ISSN: 1231-4005
e-ISSN: 2354-0133
ICID: 1133899
DOI: 10.5604/12314005.1133899

DATA ANALYSIS AND PROCESSING FOR THE SYSTEM RELIABILITY NEURAL NETWORK BASED ON EXPERT JUDGMENT

Hoang Nguyen

Gdynia Maritime University, Department of Engineering Sciences
Morska Street 81-87, 81-225 Gdynia, Poland.
tel.:+48 58 6901306
e-mail: hoang@am.gdynia.pl

Abstract

The article presents a data analysis and processing for tuning artificial neural network (ANN) of the anthrop technical system reliability, based on the opinions of experts. In general, the system reliability parameters are functions of operands – physical values – like time to failure, time between failures, duration times of specific reliability or operational states, number of failures in a time interval (event frequencies). These values are easier to be determined by an expert – operator with long year experience – than probabilistic model parameters. It is suggested that they be used in elicitation, for example linguistic estimates of the shares of reliability system elements in the system failure frequency.

The numerical – linguistic elicitation of these opinions was carried out, which turned out to be uncorrelated and not suitable for tuning the network. Data processing method was used with the appropriate adopted analytic hierarchy process (AHP) geometric scale and matrix approximation method evaluations (logarithmic least squares method). Correlation analyses were performed for received input and output data of network and error of data processing method was determined. The results are shown in the example of elicitation and data correlation analyses for tuning the reliability neural network of the ship propulsion system.

Keywords: *artificial neural network, reliability, elicitation, AHP method, expert judgment, data correlation*

1. Introduction

Neural networks can be useful if there are difficulties for formulation and also for solution of an analytical model, but where the network tuning data are available. The data may be objective or subjective, i.e. derived from human memory. Such field is, among other domains, reliability and safety of anthrop technical systems, particularly complex systems where formal models may be burdened with considerable uncertainty.

Tuning of a neural network consists in determining the values of network input/output (I/O) parameters. When objective data are not available, tuning may be based on expert judgements. There are fields of technology where experts can be found only among experienced operators. This is the case discussed in this paper.

The neural network I/O parameter values must be correlated – the non-correlated values are obviously useless for tuning. Obtaining correlated subjective data is an essential difficulty in the considered case. There are several ways of effecting the level of correlation. First of all it is proper selection of experts and methods of judgment elicitation and applying effective methods of processing the obtained data.

2. Data elicitation procedure

Expert is assumed to be well acquainted with the subject he is expected to formulate on his judgment. The knowledge is connected with experience acquired by years-long practice. Expert should also be capable of formulating his judgement. This is connected with level of his education and the language used in the elicitation process, particularly as regards the parameters the expert is expected to estimate. This may be the language of numerical or linguistic values. Numerical values are better but are more difficult to articulate – also errors in judgments are more likely. The analyst designing the reliability investigation method must select properly the category of available experts in each case, the number of experts and the elicitation language to be used. The number and qualifications of available experts may be a limitation.

In the case of reliability, tuning pertains to characteristics expressed by probabilistic values, e.g. reliability function, unreliability function, failure rate, intensity function, or to physical values – operands in those expressions – e.g. failure frequency, time to failure or time between failures.

Preferred candidates for experts are persons having experience in observing the operation process of the elicitation objects for sufficiently long time and also having the proper theoretical knowledge. The reliability analyst must determine the elicitation language and choose the available expert category. For instance, in the reliability investigation of nuclear power station operators of those objects may be counted on – high-class specialists with knowledge of the calculus of probability, and on seagoing ships – members of the crew with various education levels, generally not familiar with probability.

Man is not a good probability estimator. His judgments show biases, weak calibration, incoherence and overconfidence tendency. Dependences may occur between expert judgments. These flaws cannot be fully removed in the elicitation phase [4, 9].

Table 1 contains data on presentation forms of probabilistic judgments. The type of probability distribution is connected with character of the respective event. For instance, the time to failure or maintenance time distributions is continuous and the human error probability is generally estimated by discrete distributions.

