

# Uncertainty Modeling in Risk Analysis: A Fuzzy Set Approach

Palash Dutta

Dept. of Mathematics  
Dibrugarh University,  
Dibrugarh-786004, India

Tazid Ali

Dept. of Mathematics  
Dibrugarh University,  
Dibrugarh-786004, India

## ABSTRACT

Risk assessment is a popular and important tool in decision making process. Risk assessment is generally performed using *models* and model is a function of some parameters which are usually affected by uncertainty. Here, we consider that model parameters are affected by epistemic uncertainty. To represent epistemic uncertainty in general triangular fuzzy number or trapezoidal fuzzy numbers are used. In this paper, we study Gaussian fuzzy number to represent epistemic type uncertainty and try to fuse with triangular fuzzy numbers and also risk assessment is carried out under fuzzy environment.

## Keywords

Risk assessment, epistemic uncertainty, Gaussian fuzzy number

## 1. INTRODUCTION

Human being is always exposed to radiation either from natural or artificial/manmade sources in the environment. While there have been natural nuclides since the beginning the earth's existence, manmade nuclides have been released from nuclear installations and fallouts from the nuclear test and nuclear accident. Also produced water is the most significant source of waste generated in the production phase of oil and gas operations. Once discharged into the ocean, a number of heavy metals and poly aromatic hydrocarbon in produced water may introduce toxicity and bioaccumulation in marine organisms. Similarly chemical organic compounds such as chlorinated organic solvent (1,1,2-trichloroetene) etc. are also harmful substances to the aquatic organism. These compounds are harmful to fish and therefore human can be affected through intake of such fishes. Consequently, we can say that human health can also be indirectly (or directly) affected through different pathways such as inhalation, ingestion, submersion and dermal contact.

When hazardous substances are released into the environment, an evaluation is necessary to determine the possible impact such substances may have on human health and other biota. For this purpose, risk assessment is performed to quantify the potential detriment to human and evaluate the effectiveness of proposed remediation measures. The assessment is performed using 'models' and a 'model' is a function of parameters which are usually affected by aleatory and epistemic uncertainty. Aleatory uncertainty and epistemic uncertainty are two distinct facets of uncertainty in risk assessment. Aleatory uncertainty arises from heterogeneity or the random character of natural processes while epistemic uncertainty arises from the partial character of our knowledge of the natural world. Epistemic uncertainty can be reduced by further study while aleatory cannot be reduced. Here, we have considered that parameters are affected by epistemic uncertainty (fuzzy number).

If the representation of the input parameters of the risk model is in probabilistic sense, the output risk distribution may reflect probabilistic information. However, resalable and sufficient data is required to estimate and characterized the probability distribution of the input variables. If uncertainty does not arise due to randomness, or if the information are partial, not fully reliable, receive of information from more than one sources or inherent imprecision, then probability theory is inappropriate to represent such kind of uncertainty.

To overcome this limitation of probability theory Zadeh (1965) introduced fuzzy set theory which can be used to incorporate epistemic uncertainty. In general, Triangular or Trapezoidal shape fuzzy membership functions have been widely studied in literature to represent uncertainty. However, in practice, there are certain applications where to represent uncertainty, besides Triangular or Trapezoidal shape fuzzy numbers some other types of fuzzy numbers come into picture viz., Gaussian fuzzy number, lognormal shaped fuzzy number, Cauchy shaped fuzzy number etc. In [3] authors studied Gaussian fuzzy number and Lognormal shaped fuzzy number and their arithmetic operations respectively. Here, we also study Gaussian fuzzy numbers and perform arithmetic operations between Gaussian fuzzy number and triangular fuzzy number. Finally, we apply it in risk assessment models to demonstrate and check its validity. Here, we consider two scenarios. In scenario1, parameters are considered as Gaussian fuzzy numbers, scenario2 some parameters are considered as Gaussian fuzzy numbers while some parameters are considered as triangular fuzzy numbers and finally risk assessment is carried out. Also, uncertainty involved in the output risk also measured using *average with* for each scenario.

## 2. BASIC CONCEPT OF FUZZY SET THEORY

In this section, some necessary backgrounds and notions of fuzzy set theory ([1], [2]) are reviewed.

