XX Jubileuszowa-Krajowa, VIII Międzynarodowa Konferencja Naukowo - Techniczna
„ZAOPATRZENIE W WODĘ, JAKOŚĆ I OCHRONA WÓD"
XX-th Jubilee-National, VIII-th International Scientific And Technical Conference
„WATER SUPPLY AND WATER QUALITY"
Poland 15-18 June 2008

Barbara TCHÓRZEWSKA-CIEŚLAK

Faculty of Civil and Environmental Engineering
Rzeszów University of Technology
Rzeszów

UNCERTAINTY IN ANALYSIS OF RISK CONNECTED WITH WATER DISTRIBUTION SUBSYSTEM FUNCTIONING

NIEPEWNOŚĆ W ANALIZIE RYZYKA ZWIĄZANEGO Z FUNKCJONOWANIEM PODSYSTEMU DYSTRYBUCJI WODY

In the paper the problem of collecting and processing data necessary to analyse risk of water distribution subsystem functioning, has been presented. The basic definitions concerning the uncertainty in data analysis and sources of uncertainties, have been shown. The method of risk analysis taking into account the so called uncertainty of the obtained data and the possibility of using the fuzzy sets theory to asses risk of water distribution subsystem, have been proposed.

W pracy zaprezentowano problem gromadzenia oraz przetwarzania danych potrzebnych do analizy ryzyka funkcjonowania podsystemu dystrybucji wody. Przedstawiano podstawowe definicje dotyczące niepewności w analizie danych oraz źródła niepewności. Zaproponowano metodę analizy ryzyka uwzględniającą tzw. niepewność pozyskanych danych oraz możliwość wykorzystania teorii zbiorów rozmytych do oceny ryzyka podsystemu dystrybucji wody.

1. Introduction

The objective reality in the water distribution subsystem (WDS) functioning is the possibility of the occurrence of various types of the undesirable events (failures) , which have a direct impact on reliability and safety of the whole water supply system (WSS). These failures do not appear without reason, they can be a result of a serious of events (emergency scenario), the so called "domino effect". They also occur as a result of incorrect decisions, which cause the negative consequences during the WDS functioning. For the right analysis of risk connected with the WDS functioning an appropriate amount of different information, data recording and the possibility of their processing, which is not simple in practice, are necessary.

Risk analysis connected with the WDS functioning is often performed in the so called „uncertain information conditions", which is connected with uncertain (incomplete, imprecise or undependable) data concerning subsystem operating. The measure of data inaccuracy can be the so called quantitative uncertainty. This notion was introduced through the document "Guide to the Expression of Uncertainty in Measurement", released in 1993 by the International Organization for Standardization ISO, which became an international standard being in force also in Poland [10]. According to the mentioned above standard , uncertainty of measurement is defined as a parameter connected with a result of measurement or data accuracy. It is important to distinguish between the notion of measurement error and measurement uncertainty. Error is a random variable and uncertainty is a parameter of the error probability distribution.

The purpose of this paper is to present a problem of uncertainty of data necessary to estimate and analyse risk in the WDS. The paper contains the basic information concerning uncertainty in data analysis and the proposal of using of the fuzzy logic theory to analyse risk of the WDS functioning, if the uncertain data occur.

2. Risk in the WDS functioning

As a result of the occurrence in the WDS the so called representative emergency scenario (RES), marked as S_i, water consumers are subjected to the possibility of the loss of safety. The measure of the loss of drinking water consumers safety is risk connected with the possibility that they will use low quality water or will suffer the lack of water [7]. Risk (r) is a function of three parameters: the probability P_{Si} that *i* representative emergency scenario S_i will occur, the magnitude of losses C_{Si} caused by *i* representative emergency scenario S_i and consumers protection against *i* representative emergency scenario S_i, marked O_{Si},$r = f(P_{Si}, C_{Si}, O_{Si})$ [6,9]. The formula to determine the size of risk connected with the WDS functioning is the following:

$$r = \bigcup_{i=1}^{N} P_{Si} \cdot W_{Si} \tag{1}$$

$$W_{Si} = \frac{C_{Si}}{O_{Si}} \tag{2}$$

where:

S_i - *i* representative emergency scenario (RES), described as a series of the successive undesirable events (failures),

P_{Si} - the probability that *i* representative emergency scenario will occur,

W_{Si} – a weight coefficient associated with a magnitude of losses caused by *i* RES and protection against *i* RES,

C_{Si} – a point value for the parameter of losses caused by *i* RES, assumed according to the following descriptive and point scale: little –1, medium-2, large-3,

O_{Si} - a point value for a protection parameter against *i* RES, assumed according to the following descriptive and point scale: low-1, medium-2, high-3,

N - a number of RES which can occur in WDS.

For the parameter of losses C_{Si} the following criteria and point weights for the assumed descriptive and point scale were used:

- little, $C_{Si}=1$ – perceptible organoleptic changes in water, isolated consumer complaints , financial losses up to $5 \cdot 10^3$ PLN,
- medium, $C_{Si}=2$ – considerable organoleptic difficulty (odour, changed colour and turbidity), consumers health problems, numerous complaints, information in local public media , financial loss up to 10^5 PLN,
- large, $C_{Si}=3$ – the endangered people require hospitalisation, professional rescue teams involved, serious toxic effects in test organisms , information in nationwide media, financial loss over 10^5 PLN.

For the monitoring-warning- blocking parameter O_{Si}, the following criteria and point weights for the assumed descriptive and point scale were used: :

- low protection level - $O_{Si} = 1$, municipal water quality standard monitoring, according to current regulation concerning drinking water quality (monitoring in the selected WSS points, evenly located in the whole area of water supply, and especially: water intake, water treatment station control points, places where water flows into network, in the selected water pipe network points),
- medium protection level - $O_{Si} = 2$, the WSS functioning over standard monitoring, (full water pipe network monitoring, e.g. using SCADA software),
- high protection level - $O_{Si} = 3$, the WSS functioning special monitoring (e.g. within the framework of the multi-barrier system including raw water biomonitoring based on test organisms, and using industrial television with movement detectors for strategic objects).

To determine a value of the probability that the given representative emergency scenario P_{Si} will occur, one needs to determine the probability that the particular events included in *i* RES will occur, based on the statistical data and the assumed probability distribution. Calculation of the probability values P_{Si} can be made by means of the Event Tree Analysis or the Fault Tree Analysis.

In order to determine risk using the matrix method according to the formula (1) it is necessary to use appropriate descriptive and point scale for the probability category [7].

For the probability category P_{Si} the following criteria for the assumed 5 degree descriptive and point scale were proposed:

- probable, frequency: a few times a year, point weight $Ps_i = 5$,
- quite probable, frequency: once a year , point weight $Ps_i = 4$,
- little probable, frequency: 1 in 10 years , point weight $Ps_i = 3$,
- impossible, frequency: 1 in 50 years; point weight $Ps_i = 2$,
- very improbable, frequency: 1 in 100 years, point weight $Ps_i = 1$.

In table 1 the values of the coefficient W_{Si} depending on the point weights C_{Si} and O_{Si} are presented:

Tab. 1. The values of W_{Si}

Tab. 1. Wartości W_{Si}

C_{Si} \ O_{Si}	1	2	3
	W_{Si}		
1	1	0,5	0,33
2	2	1	0,67
3	3	1,5	1,00

In this way the size of risk r calculated from the formula (1) for the individual RES (N=1) takes the values within the range (0,33 ÷15]. In three degree scale the following values of risk range, presented in table 2, were used:

Tab. 2. Risk scale

Tab. 2. Skala ryzyka

Risk scale	r
tolerable	(0,33÷3]
controlled	(3÷8]
unacceptable	(8÷15]

If there are several RES, the highest risk value should be used.