Tab. 1. The forms of probabilistic judgments used in reliability

Distribution type	Discrete: two-point or multipoint. Continuous: functional or empirical.
Models of probability distributions	Empirical. Formal – e.g. exponential, normal or Markov processes.
Distribution dimension	Single-dimensional or multi-dimensional.
Frequency of events	Frequent ($p > 0.01$) or rare ($p < 0.01$).
Calibration	With or without calibration by objective data.

Differentiation of frequent and rare events is essential. The latter may be out of the experience of experts, who have not observed them. Estimation of the probability of occurrence of rare events is based on intuition. As regards information in the last row of Tab. 1, significant is the fact of having or not having objective information, which could be used for calibration of the expert judgments. Without such information, the estimation results may bear considerable uncertainty.

Reference [4] describes conditions to be fulfilled in the expert judgment elicitation phase. The main conditions pertain to the selection of experts, instructions, questionnaires and the way they should be filled-in and also independence of judgments and duration of the interview. They are supposed to formulate their judgments entirely on their own, relying on their personal experience [1, 4, 9].

3. Formal model of the system reliability

The anthrop technical object of interest will be treated as a reliability system. It may be a no repairable or repairable system with negligible or non-negligible renewal time. The catastrophic failure state will be modelled as absorbing state.

The first step in programming an investigation is to define its objective and assumptions concerning the investigation subject (definitions of the system and its operational states, formal reliability model, characteristics of the environment). With these assumptions, the system fault tree

(FT) can be constructed. The fault tree allows determining the sets of elements effecting the system reliability and also indirect relations, if it appears helpful in the elicitation process.

We shall continue the consideration assuming the system reliability model as a Markov chain with two operational states: 1 – operational use state with $\lambda(t)$ failure rate and 2 – maintenance state with $\mu(t)$ repair rate. It is a non-homogeneous process with finite renewal time. When transition rates are not time-dependent then the process becomes homogeneous and distributions of the state 1 and 2 duration time are exponential. The availability formula of a homogenous version of the system takes the form:

$$a(t) = P_1(t) = \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu}\exp[-(\lambda+\mu)t]. \tag{1}$$

Parameters of model (1) are the failure $\lambda(t)$ and repair $\mu(t)$ rates. In general, they are time-dependent but may be approximated by constant or constant in intervals values λ and μ. Statistical verification of such simplification is recommended [6, 11]. From the renewal equations:

$$\lim_{t\to\infty}\frac{H(t)}{t} = \frac{1}{T_0}, \tag{2}$$

where $H(t) = E[\nu(t)]$ = expected value of $\nu(t)$; $\nu(t)$ = number of failures in time interval t, T_o = mean time between failure =*MTBF*.

From formula (2) – after sufficiently long time:

$$MTBF = \frac{t}{\overline{\nu(t)}}, \tag{3}$$

where: $\overline{\nu(t)}$ = mean number of failures in time interval t, which can be easily determined from the expert judgments.

It is generally assumed that maintenance times have also exponential distributions with time-independent transition rates μ. This assumption pertains to direct maintenance work time without organizational preparation time and waiting time for beginning the work. In practice, that preparation period may be chaotic, which makes probabilistic estimation of parameter μ difficult or even impossible. The following attitude to the estimation of parameter μ is proposed:

a) adopting the model with negligible renewal time when that time is short compared with usage time:

$$R(t) = \exp(-\lambda\, t), \tag{4}$$

where: $\lambda = {}^{1}/_{MTBF}$ = failure rate; t = time;

b) adopting constant renewal time of individual devices;

c) determining μ from the formula:

$$\mu = \frac{1}{\bar{\theta}}, \tag{5}$$

where: $\bar{\theta}$ = mean renewal time, which can be estimated by the experts.

