Tomasz RYMARCZYK[1,2], Bartosz PRZYSUCHA[3], Paweł PAWLIK[4]

Research & Development Centre Netrix S.A.(1), University of Economics and Innovation (2)
Lublin University of Technology (3)
AGH University of Science and Technology (4)

doi:10.15199/48.2020.09.20

Analysis of multi-source data for monitoring and control of intelligent technological systems

***Abstract**. The article presents a part of cyber-physical system for acquiring, processing and controlling from measurement data. The technology was based on, intelligent measurement sensors, internet of Things as a solution for Industry 4.0. The aspect raised in the article concerns data reduction and selection of an appropriate covariant in the modeling optimization of modeling faults by the Cox model for a specific mechanical system.*

***Streszczenie.** Artykuł przedstawia część cyber-fizycznego systemu do zbierania, przetwarzania i sterowania przy pomocy informacji pochodzącej z danych pomiarowych. Technologia ta została oparta na inteligentnych czujnikach pomiarowych z użyciem internetu rzeczy jako rozwiązania dla Przemysłu 4.0. Aspekt poruszony w pracy dotyczy redukcji danych i wyboru odpowiedniego kowariantu w optymalizacji modelowania usterek modelem Coxa dla konkretnego układu mechanicznego (**Analiza danych wieloźródłowych do monitorowania i kontroli inteligentnych systemów technologicznych**).*

Słowa kluczowe: Model Coxa, predykcja uszkodzeń, sterowanie produkcją, inteligentna platforma.
Keywords: Cox model, time to failure prediction, production control, intelligent platform

Introduction

The article presents the results of research on the use of sensors for the analysis of industrial processes using dedicated measuring devices [1]. Advanced automation and control of production processes play a key role in enterprises. Technological equipment and production lines can be considered the heart of industrial production, while information technologies and control systems are its brains. They provide high flexibility, quick adaptation of production processes to changing market requirements, as well as safety and efficiency at optimal resource and energy costs.

The presented concept of Cyber Intelligent Enterprise consists directly in the application of vision in the field of corporate systems and the Internet of Things (IoT). Tight integration of physical devices and business processes gives new opportunities and increases enterprise efficiency [2]. Cyberphysical production systems (CPPS) are based on the latest achievements in the field of information technology, electronics, information and communication technologies. There are many methods for solving optimization problems [3-22]. Autonomy, cooperation, optimization, integration of analytical approaches and simulations is related to the operation of sensor networks, large amounts of data as well as searching, analysis and interpretation of information, with particular emphasis on security aspects [23-25].

Modeling survival times

The reliability of equipment and components in the production process is a necessary condition for ensuring continuity of production. Predicting damage time gives you the ability to monitor the current degradation status of the device and improve the control planning process, preventive replacement of parts, warranty service or storage of spare parts in order to maintain production continuity and optimal production efficiency. One of the most popular models for determining the expected failure time is the Cox model. This model is a regressive survival analysis model. It allows you to calculate the probability of failure occurrence in a given time or average time of failure occurrence, etc. [26-29].

The Cox model, proportional hazard is defined by the following terms. The survival function is called the function given by the formula:

$$S(x) = P(T \geq x) = 1 - F(x^-),$$

$F -$ lifetime distributor, $f -$ lifetime density. It determines the probability that the object will live longer than the time x.

Description of the test stand

The testing fig consisted of a drive engine connected to the shaft by a claw clutch. The clutch in this case had 4 claws. The shaft was placed in three bearing nodes using ball bearings. The drive system was misaligned by moving the middle bearing 5mm from the shaft axis. The testing rig scheme with measuring point was presented in the Fig. 1.