**Definition 2.1:** Let  $X$  be a universal set. Then the fuzzy subset  $A$  of  $X$  is defined by its membership function

$$\mu_A : X \rightarrow [0,1]$$

Which assign a real number  $\mu_A(x)$  in the interval  $[0, 1]$ , to each element  $x \in A$ , where the value of  $\mu_A(x)$  at  $x$  shows the grade of membership of  $x$  in  $A$ .

**Definition 2.2:** Given a fuzzy set  $A$  in  $X$  and any real number  $\alpha \in [0, 1]$ . Then the  $\alpha$ -cut or  $\alpha$ -level or cut worthy set of  $A$ , denoted by  ${}^\alpha A$  is the crisp set

$${}^\alpha A = \{x \in X : \mu_A(x) \geq \alpha\}$$

The strong  $\alpha$  cut, denoted by  ${}^{\alpha+}A$  is the crisp set

$${}^{\alpha+}A = \{x \in X : \mu_A(x) > \alpha\}$$

**Definition 2.3:** The support of a fuzzy set  $A$  defined on  $X$  is a crisp set defined as

$$\text{Supp}(A) = \{x \in X : \mu_A(x) > 0\}$$

**Definition 2.4:** The height of a fuzzy set  $A$ , denoted by  $h(A)$  is the largest membership grade obtain by any element in the set and it is denoted as  $h(A) = \sup_{x \in X} \mu_A(x)$

**Definition 2.5:** A fuzzy number is a convex normalized fuzzy set of the real line  $R$  whose membership function is piecewise continuous.

**Definition 2.6:** A triangular fuzzy number  $A$  can be defined as a triplet  $(a, b, c)$ . Its membership function is defined as:

$$\mu_A(x) = \begin{cases} \frac{x-a}{b-a}, & a \leq x \leq b \\ \frac{c-x}{c-b}, & b \leq x \leq c \end{cases}$$

**Definition 2.7:** A trapezoidal fuzzy number  $A$  can be expressed as  $(a,b,c,d)$  and its membership fuzzy number is defined as:

$$\mu_A(x) = \begin{cases} \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & b \leq x \leq c \\ \frac{d-x}{d-c}, & c \leq x \leq d \end{cases}$$

### 3. GAUSSIAN FUZZY NUMBER

Although, Triangular or Trapezoidal shape fuzzy membership functions have been widely studied in literature to represent uncertainty. However, in practice, there are certain applications where to represent uncertainty besides Triangular or Trapezoidal shape fuzzy numbers some other types of fuzzy numbers come into picture viz., Gaussian fuzzy number [3], lognormal shaped fuzzy number etc.

To define a fuzzy number in the form of Gaussian distribution and so, the membership function required for building the Gaussian shaped fuzzy set must be expressed as follows:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)$$

where  $\mu$  and  $\sigma$  represent the mean and the standard deviation respectively.

By definition, a fuzzy number is a fuzzy set whose membership function  $\mu_A : \rightarrow [0,1] \Leftrightarrow \mu_A(A) = 1$  i.e., there must be at least one domain element whose membership grade equals 1, and  $\exists x_1, x_2 \in R : x_1 < x_2 \forall x \in [x_1, x_2] : \mu_A(x) = 0$ .

Hence, by normalizing the fuzzy set, we obtain the following expression for membership function of the fuzzy set:

$$\mu_A(x) = \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) \quad \dots(1)$$

For fuzzy set to become convex, one calculates the inflexion points of the domain of the Gaussian by making the second derivative equal to zero, i.e.,  $f''(x) = 0$ . Thus, the domain of the normalized and convex fuzzy set is formed by the interval  $[\mu - \sigma, \mu + \sigma]$ . However, for this domain, only 68% of the information contained in the Gaussian will be represented by the fuzzy number. Due to small amount of Gaussian information contained in the fuzzy set, a new interval is considered,  $[\mu - 3\sigma, \mu + 3\sigma]$ , as the domain of the fuzzy number which represents approximately 99.7% of the information contained in the Gaussian and established an  $\alpha$ -cut at 0.01. Here, convexity constrained is relaxed and operation on Gaussian fuzzy number is performed.

The domain  $R$  of the fuzzy number is bounded to the interval  $[\mu - 3\sigma, \mu + 3\sigma]$  and whose  $\alpha$ -cut is placed at 0.01.