In general, the reliability model parameters are functions of operands – physical values – like time to failure, time between failures, duration times of specific reliability or operational states, number of failures in a time interval (event frequencies). These values are easier to be determined by an expert than probabilistic model parameters. It is suggested that they be used in elicitation. It is possible to obtain linguistic values of shares of the basic events (system element failures) in the top event (system failure frequencies) of the system fault tree. The linguistic variables have associated sets of values (very rare, rare, occasional, frequent, very frequent). In the case of a large reliability system, the elicitation process of the shares of system elements in the failure frequencies may be subdivided into "layers", for instance in the case of two layers of the system FT – higher and lower – first determine shares of the higher layer elements in the system failure frequencies and then shares of the lower layer elements in the failure frequencies of the higher layer objects. The principles of completeness and disjointness should be maintained. This procedure will increase the expert estimates as it shortens the "distance" of the related objects. Fig. 1 presents flow diagram of the elicitation and data processing algorithm for obtaining correlated sets suitable for NN tuning.

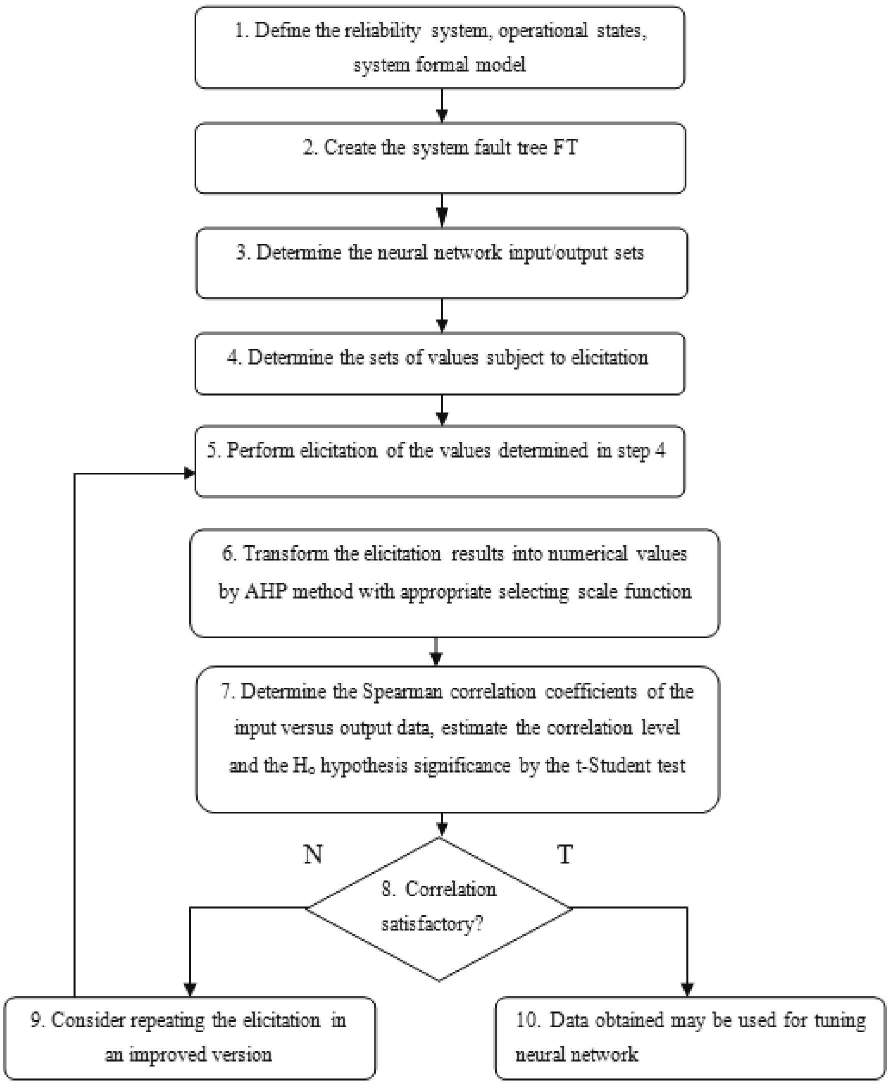

Fig. 1. Algorithm of data analysis and processing for tuning the reliability neural network