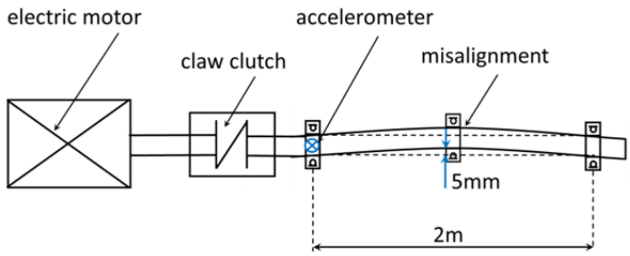

Fig. 1. Testing rig scheme with measuring point

Acceleration of vibrations was recorded on the first bearing node closest to the clutch. The piezoelectric sensor (Type PCB 356B08) mounted on the bearing housing in a vertical direction transverse to the shaft axis was used (Fig. 1). The measurement was carried out with a sampling frequency of 50kHz using a measuring card (Type NI 9233) connected to a PC. Damage consisting of overheating and displacement of the claw clutch due to misalignment of the system was recorded. The experiment was repeated 13 times. Vibration acceleration signal waveforms were recorded until the fault occurred. From signel were calculated parameters with a time interval of 0.1s: kurtosis and root mean square (RMS), which were used as covariates in the model PHM. The duration of signals from the beginning of measurements to the time of occurrence of the fault is given in Table 1.

Table 1. Duration of measurements until the fault occurs

Nr	1	2	3	4	5
$[t_0, t_1]$	[20;133]	[20;159]	[20;165]	[80,111]	[96;127]
Nr	6	7	8	9	10
$[t_0, t_1]$	[75;106]	[20,141]	[25,164]	[25,165]	[82,113]
Nr	11	12	13	-	-

Results and data analysis

As a result, models were obtained for which the p-value values were approaching the required level of 0.05 but did not reach it. Models with one covariant were also tested and it was obtained that for the model with RMS covariant in the form of maximum values from 10-second intervals, the model turned out to be statistically significant. The model parameters are presented in table 3.

Table 3. Statistics in the Cox model estimation process for one maximum from 10 sec RMS Covariant model.

	Value	St. error	p-value
β_{RMS}	51	18.4	0.017

After adding another fault to the model, the value of the p-value parameter for the β_RMS coefficient decreased to p = 0.0055.

Based on Cox modeling, the RMS time course value was selected as a significant covariant. Then, based on the time signal analysis, 500 RMS covariate time series were generated. Each RMS sequence RMS $\{x_t\}_{t\geq 1}$ was generated as follow:

Three values from random variables with a Poisson distribution were generated. These values are responsible for changing the shape of the time series

$a_1 -$ for $\lambda = 1500$,

$a_2 -$ for $\lambda = 1200$,

$a_3 -$ for $\lambda = 300$,

and α with normal distribution $N(0{,}0.005)$.

Sequence $\{x_t\}_{t\geq 1}$ is defined:

$x_t = \xi_{1t}$, for $1 \leq t \leq a_1$, $\xi_{1t} \sim N(0{,}0.25)$

$x_t = |\alpha|(t - a_1) + \xi_{2t}$, for $a_1 \leq t \leq a_1 + a_2$, $\xi_{2t} \sim N(0{,}0.25)$

$x_t = \sum_{j=a_1+a_2}^{t} |\alpha|(j - (a_1 + a_2)) + \xi_{3t}$, for $t > a_1 + a_2$, $\xi_{3t} \sim N(0{,}0.0025)$

The first part of the time series $\{x_t\}_{t\geq 1}$ until a_1 contains the values of random variables from a variable with a normal distribution. From $a_1 + 1$ to $a_1 + a_2$, values are generated as the sum of the trend plus a disruption with normal distribution, from $a_1 + a_2 + 1$, the elements are generated as an integration of the one-degree deterministic trend plus a disruption with a normal distribution.

Based on the generated RMS signals, the impact of the number of defects on the statistical significance of the Cox model with the RMS covariant was analyzed.