3. The notion of uncertainty

The following basic definitions are connected with the notion of uncertainty and errors in risk analysis:

- **measuring error:** deviation of a result of an individual measurement from a real value which generally is not known;
- **statistical error:** measuring error resulting from the whole environmental impact, which is often impossible to identify and eliminate, properties of used measuring instrument and other reasons,
- **systematic error:** error resulting from used measuring method or other reasons (e.g. some known events have an impact on a measurement but they are unable to be eliminated),
- **standard deviation:** estimator that approximates systematic error value, adequate if there is an appropriate number of measurements in one test,

- **gross error, mistake**: it takes place when one of measurement results is significantly different than the others; we can suppose that some event, which caused a large deviation of examined value, had happened. Such results are often rejected during the statistical analysis.
- **systematic uncertainty –** it takes place when some important factor influencing analysed or measured value is not taken into account, it results from the lack of reliable source of information,
- **accidental uncertainty (statistical)** – an impact of different external or internal factors which cannot be avoided (human errors) and the used research method on the result of performed analysis, resulting from random character of the given event.
- **extended uncertainty (total uncertainty) –** a value defining a range around the result of analysis which, according to expectations, can cover a large part of distribution of values that can be assigned to the analysed value in a justified way.

In the analysis of risk of the WDS the most often the statistical uncertainty occurs, caused by the random nature of the analysed event, the impact of the external factors, as well as a time factor, which determines the change of the analysed undesirable event (failure).

There are some sources of uncertainty:

- incomplete or imprecise definition of analysed value (e.g. imprecise definition of failure in water pipe network),
- incomplete knowledge about the impact of environment on the analysed event (e.g. impact of the ground and water conditions on water pipe network failures),
- reading errors and accuracy class of instrument reading,
- inaccurate data obtained from the external sources (data about the WDS operating obtained from waterworks),
- imperfection of the used research method.

4. Uncertainty in the WDS risk analysis

4.1. The reasons of uncertainty in the risk analysis

To analyse uncertainty we usually use the probabilistic methods which require a large amount of data [1]. In many cases data concerning the description of events, e.g. failures in water pipe network, are obtained on the basis of experts information (the WSS users, experienced engineers or scientists). The most difficult is to chose the probability distribution. In practice data concerning risk analysis in the WDS are not only random but also unreliable (incomplete). Uncertainty of such data consists of many elements. Some of them are determined on the basis of data distribution, characterised by a standard deviation. The remaining elements are estimated on the basis of the assumed probability distribution, known from experience, or other information [2,3,8].

The following data, among others, are necessary to perform risk analysis in the WDS:

- data identifying analysed object (name and type of object and its basic technical data). Such data concern highly detailed studies.

- data about failures (undesirable events), repairs and other breaks in the WDS operating (information about failure date, time, duration and its description),
- data concerning the reasons of the undesirable events occurrence,
- data concerning the consequences of those events.

The sources of the data necessary to analyse risk are:

- data collected from waterworks about the WDS operating.
- measurement data (e.g. measurements of pressure and water flow in water pipe network, measurements of water leaks in water pipe network).
- data collected from experts.

The source of uncertainty in the mentioned above data analysis is, the most often, the incomplete or unreliable knowledge about:

- quantitative and qualitative data base concerning the WDS failures,
- water pipe network technical condition assessment,
- imprecise and incomplete information concerning failures (the undesirable events) localization and identification, ground conditions, and so on,
- cause and effect failure assessment,
- experts opinions and expertises.

4.2. Utilization of the fuzzy sets theory in the WDS functioning risk analysis

The probabilistic methods for risk assessment (based on the undesirable events probability distributions) require unequivocal determination of statistical characteristics, which is connected with the necessity of having appropriate amount of highly reliable data. If the obtained data are "highly unreliable", using such methods leads to getting faulty result of the performed analysis.