For this fuzzy number, the  $\alpha$ -cut is defined as:

$${}^\alpha A = [\mu - \sigma\sqrt{-2\ln \alpha}, \mu + \sigma\sqrt{-2\ln \alpha}]$$

Where  $\mu$  and  $\sigma$  represent the mean and the standard deviation respectively and  $\alpha$  corresponds to the  $\alpha$ -cut defined in the interval  $[0.01, 1]$ . Then interval arithmetic can be applied to perform operations on Gaussian fuzzy numbers, where each possibilistic interval defined by the  $\alpha$ -cut can be treated independently.

### 4. COMBINATION OF GAUSSIAN AND TRIANGULAR FUZZY NUMBER

Here, we perform arithmetic operation between Gaussian and triangular fuzzy numbers and apply in health risk assessment models. Since values of model parameters of health risk

assessment are positive so interval arithmetic are enough to combine these types of fuzzy numbers.

As the general form of the  $\alpha$  -cut of Gaussian fuzzy number is

$${}^\alpha A_1 = [\mu - \sigma\sqrt{-2\ln\alpha}, \mu + \sigma\sqrt{-2\ln\alpha}] \quad \dots(2)$$

where  $\mu$  and  $\sigma$  represent the mean and the standard deviation respectively and  $\alpha$  corresponds to the  $\alpha$  -cut defined in the interval [0.01, 1].

The general form of the  $\alpha$  -cut of triangular fuzzy number is

$${}^\alpha A_2 = [(b-a)\alpha + a, c - (c-b)\alpha] \quad \dots(3)$$

where  $a$ ,  $b$ , and  $c$  are minimum, most likely and maximum value respectively and  $\alpha \in [0,1]$ . Here, we will relax the convexity condition to combine Gaussian fuzzy number and triangular fuzzy number using  $\alpha$  -cut by considering that  $\alpha$  corresponds to the  $\alpha$  -cut defined in the interval [0.01, 1]. It will not affect the uncertainty involved in the fuzzy number.

**Addition:**

$${}^\alpha A_1 + {}^\alpha A_2 = [\mu - \sigma\sqrt{-2\ln\alpha} + (b-a)\alpha + a, \mu + \sigma\sqrt{-2\ln\alpha} + c - (c-b)\alpha]$$

**Subtraction:**

$${}^\alpha A_1 - {}^\alpha A_2 = [\mu - \sigma\sqrt{-2\ln\alpha} - (c - (c-b)\alpha), \mu + \sigma\sqrt{-2\ln\alpha} - ((b-a)\alpha + a)]$$

**Multiplication:**

$${}^\alpha A_1 \times {}^\alpha A_2 = [(\mu - \sigma\sqrt{-2\ln\alpha}) \times ((b-a)\alpha + a), (\mu + \sigma\sqrt{-2\ln\alpha}) \times (c - (c-b)\alpha)]$$

**Division:**

$$\frac{{}^\alpha A_1}{{}^\alpha A_2} = \left[ \frac{\mu - \sigma\sqrt{-2\ln\alpha}}{(c - (c-b)\alpha)}, \frac{\mu + \sigma\sqrt{-2\ln\alpha}}{((b-a)\alpha + a)} \right]$$

## 5. AVERAGE WIDTH OF A FUZZY NUMBER

Since a fuzzy number can be considered as a generalization of an interval, we define width of a fuzzy number and called it average with. Let A be a continuous triangular fuzzy number whose membership function is defined as:

$$\mu_A(x) = \begin{cases} \frac{x-a}{b-a}, & a \leq x \leq b \\ \frac{c-x}{c-b}, & b \leq x \leq c \end{cases}$$

Then alpha-cuts of A are  $[(b-a)\alpha + a, c - (c-b)\alpha]$  where  $\alpha \in [0,1]$ .

We calculate average width of A following the steps below:

Step 1: Divide [0,1] into N numbers of subintervals and consider each value as an alpha value.

Step 2: Then find alpha-cuts for each alpha-value.

i.e.,  $[(b-a)\alpha + a, c - (c-b)\alpha]$ , where  $\alpha \in [0,1]$ .

Step 3: Calculate width of each alpha-cut (alpha-cut gives closed interval).

$$\begin{aligned} \text{That is, } & c - (c-b)\alpha - \{(b-a)\alpha + a\} \\ & = (c-a)(1-\alpha) \end{aligned}$$

Step 4: Sum up all the widths and divide by N+1.