The time of fault's occurrence from the moment of measurements was from a dozen to several dozen seconds. RMS coefficient was determined every 0.1 sec. From the 500-element measurement sample of the time series of the RMS covariant, $n = 5$, $n = 7$, $n = 8$, $n = 10$, $n = 12$, $n = 15$ elemental samples were drawn 10,000 times. The model was constructed from data containing values of the RMS coefficient in a given time period $d =$ 1sec, 2sec, 3sec, 5sec, 10sec, 12,5sec, 15sec, 17,5sec, 20sec, 22,5sec, 25sec.

The following covariant were tested:

- maximum in the time period
- mean in the time period
- median in the time period
- minimum in the time period

From the covariants calculated in the above manner, Cox models of the time of fault occurrence were created. The draw was repeated 10,000 times and the percentage of models (Monte Carlo method) was calculated for which the parameter at the covariance turned out to be statistically significant.

Maximum in time period:

Table 4. Values of parameters obtained as a result of estimation for different lengths of the time interval $d =$ 1sec, $d =$2sec, …, $d =$25sec, and sample size n. For the covariate being the maximum value of the RMS parameter in the time interval d and sample size n.

d	symbol	n: 7	8	10	12	15
1	p	2,9%	15,8%	55,1%	84,4%	97,6%
	$\sum S_e$	28722,04	17605,76	5439,14	971,36	164,43
	b_m	842,44	674,18	389,86	215,78	139,57
5	p	5,8%	26,2%	73,8%	95,0%	99,8%
	$\sum S_e$	12617,12	5646,65	707,09	76,57	30,93
	b_m	399,94	265,50	134,00	103,12	93,06
10	p	6,4%	29,0%	78,9%	96,8%	99,9%
	$\sum S_e$	2889,37	503,49	48,40	30,28	23,71
	b_m	142,34	98,14	78,95	74,60	70,81
15	p	5,0%	23,3%	73,5%	94,9%	99,7%
	$\sum S_e$	356,99	44,47	31,48	25,72	20,88
	b_m	81,07	69,93	64,49	61,66	59,54
20	p	2,7%	14,8%	57,1%	87,1	98,5
	$\sum S_e$	66,28	37,04	28,45	23,61	19,59
	b_m	62,57	58,12	55,06	53,01	51,70
25	p	1,6%	10,4%	47,9%	79,9%	96,2%
	$\sum S_e$	55,12	35,85	27,39	23,16	19,14
	b_m	55,93	53,95	51,06	49,57	48,16

$p -$ percentage number of models significant for the generated n-element samples, $\sum S_e -$ sum of parameter estimation errors at the covariant, $b_m -$ average value of the parameter at the covariant.

Mean in the time period:

Table 5. Values of parameters obtained as a result of estimation for different lengths of the time interval $d =$ 1sec, $d =$2sec,, $d =$25sec, and sample size n. For the covariate being the mean value of the RMS parameter in the time interval d and sample size n.

d	symbol	n: 7	8	10	12	15
1	p	3,2%	17,4%	56,6%	83,4%	97,5%
	$\sum S_e$	28395,56	16878,22	4763,12	1035,78	109,99
	b_m	826,34	640,32	359,71	211,39	139,93
5	p	5,5%	27,1%	75,4%	95,1%	99,9%
	$\sum S_e$	12257,01	5305,36	694,74	45,58	29,79
	b_m	380,51	261,51	132,74	101,45	90,70
10	p	8,3%	34,4%	85,3%	98,0%	99,9%
	$\sum S_e$	2362,94	670,85	34,41	27,03	21,36
	b_m	130,69	97,83	73,48	69,43	65,84
15	p	8,5%	32,0%	80,0%	95,1%	99,3%
	$\sum S_e$	230,33	75,16	26,43	22,13	18,10
	b_m	71,37	63,24	57,01	55,11	52,87
20	p	5,3%	21,3%	58,3%	78,8%	92,0%
	$\sum S_e$	36,87	28,99	22,26	18,89	15,74
	b_m	51,22	47,58	44,00	42,57	40,84
25	p	2,9%	12,6%	38,2%	57,1%	76,4%
	$\sum S_e$	32,60	27,04	21,47	18,25	15,43
	b_m	43,86	40,97	38,02	36,39	35,33

$p -$percentage number of models significant for the generated n-element samples, $\sum S_e -$ sum of parameter estimation errors at the covariant, $b_m -$ average value of the parameter at the covariant.