Having at disposal different types of data in the analysis of risk connected with WDS functioning it is necessary to develop the method which would allow to use reliable (complete) data, as well as data which are unreliable or incomplete but are important from the risk assessment and analysis point of view. The method which can be useful in such situation to assess risk is the method using fuzzy risk analysis. (FRA)

The notion of fuzzy sets was introduced in 1965 by L. Zadeh [4]. Unlike the conventional set, a limit of fuzzy set is not defined precisely, however, there is a gradual assessment from a complete lack of affiliation of an element in a set, through its partly affiliation, till its total affiliation.. This gradual assessment is defined by means of the so called affiliation function μ_A, where A means a set of fuzzy numbers. Fuzzy sets can be used to describe different linguistic notions connected with risk analysis (little, medium, large, very large). A linguistic variable is such variable which characterises the so called fuzzy, imprecise notions, expressed by means of words, e.g. about number 1, high risk, low risk value.

The basic characteristics of the affiliation function are the following [5]:

- the affiliation function (adjustment degree, compliance) associates every element x from the considered area (space) to a value from the interval [0,1] $\mu_A:X\to[0,1]$, which means that every element x from space X belongs to the fuzzy set A with some affiliation degree,
- a fuzzy set A is defined as: $A=\{\mu_A(x), x\}\}$,

- values of affiliation function μ_A are real numbers from the interval [0,1],
- if $\mu_A = 0$ it means the lack of affiliation of variable x in set A,
- if $\mu_A = 1$ it means the total affiliation of variable x in set A,
- if $0 < \mu_A < 1$ it means the partial affiliation of variable x in set A,
- a real value of affiliation function is called affiliation degree, which can be defined by means of functional dependency or in a discreet way,
- the affiliation function can have different shapes, the most often used are the Gauss' function, the triangular function and the trapezoidal function.
- the affiliation function in the triangular form s defined as:

$$\mu_A(x) = \begin{cases} 1 - \dfrac{|x-m|}{s} & \text{for } x \in [m-s, m+s] \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where:
x – a variable (fuzzy number),
m- a central point for which $\mu A = 1$,
s – range width, m-s = l – minimum fuzzy number value, m+s = h – maximum fuzzy number value.

In figure 1 a graphic interpretation of the triangular affiliation function is presented.

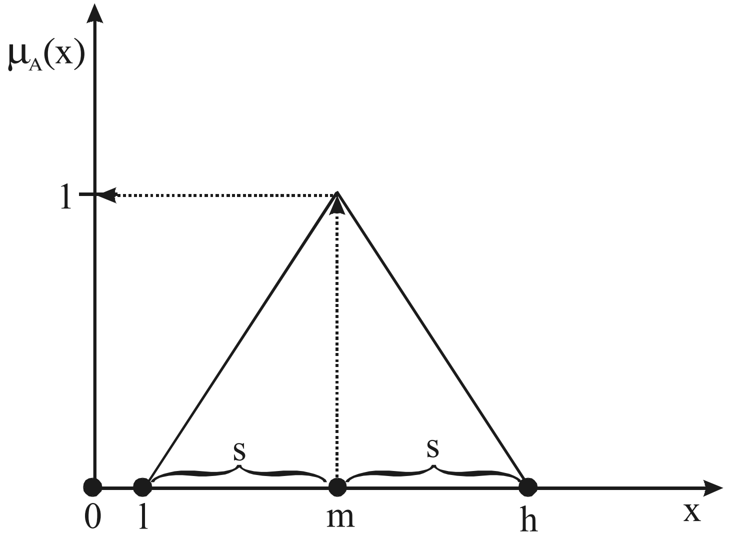

Fig.1 A graphic interpretation of the triangular affiliation function

Rys.1 Graficzna interpretacja trójkątnej funkcji przynależności

The affiliation function defined in such way can be used to change the linguistic type variables into the fuzzy type variables. The particular parameters, e.g. characterizing risk value are described by means of "n" linguistic variables. Then the particular linguistic assessments are assigned to fuzzy numbers x_j, which are defined as threes $x_j=(l_j, m_j, h_j)$, where : j=1,2,…n, with this condition satisfied:

$$0 \le l_j \le m_j \le h_j \le 1 \quad (4)$$

The fuzzy numbers values assigned to the particular linguistic variables are determined according to the following dependences [4]:

- for j=1

$$x_j = (0;0;\frac{1}{n-1}) \quad (5)$$

- for 2≤ j≤n-1

$$x_j = (\frac{j-2}{n-1};\frac{j-1}{n-1};\frac{j}{n-1}) \quad (6)$$

- for j=n

$$xj = (\frac{n-2}{n-1};1;1) \quad (7)$$

where:
x_j – a form of *j* fuzzy number,
n – a number of linguistic variables describing given parameter (event, value, e.g. damage, loss, protection degree, probability, risk, safety),
j – a successive number of linguistic variable, j=1,2,…n.

Linguistic type variables are often used to describe parameters characterizing size of risk , e.g. P_{Si}, C_{Si} , O_{Si}, as well as risk itself, so there is the possibility to change linguistic variables into fuzzy numbers characterizing risk in the WDS.

For the parameters P_{Si}, C_{Si} , O_{Si} the following linguistic variables were assumed:

- for the probability parameter P_{Si}, n=5;
  j =1 - very improbable, j = 2 – improbable, j = 3 – little probable, j = 4 – quite probable, j = 5 – probable
- for the consequences parameter C_{Si}, n=3:
  j = 1 – little, j = 2 – medium, j = 3 – large,
- for the protection parameter O_{Si}, n=3:
  j = 1 -low, j =2 – medium, j=3 – high,

For these assumed linguistic variables the fuzzy numbers values were determined, according to the formulas (5), (6), (7), which is presented in tables 3 and 4.

Tab. 3. Fuzzy numbers for linguistic variables describing parameters P_{Si}, C_{Si}

Tab. 3. Liczby rozmyte dla zmiennych lingwistycznych opisujących parametry P_{Si}, C_{Si}

j	Description for P_{Si}	Fuzzy number
1	very improbable	(0,0; 0,0; 0,25)
2	improbable	(0,0; 0,25; 0,5)
3	little probable	(0,25; 0,5; 0,75)
4	quite probable	(0,5; 0,75; 1,0)
5	probable	(0,75;1,0;1,0)

Tab. 4. Fuzzy numbers for linguistic variables describing parameters O_{Si}, C_{Si}

Tab. 4. Liczby rozmyte dla zmiennych lingwistycznych opisujących parametry O_{Si} C_{Si}

j	Description for O_{Si}, / C_{Si}	Fuzzy number
1	low/little	(0,0; 0,0; 0,5)
2	medium	(0,0; 0,5; 1,0)
3	high	(0,5; 1,0; 1,0)

In fig.2 a graphic interpretation of the affiliation function for risk parameters of WDS functioning is presented.

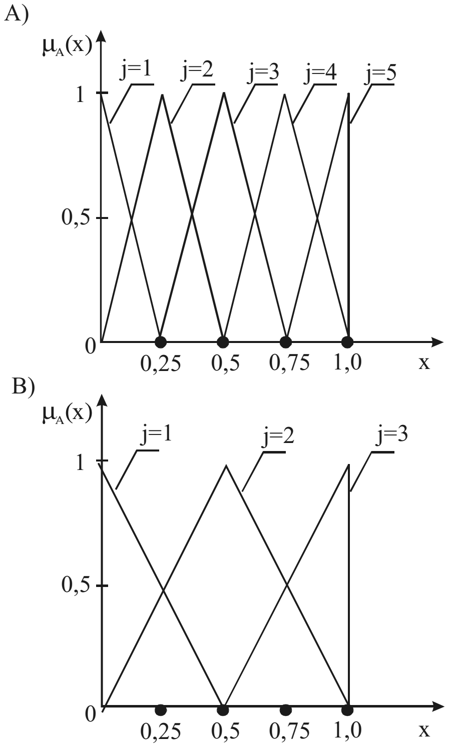

Fig.2 The triangular affiliation function for: A) parameter P_{Si}, B) parameters O_{Si} and C_{Si}