That is,  $\frac{(c-a)(1-\alpha)}{N+1}$ , which will give the average width of the fuzzy number A.

In a similar fashion, we can calculate average width of non triangular fuzzy number.

For example, let A be a fuzzy number whose membership function is given as:

$$\mu_A(x) = \begin{cases} \frac{\sqrt{x}-2}{2}, & 4 \leq x \leq 16 \\ 5 - \sqrt{x}, & 16 \leq x \leq 25 \end{cases}$$

Alpha-cuts of A are  $[(2\alpha + 2)^2, (5 - \alpha)^2]$

Dividing [0,1] into 1000 subintervals and considering each value as alpha value, the average width of A is

$$\begin{aligned} & \{sum((5-\alpha)^2 - (2\alpha+2)^2)\} / 1001 \\ & = 10.995 \end{aligned}$$

## 6. CASE STUDY

To demonstrate and make use of the Gaussian fuzzy number and triangular fuzzy numbers and their combination a hypothetical case study for health risk assessment is presented here.

Suppose the exposure of a population to chlorinated organic solvent (1, 1, 2-trichloroethene), via the consumption of contaminated drinking water. The general form of a risk assessment model for a lifetime daily dose of exposure as provided by EPA, 2001 is follows [4]:

$$D = \frac{CDI \times IR \times EF \times ED}{BW \times AT} \quad \dots(4)$$

Where,  $D$  is the dose of exposure (mg/Kg.day),  $IR$  is the intake of water (l/d),  $CDI$  is the concentration of 1, 1, 2-trichloroethene in the water (mg/l),  $EF$  is the exposure frequency (d/yr),  $ED$  exposure duration (yr),  $BW$  is the body weight (kg),  $AT$  is the average time (d).

The excess risk is express as:

$$ER = D \times UER \quad \dots(5)$$

$UER$  is the unit excess risk: a probability of excess cancer per unit daily dose of exposure.

**Scenario1:**

In this scenario, the parameter concentration of 1, 1, 2-trichloroethene ( $CDI$ ) in the water and exposure duration are considered to be Gaussian fuzzy number while other parameters are taken to be constant. The values of the parameters are given in table1.

**Table 1: Parameters used in risk assessment.**

Variable	Representation	Value
CDI	Possibilistic	GFN(0.2, 0.02)
IR	Deterministic	1.5
EF	Deterministic	250
ED	Possibilistic	GFN(30,2)
BW	Deterministic	70
AT	Deterministic	70
UER	Deterministic	0.024

Here, ( $GFN(\mu, \sigma)$ ) indicates Gaussian fuzzy number with mean  $\mu$  and standard deviation  $\sigma$ .

The results of the risk assessment using (5) of scenario 1 is depicted in figure (1).

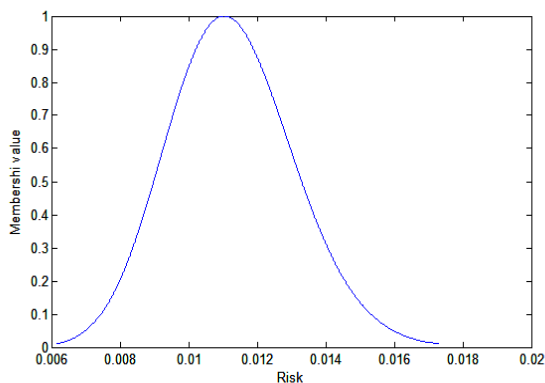


Fig1: Risk Assessment of Scenario1.

Here, some inputs parameters are Gaussian fuzzy number and some are constant therefore, the resulting risk is also obtained in the form of a fuzzy number. 1.102e-02 is the core of the output fuzzy risk while the range is [0.6123e-02, 1.727e-02] and the uncertainty involved in the output risk as 0.0045.

**Scenario2:**

In this scenario, the parameter intake rate of water ( $IR$ ) is considered as triangular fuzzy number and exposure duration ( $ED$ ) is considered to be Gaussian fuzzy number while other parameters are taken to be constant. The parameters values used in the risk assessment are given in table2.