Median in the time period:

Table 6. Values of parameters obtained as a result of estimation for different lengths of the time interval $d = 1$sec, d =2sec,, d =25sec, and sample size n. For the covariate being the median value of the RMS parameter in the time interval d and sample size n.

d	symbol	n = 7	8	10	12	15
1	p	2,7%	16,4%	55,9%	84,2%	97,6%
	$\sum S_e$	28173,35	16953,89	4948,76	831,37	120,51
	b_m	822,37	660,49	359,16	205,21	141,23
5	p	5,8%	27,7%	76,6%	96,0%	99,8%
	$\sum S_e$	12406,37	5204,29	516,51	60,45	29,18
	b_m	384,31	251,94	126,67	99,48	89,20
10	p	8,4%	34,5%	84,9%	97,9%	99,9%
	$\sum S_e$	2373,74	586,78	34,07	26,70	21,18
	b_m	129,78	93,60	72,9	68,42	65,12
15	p	8,7%	32,3%	78,9%	94,7%	99,2%
	$\sum S_e$	238,32	35,83	26,25	21,61	17,84
	b_m	71,59	61,43	56,31	53,88	52,04
20	p	5,3%	20,3%	57,1%	77,3%	91,32
	$\sum S_e$	44,44	27,88	21,70	18,39	15,41
	b_m	49,74	45,60	42,62	41,11	39,64
25	p	2,7%	10,6%	33,1%	50,8%	70,2%
	$\sum S_e$	31,09	25,89	20,53	17,61	14,92
	b_m	40,62416	37,83	35,24	33,95	32,78

p –percentage number of models significant for the generated n-element samples, $\sum S_e$ – sum of parameter estimation errors at the covariant, b_m – average value of the parameter at the covariant.

Minimum in the time period

Tabela 7. Values of parameters obtained as a result of estimation for different lengths of the time interval $d = 1$sec, d =2sec,, d =25sec, and sample size n. For the covariate being the minimum value of the RMS parameter in the time interval d and sample size n.

d		n = 7	8	10	12	15
1	p	2,6%	16,2%	55,4%	83,1%	97,8%
	$\sum S_e$	27378,31	16301,92	4226,90	862,48	81,63
	b_m	820,58	642,22	339,53	206,04	138,31
5	p	7,1%	31,8%	82,1%	97,1%	99,8%
	$\sum S_e$	9,899e+03	3,630e+03	522,38	39,15	26,16
	b_m	326,96	214,15	115,07	88,61	81,58
10	p	11,7%	40,7%	85,7%	96,7%	99,5%
	$\sum S_e$	1,828e+03	392,10	26,71	20,86	17,01
	b_m	109,51	74,43	58,76	54,78	52,67
15	p	8,3%	26,6%	59,2%	76,9%	90,1%
	$\sum S_e$	153,95	25,65	19,49	16,46	13,71
	b_m	49,31	42,77	39,35	37,55	36,08
20	p	2,5%	7,8%	20,0%	30,5%	45,3%
	$\sum S_e$	23,87	20,24	16,07	13,80	11,70
	b_m	27,38	25,39	23,56	22,27	21,73
25	p	0,4%	1,4%	3,7%	5,9%	9,3%
	$\sum S_e$	22,7	19,32	15,68	13,58	11,54
	b_m	17,89	16,26	15,09	14,48	13,74

p –percentage number of models significant for the generated n-element samples, $\sum S_e$ – sum of parameter estimation errors at the covariant, b_m – average value of the parameter at the covariant.

