, economic shocks).

Acknowledgements

This paper was written with the support of the Student Project Grant Competition 2012; grant No. 3838.

Literature

[1] HAMILTON, J. D.: *Time Series Analysis.* Edition 1. Princeton University Press, Princeton, 1994. ISBN 06-91042-89-6.

[2] HINDLS, R.; HRONOVÁ, S.; NOVÁK, I.: *Metody statistické analýzy pro ekonomy.* Edition 2. Management Press, Praha, 2000. ISBN 80-7261-013-9.

[3] CHATFIELD, CH.: *The analysis of time series: an introduction.* Edition 6. CRC Press, London, 2003. ISBN 15-848831-7-0.

[4] TSAY, S. R.: *Analysis of financial time series.* Edition 2. John Wiley and Sons, Chicago, 2005. ISBN 04-71690-74-0.

[5] ČAP. *Výroční zprávy 2001–2011.* [online]. ČAP, Praha, 2001–2011. [accessed 2012-09-07]. Available from WWW: <http://www.cap.cz>.

[6] ČNB. *Pojistné ukazatele 2001–2011.* [online]. ČNB, Praha, 2001–2011. [accessed 2012-10-15]. Available from WWW: <http://www.cnb.cz>.

ANALÝZA ČASOVÝCH ŘAD A PREDIKCE VÝVOJE PŘEDEPSANÉHO HRUBÉHO POJISTNÉHO POJIŠTĚNÍ ODPOVĚDNOSTI ZAMĚSTNAVATELE ZA PRACOVNÍ ÚRAZY A NEMOCI Z POVOLÁNÍ V RÁMCI ČESKÉHO POJISTNÉHO TRHU

Tato práce se zabývá analýzou časových řad a predikcí vývoje hrubého předepsaného pojistného u pojištění odpovědnosti zaměstnavatele za pracovní úrazy a nemoci z povolání v rámci českého pojistného trhu pro roky 2012 a 2013. Časové řady jsou definovány jako sekvence datových bodů, měřených obvykle po sobě jdoucím čase, které jsou umístěny v časových intervalech. Daty v této modelaci jsou hrubé předepsané pojistné u pojištění odpovědnosti zaměstnavatele za pracovní úrazy a nemoci z povolání členů České asociace pojišťoven a také data České národní banky za období let 2001–2011. Tato analýza nezahrnuje ekonomické faktory (například: inflaci, hospodářský pokrok, hospodářskou recesi, ekonomické šoky).

ZEITREIHENANALYSE UND PROGNOSE DER BRUTTOPRÄMIEN DER UNFALLVERSICHERUNG FÜR ARBEITSUNFÄLLE UND BERUFSKRANKHEITEN AUF DEM TSCHECHISCHEN VERSICHERUNGSMARKT

Diese Arbeit beschäftigt sich mit der Zeitreihenanalyse und der Prognose von Bruttoprämien Unfallversicherung für Arbeitsunfälle und Berufskrankheiten auf dem tschechischen Versicherungsmarkt für die Jahre 2012 und 2013. Zeitreihen werden als Folge von üblicherweise in zeitlicher Abfolge gemessenen Datenpunkten definiert, die in zeitlichen Intervallen angeordnet sind. Die Daten in dieser Modellierung bestehen in den Bruttoprämien der Unfallversicherung für Arbeitsunfälle und Berufskrankheiten der Mitglieder des Tschechischen Versicherungsverbands und auch die Daten der Tschechischen Nationalbank für den Zeitraum 2001–2011. Diese Analyse schließt wirtschaftliche Faktoren wie Inflation, wirtschaftlichen Fortschritt, wirtschaftliche Rezession ökonomische Schocks usw. nicht mit ein.

ANALIZA SZEREGÓW CZASOWYCH I PROGNOZOWANIE SKŁADKI PRZYPISANEJ BRUTTO ODPOWIEDZIALNOŚCI PRACODAWCÓW UBEZPIECZEŃ Z TYTUŁU WYPADKÓW PRZY PRACY I CHORÓB ZAWODOWYCH W CZESKIM RYNKU UBEZPIECZEŃ

Niniejsze opracowanie dotyczy analizy szeregów czasowych i prognozowania składek przypisanych brutto w odpowiedzialności pracodawców ubezpieczeń z tytułu wypadków przy pracy i chorób zawodowych na czeskim rynku ubezpieczeń w latach 2012 i 2013. Szeregi czasowe określone są jako ciąg punktów danych, mierzonych zazwyczaj kolejno w pewnych odstępach czasu. Dane w omawianym modelowaniu stanowią składki przypisane brutto odpowiedzialności pracodawców ubezpieczeń z tytułu wypadków przy pracy i chorób zawodowych Czeskie Zrzeszenie Ubezpieczycieli i dane z Czeskiego Banku Narodowego w okresie 2001–2011. Analiza ta nie obejmuje czynników ekonomicznych (takich jak: inflacja, postęp gospodarczy, recesja gospodarcza, wstrząsy gospodarcze).