4. Processing of the expert data by the AHP method

Linguistic estimates of the shares of reliability system elements in the system failure frequency consist in expert choice of the share value from the set of five values. The estimates are given numbers from 1 to 5. Differences of experts' judgments indicate the scale of preferences in the in pairs comparing the linguistic estimates. Depending on these differences, the preferences are assigned weights $r(s)$ in accordance with a scale function. Then the linguistic judgment matrix R is determined as:

$$\boldsymbol{R} = \boldsymbol{r_{ij}} = \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ \vdots & \ddots & \vdots \\ r_{n1} & \cdots & r_{nn} \end{bmatrix}, \tag{6}$$

where: r_{ij} = preference of *i-th* to *j-th* share (*i,j* = 1,2,…,*n*), with properties: $r_{ij} > 0, r_{ij} = \frac{1}{r_{ji}}\ \forall i, j$.

Matrix $\boldsymbol{R}$ is consistent if its elements fulfil the condition: $r_{ij}r_{jk} = r_{ik},\ \forall i, j$,k = 1,2,…,*n*.

Priority vector $\boldsymbol{p}= (p_1, \dots, p_n)^T$ is determined by approximation of matrix $\boldsymbol{R}$ with matrix $\boldsymbol{P}$, where:

$$\boldsymbol{P} = \begin{bmatrix} {p_1}/{p_1} & \cdots & {p_1}/{p_n} \\ \vdots & \ddots & \vdots \\ {p_n}/{p_1} & \cdots & {p_n}/{p_n} \end{bmatrix}. \tag{7}$$

The preferred method of determining the priority vector is the logarithmic least squares method [10]. Measure of consistency of the processing is the difference between matrices $\boldsymbol{P}$ and $\boldsymbol{R}$ [14]:

$$\mathrm{d(R,P)} = \sqrt{\frac{2}{\mathrm{n(n-1)}} \sum_{\mathrm{i=1}}^{\mathrm{n}} \sum_{\mathrm{j=1}}^{\mathrm{n}} (\mathrm{r_{ij}} - \mathrm{p_{ij}})^2}. \tag{8}$$

The Xu scale [14] is a geometrical scale with parameter c = 2, in the form:

$$r(s) = \left(\sqrt{c}\right)^{I(s)}. \tag{9}$$

where: $I(s)$ = index of the preference symbol s; c = parameter.

It was proved [14] that with this particular value of c parameter the difference between the linguistic judgment matrix and the matrix derived from the priority vector is at a minimum. Tab. 2 shows the Xu scale data with parameter c = 2.

Tab. 2. AHP geometrical scale data (c = 2)

Differences of expert judgments	AHP geometrical scale with parameter c = 2			
	s_i	$I(s)$	$r(s)$	Description of preference
0	s_0	0	1	equally important
1	s_2	2	2	moderately more important
2	s_4	4	4	strongly more important
3	s_6	6	8	evidently more important
4	s_8	8	16	extremely more important
-1	s_{-2}	-2	0.5	moderately less important
-2	s_{-4}	-4	0.25	strongly less important
-3	s_{-6}	-6	0.125	evidently less important
-4	s_{-8}	-8	0.0625	extremely less important

5. The data correlation problem

Data elicited from experts may have a numerical or linguistic form. The latter are adjectives valuating intensity of the measured variable of an object, process or phenomenon. These adjectives may be assigned natural numbers, ascending with increasing intensity, i.e. perform ranking of the measured value. Estimates of the linguistic values are done by means of ordered scales. The scales have order relations. As indicated above, the numerical values, in the case of physical objects observed in the operation process, pertain to the values of independent variables in expressions defining the reliability model parameters. They are estimated in the interval scales. Such scales have constant unit of measurement, order relation and optionally chosen zero point.