Tables 4-7 present the results of the experiment. The average value of the coefficient for the covariance b_m, the sum of standard errors of all models $\sum S_e$, and the percentage of models with a significant parameter at the covariate, p depending on the covariate are presented.

The relationships obtained in Tables 4-7 are shown in Fig 2-5. Graphs of the percentage of significant models for covariants, respectively, median, average value, maximum and minimum RMS in the time period $d = 1$ sek, 2 sec, 3 sec, 5 sec, 6 sec, 8 sec, 10 sec, $12{,}5$ sec, 15 sec, $17{,}5$ sec, 20 sec, $22{,}5$ sec, 25 sec. On each graph a different curve showing the percentage of significant models for n-element samples in time interval of length d in seconds (x axis).

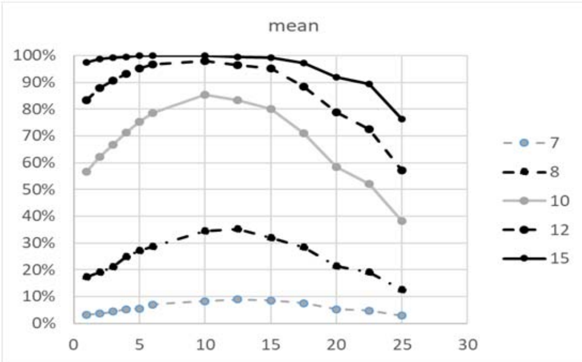

Fig 2. Graph of the percentage of significant models depending on the averaging time d for various numbers of measurement samples n for the covariate being the average value of the RMS value over the time interval d.

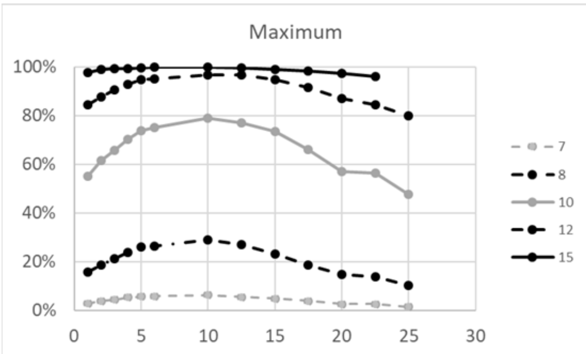

Fig 3. Graph of the percentage of significant models depending on the averaging time d for various numbers of measurement samples n for the covariate being the maximum value of the RMS value over the time interval d.

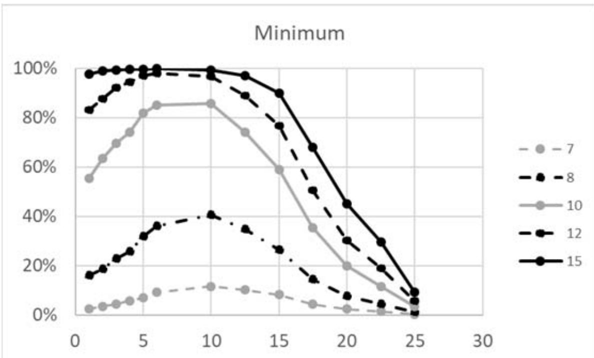

Fig 4. Graph of the percentage of significant models depending on the averaging time d for various numbers of measurement samples n for the covariate being the minimum value of the RMS value over the time interval d.

Conclusion

The article deals with the problem of selection of covariants in the Cox model and data reduction in order to optimize the quality of obtained models. The analysis was carried out on data from an experiment consisting in the introduction of a fault in the mechanical system of shaft alignment of the clutch. Based on the received data, statistically significant covariates were determined, and 500-damage simulations were performed. Based on the generated RMS time series data, the impact of the time window length of data reduction and the type of covariant on the quality of the model was tested. For each covariate, the maximum percentage of significant models can be observed depending on the length of the time window. Optimal values are obtained for a time window between 10 and 12 second.