**Table 2: Parameter values used in the risk assessment.**

Variable	Representation	Value
CDI	Possibilistic	0.2
IR	Deterministic	TFN(1, 1.5, 2)
EF	Deterministic	250
ED	Possibilistic	GFN(30,2)
BW	Deterministic	70
AT	Deterministic	70
UER	Deterministic	0.024

Here, ( $GFN(\mu, \sigma)$ ) indicates Gaussian fuzzy number with mean  $\mu$  and standard deviation  $\sigma$ . TFN(a, b, c) indicates triangular fuzzy number with minimum value a, most likely value b and maximum value.

The results of the risk assessment using (5) of scenario 2 is depicted in figure (2).

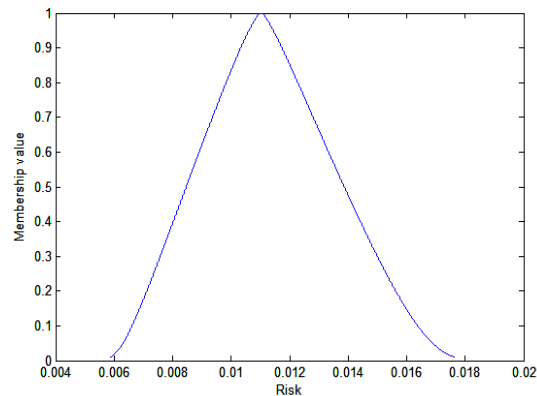


Fig2: Risk assessment of scenario2

Here, the parameter  $IR$  is available in the form of triangular fuzzy number while  $ED$  is available in the form of Gaussian fuzzy number and other parameters are constant. Therefore, output risk is obtained in the form of fuzzy number around 1.102e-02 ranging from 0.597e-02 to 1.762e-02. The uncertainty involved in the output risk is 0.0054.

**7. CONCLUSION**

Risk assessment is a popular and important tool in decision making process. Risk assessment is generally performed using models and model is a function of some parameters which are usually affected by uncertainty. Since we have considered that the parameters are tainted with epistemic uncertainty, the same is propagated to the output through the model. In general, Triangular or Trapezoidal shape fuzzy membership functions have been widely studied in literature to represent uncertainty. However, in practice, there are certain applications where to represent uncertainty besides Triangular or Trapezoidal shape fuzzy numbers some other types of fuzzy numbers come into picture viz., Gaussian fuzzy number, lognormal shaped fuzzy number etc. In this paper, we have studied Gaussian fuzzy numbers and try to fuse with triangular fuzzy number. Finally, we applied it in risk assessment models to demonstrate and

check its validity. Here, we consider two scenarios. In scenario1, parameters are considered as Gaussian fuzzy numbers, and some are constant therefore, the resulting risk is also obtained in the form of a fuzzy number. 1.102e-02 is the core of the output fuzzy risk while the range is [0.6123e-02, 1.727e-02] and the uncertainty involved in the output risk as 0.0045. In scenario2, some parameters are considered as Gaussian fuzzy numbers while some parameters are considered as triangular fuzzy numbers, therefore, output risk is obtained in the form of fuzzy number around 1.102e-02 ranging from 0.597e-02 to 1.762e-02. The uncertainty involved in the output risk is 0.0054.

## **8. ACKNOWLEDGEMENT**

The work done in this paper is under a research project funded by Board of Research in Nuclear Sciences, Department of Atomic Energy, Govt. of India.

## **9. REFERENCES**

- [1] Bojadziev G. and Bojadziev, M. (1995). Fuzzy set, Fuzzy logic, application, (Singapore: world Scientific.)
- [2] Dutta P., Boruah H. and Ali T. (2011). Fuzzy arithmetic with and without using  $\alpha$ -cut method: a comparative study”, International Journal of Latest Trends in Computing, vol-2, pp : 99-108.
- [3] Pacheco M. A.C., Vellasco M .B. R.(2009) Intelligent Systems in Oil Field Development under Uncertainty. (Springer-Verlag, Berlin Heidelberg.)
- [4] US EPA. 2001. Risk Assessment Guidance for Superfund, Volume I: Human Health Evaluation Manual (Part E, Supplemental Guidance for Dermal Risk Assessment). Office of Emergency and Remedial Response, EPA/540/R/99/005, Interim, Review Draft. United States Environmental Protection Agency. September 2001.