The correlation analysis of values measured on the above presented scales is carried out by non-parametric methods. They compensate effects of the standing-out measurements and non-normality of the elicited values [13].

To the correlation, analysis of data obtained from the elicitation process the R. Spearman's method is applied. The output and input data are ranked by assigning them ascending natural numbers starting from 1. The numbers are ranks. The ranking process may be performed also with decreasing sequence. The Spearman's rank correlation coefficient is determined from the following formula [13]:

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}, \tag{10}$$

where: d_i = difference between the ranks of corresponding characteristic values; correlation coefficient $-1 \leq r_s \leq 1$.

Correlation coefficient determined by formula (10) is applicable only to a random sample and correlation of the general population should also be checked. For that the zero hypothesis H_o: ρ = 0, where ρ is the correlation coefficient of the general population, is verified against the alternative hypothesis $H_1 \neq 0$. The *t-Student* test is used for verification. It is assumed that the population has the Student distribution with $n - 1$ degrees of freedom. The test has the form:

$$t = \frac{r_s}{\sqrt{1-r_s^2}}\sqrt{n-1}, \tag{11}$$

where: n = size of the sample.

The test value t is compared with critical value t_p determined from the Student tables for an assumed significance level p and $n - 2$ degrees of freedom. If $t > t_p$, then hypothesis H_0 is rejected and hypothesis H_1 is accepted (correlation exists also in the general population) [13].

6. Example

6.1. Object of analysis, elicitation

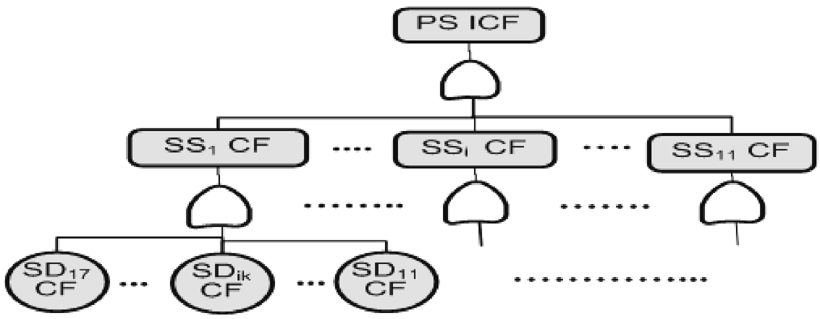

Fig. 2. Fault tree of a ship propulsion system ICF, Legend: PS – propulsion system, ICF – immediate catastrophic failure, CF – catastrophic failure. SS_i – subsystem, i = 1 – fuel oil subsystem, 2 – sea water cooling subs., 3 – low temperature fresh water cooling subs., 4 – high temperature fresh water cooling subs., 5 – starting air subs., 6 – lubrication oil subs., 7 – cylinder lubrication oil subs., 8 – electrical subs ., 9 – main engine subs., 10 – remote control subs., 11 – propeller + shaft line subs. SD_{1k} – set of devices, ik = 11 – fuel oil service tanks, 12 – f. O. Supply pumps, 13 – f. o. Circulating pumps, 14 – f. o. heaters, 15 – filters, 16 – viscosity control arrangement, 17 – piping heating up steam arrangement

The example illustrates reliability analysis of a container carrier propulsion system (PS) with slow-speed piston internal combustion engine and screw propeller, operating in the North Atlantic. Reliability was analysed for immediate catastrophic failures (ICF) of the PS. Fig. 2 presents the FT of the propulsion system. It was assumed that ICFs could occur only during active usage state of the system, i.e. during the ship sea voyage. Share of that operating state time in the entire ship usage time was $\bar{\kappa}$ = 0.8396 (mean value of 50 expert judgments). Detailed data of the example can be found in [1-3].