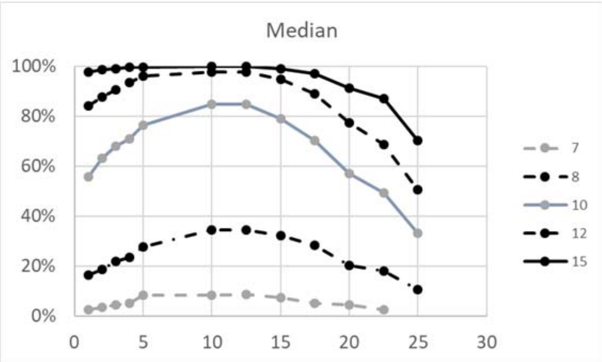

Fig 5. Graph of the percentage of significant models depending on the averaging time d for various numbers of measurement samples n for the covariate being the median value of the RMS value over the time interval d.

***Authors**: Tomasz Rymarczyk, Ph.D. Eng., University of Economics and Innovation, Projektowa 4, Lublin, Poland/ Research & Development Centre Netrix S.A. E-mail: tomasz@rymarczyk.com; , dr Bartosz Przysucha, Faculty of management, Lublin University of Technology, 20-618 Lublin 38A Nadbystrzycka str., e-mail: b.przysucha@pollub.pl, dr inż. Paweł Pawlik, AGH University of Science and Technology, Department of Mechanics and Vibroacoustics, 30-059 Kraków, al. A. Mickiewicza 30, Poland, e-mail: pawlik@agh.edu.pl*