Rys.2 Postać trójkątnej funkcji przynależności dla: A) parametru P_{Si}, B) parametrów O_{Si} i C_{Si}

5. Conclusions

- Analysis of risk connected with the WDS functioning should be the main element of complex WSS risk management.
- For the correct and complete risk analysis and assessment it is necessary to possess large base of different data about subsystem operating.
- If it is impossible to obtain accurate and complete data, then risk analysis is carried out in uncertainty conditions. A solution for this problem can be to apply risk analysis methods using fuzzy set theory (FRA).
- The probability can be defined as a fuzzy value, especially when it is estimated and not defined, which often takes place when the analysis of the undesirable events in the WDS is performed.

- The present paper is a proposal of the utilization of known methods of fuzzy logic theory in analysis and assessment of risk in the WDS for subsystems in which base of operating data is small, incomplete or with a low degree of reliability.

Scientific work was financed from the measures for research in years 2007–2010, as a research project Nr N N523 3765 33 entitled " Development of the methodology of analysis and assessment of risk of failure in water supply systems, with regard to water consumers safety. "

References

[1] Burciu Z. Podejmowanie decyzji przez koordynatora w aspekcie gotowości systemu SAR. Materiały XXXVI szkoły niezawodności. PAN. Wydział Transportu Politechniki Warszawskiej..s:58-66 . 2008

[2] Kluska J., Górski J., Rak J. Możliwość zastosowania logiki rozmytej w systemach techniki grzewczej mat. XI Konferencji Ciepłowników. Wydawnictwo PZITS/O Rzeszów. Solina.s.111-120.1999.

[3] Kwiesielewicz M. Modelowanie niepewności przy użyciu przybliżonych miar prawdopodobieństwa. Materiały Wydziału Elektrotechniki i Automatyki . Politechnika gdańska. Gdańsk 1998r.

[4] Liberacki R. Możliwość zastosowania elementów logiki rozmytej do oceny ryzyka środowiskowego statków. XXVII Sympozjum Siłowni Okrętowych. Zeszyty naukowe Akademii Marynarki Wojennej .Wydawnictwo Akademii Marynarki Wojennej. Gdańsk. s301-308.2006

[5] Osowski S. Sieci neuronowe do przetwarzania informacji. Oficyna Wydawnicza Politechniki Warszawskiej. Warszawa 2000.

[6] Rak J., Tchórzewska-Cieślak B.: Review of matrix methods for risk assessment in water supply system. Wydawnictwo Instytutu Technicznego Wojsk Lotniczych. Journal of Konbin, t.1, z.1, s.67-76, 2006.

[7] Rak J., Tchórzewska-Cieślak B.: Metody analizy i oceny ryzyka w systemie zaopatrzenia w wodę. Oficyna Wydawnicza Politechniki Rzeszowskiej. Rzeszów, 2005

[8] Tchórzewska-Cieślak B. Use of maintenance technique directed to reliability to manage risk connected with water supply system operation. XIX Krajowa Konferencja i VII Międzynarodowa Konferencja "Zaopatrzenie w wodę, jakość i ochrona wód". Poznań - Zakopane, t.2, s.631-638, 2006.

[9] Tchórzewska-Cieślak B.: Method of assessing of risk of failure in water supply system. European safety and reliability conference ESREL. Risk, reliability and societal safety. Taylor & Francis, t.2, s.1535-1539, Norway, Stavanger 2007.

[10] Guide to the Expression of Uncertainty in Measurement" . Norma ISO.1993r