A questionnaire was presented with definitions of the investigated object, "catastrophic failure" and "sea traffic" as well as tables to be filled in by the experts and suggestions how to do it. The questionnaire was filled in by 50 experts – ship engineers with multi-year experience. Questions were asked about annually frequency of the propulsion system ICF type events, share of subsystem (SS) failures in the PS system failure frequency and share of module (set of devices – SD) ICF type failures in the SS failure frequencies.

6.2. Analysis of the correlation of elicitation results

As regards the propulsion system as a whole, experts gave their subjective estimates of the ICF type failures per year in numbers and linguistically by marking one of fields in the order scale containing numbers and descriptions of that failure frequency. Fig. 3 presents a histogram of the system PS failure frequency per year. The histogram shows a distribution close to the normal distribution, which may be considered correct in the case of observation of dangerous events with more or less steady frequency of occurrence.

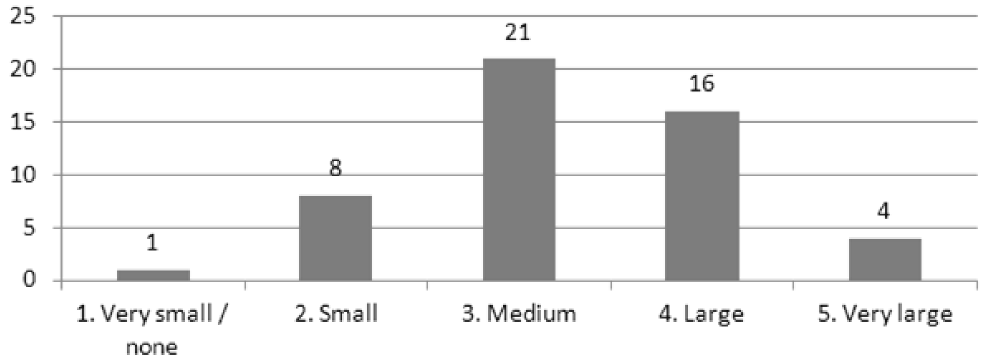

Fig. 3. Distribution of the propulsion system ICF failure frequency

Elicitation of the ICF type failures of SS subsystems and their SD modules consisted in marking appropriate fields in the questionnaire order scales (see Appendix 1). They indicated the share of a given SS or SD in the ICF type failure frequency of a direct higher-level object, i.e. of the propulsion system in the case of SSs and of a specific SS in the case of SD modules. The shares were in pairs compared and the respective differences of numbers were treated as numerical estimates of experts' preferences. The preferences were assigned values in accordance with the geometrical scale function.

Table 1-3 present selected verification results of correlation between the system PS and subsystems SS_i (i = 1,2,…,11), between the subsystem SS_1 and its sets of devices SD_{1k}, (k=1,2,…,7) and between the system PS and sets of devices SD_{1k} (k=1,2,…,7) of the first subsystem SS_1 data appropriately. The correlation coefficients were within the 0.971 – 0.990 range, so the data sets appeared well-correlated (nearly total correlation according to [13]). The zero correlation in the general population H_0 hypothesis was rejected at the 0.01 level. The consistency measure Xu (8) was zero.

Tab. 3. Spearman correlations between system PS and SSi, (i=1,2,...,11) estimates after processing

Subsystems	**SS_1**	**SS_2**	**SS_3**	**SS_4**	**SS_5**	**SS_6**	**SS_7**	**SS_8**	**SS_9**	**SS_{10}**	**SS_{11}**
Spearman coefficients	0.990	0.983	0.983	0.984	0.987	0.985	0.982	0.992	0.993	0.987	0.982
t test	50.3	38.0	37.4	39.2	43.7	40.6	37.1	55.3	58.6	44.6	36.9
t_p (p=0.01)	2.68										
Hypothesis H_0	rejected	rejected	rejected	rejected	rejected	rejected	rejected	rejected	rejected	rejected	rejected

Tab. 4. Spearman correlations between subsystem SS1 and SD1k, (k= 1,2,...,7) estimates after processing