REFERENCES

[1] Rymarczyk, T., Przysucha, B., Intelligent sensor platform for multi-source data analysis for monitoring and control of technological systems, 2019 Applications of Electromagnetics in Modern Engineering and Medicine, PTZE 2019, 171-175
[2] Repta D., Sacala I., Moisescu M., Stanescu A. Towards the development of a Cyber-Intelligent Enterprise System Architecture, 19th World Congress The International Federation of Automatic Control, Cape Town, South Africa. August 24-29, 2014.
[3] Dušek J., Hladký D., Mikulka J., Electrical Impedance Tomography Methods and Algorithms Processed with a GPU, In PIERS Proceedings, 2017, 1710-1714.
[4] Goetzke-Pala A., Hoła A., Sadowski Ł., A non-destructive method of the evaluation of the moisture in saline brick walls using artificial neural networks. Archives of Civil and Mechanical Engineering, 18 (2018), No 4, 1729-1742.
[5] Grudzien K., Romanowski A., Chaniecki Z., Niedostatkiewicz M., Sankowski D., Description of the silo flow and bulk solid pulsation detection using ECT, Flow Measurement and Instrumentation, 21 (2010), No. 3, 198-206.
[6] Kozlowski E., Mazurkiewicz D., Kowalska B., et al., Binary Linear Programming as a Decision-Making Aid for Water Intake Operators, 1st International Conference on Intelligent Systems in Production Engineering and Maintenance (ISPEM), Wroclaw, Poland, 2017.
[7] Galazka-Czarnecka, I.; Korzeniewska E., Czarnecki A. et al., Evaluation of Quality of Eggs from Hens Kept in Caged and Free-Range Systems Using Traditional Methods and Ultra-Weak Luminescence, Applied sciences-basel, 9 (2019), No. 12, 2430.
[8] Korzeniewska E., Walczak M., Rymaszewski J., Elements of Elastic Electronics Created on Textile Substrate, Proceedings of the 24th International Conference Mixed Design of Integrated Circuits and Systems - MIXDES 2017, 2017, 447-45.
[9] Kryszyn J., Smolik W., Toolbox for 3d modelling and image reconstruction in electrical capacitance tomography, Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska (IAPGOŚ) , 7 (2017), No. 1, 137-145.
[10] Mosorov V., Grudzień K., Sankowski D., Flow velocity measurement methods using electrical capacitance tomography, Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska (IAPGOŚ), 7 (2017), No.1 ,30-36
[11] Kowalska A., Banasiak R., Romanowski A., Sankowski D., Article 3D-Printed Multilayer Sensor Structure for Electrical Capacitance Tomography, 19 (2019), Sensors, 3416
[12] Rymarczyk T, Kłosowski G. Innovative methods of neural reconstruction for tomographic images in maintenance of tank industrial reactors. Eksploatacja i Niezawodnosc – Maintenance and Reliability, 21 (2019); No. 2, 261–267
[13] Rymarczyk, T.; Kozłowski, E.; Kłosowski, G.; Niderla, K. Logistic Regression for Machine Learning in Process Tomography, Sensors, 19 (2019), 3400.
[14] Rymarczyk T., Characterization of the shape of unknown objects by inverse numerical methods, Przegląd Elektrotechniczny, 88 (2012), No 7b, 138-140
[15] Rymarczyk T., Adamkiewicz P., Polakowski K., Sikora J., Effective ultrasound and radio tomography imaging algorithm for two-dimensional problems, Przegląd Elektrotechniczny, 94 (2018), No 6, 62-69
[16] Rymarczyk T., Szumowski K., Adamkiewicz P., Tchórzewski P., Sikora J., Moisture Wall Inspection Using Electrical Tomography Measurements, Przegląd Elektrotechniczny, 94 (2018), No 94, 97-100
[17] Vališ, D., & Mazurkiewicz, D. (2018). Application of selected Levy processes for degradation modelling of long range mine belt using real-time data. Archives of Civil and Mechanical Engineering, 18 (2018), No. 4, 1430-1440.
[18] Wajman R., Fiderek P., Fidos H., Sankowski D., Banasiak R., Metrological evaluation of a 3D electrical capacitance tomography measurement system for two-phase flow fraction determination, Measurement Science and Technology, 24 (2013), No. 6, 065302.
[19] Chen B., Abascal J., Soleimani M., Electrical Resistance Tomography for Visualization of Moving Objects Using a Spatiotemporal Total Variation Regularization Algorithm, 18 (2018), Sensors 2018, 1704
[20] Kozłowski E., Mazurkiewicz D., Żabiński T., Prucnal S., Sęp J., Assessment model of cutting tool condition for real-time supervision system, Eksploatacja i Niezawodnosc – Maintenance and Reliability, 21 (2019); No 4, 679–685
[21] Li X, Li J, He D, Qu Y. Gear pitting fault diagnosis using raw acoustic emission signal based on deep learning. Eksploatacja i Niezawodnosc – Maintenance and Reliability, 21 (2019), No. 3, 403–410
[22] Vališ D, Hasilová K., Forbelská M, Vintr Z, Reliability modelling and analysis of water distribution network based on backpropagation recursive processes with real field data, Measurement 149 (2020), 107026
[23] Bergweiler S. Intelligent Manufacturing based on Self-Monitoring Cyber-Physical Systems, UBICOMM 2015 The Ninth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, 2015.
[24] Monostori L. Cyber-physical production systems: Roots, expectations and R&D challenges, Variety Management in Manufacturing. Proceedings of the 47th CIRP Conference on Manufacturing Systems, Procedia CIRP 17, 2014.
[25] Qian F., Xu G., Zhang L., Dong H. Design of Hybrid NC Control System for Automatic Line, International Journal of Hybrid Information Technology, 2015, vol. 8/4, 185-192.
[26] Deszyńska A,. Modele hazardów proporcjonalnych Coxa, Matematyka stosowana Tom 13/54 2011.
[27] Cox D., Snell E., Ageneral definition of residuals, Journal of thr Royal Statistical Society Series B (Methodological), 30, 248-275, (1968).
[28] Yishu Xue, Elizabeth D. Schifano, Diagnostic for Cox model, Communications for statistical Applications and Methods 2017, Vol. 24, No. 6, 583-604.
[29] Schoenfeld D, Partial residuals for the proportional hazards regression model, Biometrika, 69, 239-241, (1980).