Sets of devices	**SD_{11}**	**SD_{12}**	**SD_{13}**	**SD_{14}**	**SD_{15}**	**SD_{16}**	**SD_{17}**
Spearman coefficients	0.971	0.985	0.986	0.976	0.982	0.987	0.978
t test	28.7	41.0	42.0	31.4	37.3	44.1	33.3
t_p (p = 0.01)	2.68						
Hypothesis H_0	rejected	rejected	rejected	rejected	rejected	rejected	rejected

Tab. 5. Spearman correlations between PS and sets of devices SD1k (k=1,2,…,7) estimates in the subsystem SS1, after processing

SD_{1k}*	SD_{11}	SD_{12}	SD_{13}	SD_{14}	SD_{15}	SD_{16}	SD_{17}
Spearman coefficients	0.982	0.990	0.990	0.987	0.990	0.993	0.987
t test	37.1	51.3	50.3	43.5	51.4	59.2	44.6
t_p (p = 0.01)	2.68						
Hypothesis H_0	rejected	rejected	rejected	rejected	rejected	rejected	rejected

7. **Conclusion**

The above presented data allow concluding that the used neural network tuning procedure gives correct results. It appears appropriate when experts are experienced operators of the reliability analysis objects. It may be useful for network tuning in the reliability analyses and for the technical system risk management. Further study would be focused on adopting the available probabilistic models for calibrating the expert data, which are extremely crucial, but often uncertain in the subjective investigations.

References

[1] Brandowski, A., *Estimation of the Probability of Propulsion Loss by a Seagoing Ship Based on Expert Opinions,* Polish Maritime Research 1(59), Vol. 16, p. 73-77, 2009.
[2] Brandowski, A., Mielewczyk, A., Nguyen, H., Frackowiak, W., *Certain propulsion risk prediction model of a seagoing ship,* Proceedings of ESREL, Conference, 2010
[3] Brandowski, A., Mielewczyk, A., Nguyen, H., Frąckowiak, W., *A fuzzy – neuron model of the ship propulsion risk prediction,* Journal of KONBiN, No 1(13), pp. 117-128, 2010.
[4] Cooke, R., M., *Experts in Uncertainty*, Oxford University Press, New York, Oxford 1991.
[5] Elliot, M., A., *Selecting numerical scales for pairwise comparisons,* Reliability Engineering and System Safety 95, pp. 750-763, 2010.
[6] Gniedienko, B. W., Bielajew, J. K., Sołowiew, A. D., *Mathematical methods in the reliability theory,* Wydawnictwa Naukowo-Techniczne, Warszawa 1968.
[7] IMO – Resolution A.849(20), *Code for the investigation of marine casualties and incidents,* London 1997.
[8] Jaźwiński, J., Smalko, Z., *The use of expert method for estimation of the beta distribution parameters for evaluation of non-defectibility and safety of technical means of transport,* Wyd. Instytutu Technologii Eksploatacji, Radom 2001.
[9] Kahneman, G., Slovic, P., Tversky, A., *Judgment under uncertainty: Heuristics and biases*, Edited by: Cambridge University Press, 2001.
[10] Kwiesielewicz, M., *Analytical hierarchical decision process, Non-fuzzy and fuzzy pairwise comparison,* Instytut Badań Systemowych, PAN, Warszawa 2002.
[11] Modarres, M., Kaminskiy, M., Krivtsov, *Reliability Engineering and Risk Analysis,* New York, Basel: Marcel Dekker Inc., 1999.
[12] Saaty, T. L. *The analytic hierarchy process*, McGraw Hill, NewYork 1980.
[13] Stanisz, A., *Intelligible course of statistics, Vol. 1,* StatSoft, Kraków 2006.
[14] Yucheng, Dong, Yinfeng, Xu, Hongyi, Li, Min, Dai, *A comparative study of the numerical scales and the prioritization methods in AHP*, European Journal of Operational Research 186, pp. 229-242, 2008.