Modeling of Uncertainty in Dose Assessment using Probability-Possibility Transformation

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

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

ABSTRACT

Uncertainty parameters in risk assessment can be modeled by different ways viz. probability distribution, possibility distribution, belief measure, depending upon the nature and availability of the data. Different transformations exist for converting expression of one form of uncertainty to another form. In this paper, we reviewed the consistency principles as given by different researchers. Then we have carried out dose assessment using probability- possibility transformation satisfying consistency conditions.

KEYWORDS

Uncertainty, Risk Assessment, Probability-possibility transformation

1 INTRODUCTION

In some situations, some models parameters of radiological risk assessment may be affected by variability and uncertainty simultaneously. Basically, transforming probabilistic data to possibilistic data is useful when weak source of information make probabilistic data unrealistic. Also, it is useful in order to explore the advantages of possibilistic theory at combination steps, or perhaps to reduce the complexity of the solution when computing with possibility values rather than with probability values.

Transforming from possibility to probability may be meaningful in the case of decision making where a precise outcome is often preferred, such that, the decision maker is interested to know "what is likely to happen in future", instead of "what is possible in future" [8].

According to [3] transforming possibility measure into probability measure or conversely can be useful in any problem where heterogeneous uncertain and imprecise data must be dealt with (e.g. subjective, linguistic like evaluation and statistical data).The possibilistic representation is weaker because it explicitly handles imprecision (i.e., incomplete knowledge) and because possibility measure are based on ordering structure rather than additive one. Therefore, it can be concluded that transforming a probabilistic representation to possibilistic representation, some information is lost because we go from point value probabilities to interval values ones. The converse transformation from possibility adds information to some possibilistic incomplete Knowledge.

Different transformations exist for converting expression of one form of uncertainty to another form. They differ from one another substantially, ranging from simple ratio scaling to more sophisticated transformation based upon various principles. These transformations should satisfy certain consistency principles. Here, we reviewed the consistency principles as given by various authors viz., Zadeh, Klir, Dubois & Prade. Then dose assessment is carried out using Probability-Possibility transformations (as [1]).

2 TRANSFORMATION CONSISTENCY PRINCIPLES

When information regarding some phenomenon is given in both probabilistic and possibilistic terms, the two descriptions should be in some sense consistent. That is, given probabilistic representation p_i and possibilistic representation π_i on X, the two representations should satisfy some consistency condition. Although various consistency conditions may be required, the weakest one acceptable on intuitive groups can be expressed as follows:

An event that is probable to some degree must be possible at least to the same degree. That is, the weakest consistency condition is expressed formally by the inequality

$$p_i \le \pi_i$$

On the other hand, the strongest consistency condition would require that any event with nonzero probability must be fully possible.

$$p_i > 0 \Rightarrow \pi_i = 1.$$

In this section different consistency principles [2], [5], [7], [10] are reviewed

2.1. Zadeh consistency principle

Zadeh defined the probability-possibility consistency principle such as "a high degree of possibility does not imply a high degree of probability, nor does a low degree of probability imply a low degree of possibility" (Zadeh 1978).

Let *U* be a finite set. *X* is a variable taking a value in *U*. π_i and p_i are possibility and probability that $X = u_i \in U$ respectively. Then, Zadeh's consistency principle can be expressed by

$$\pi_i = 0 \Rightarrow p_i = 0 \text{ and } \pi_i > \pi_j \Rightarrow p_i > p_j .$$

He defined the degree of consistency between a probability $p = (p_1, p_2, p_3 \ldots\ldots p_n)$ and a possibility distribution $\pi = (\pi_1, \pi_2, \pi_3 \ldots. \pi_n)$ as:

$$r = \sum_{i=1}^{n} \pi_i p_i ....(1)$$

From (1), it can be check that

(i) if $\pi_i = 0, \forall i$ then $r = 0$, no consistency available. An impossible event cannot be probable.

(ii) If $\pi_i = 1, \forall i$ then $r = 1$, the maximum consistency value is reached. Any probability measure is still consistent with this probability measure.

Maximizing the degree of consistency, however, poses us a very restrictive condition that; $\pi_i < 1 \Rightarrow p_i = 0.$

2.2: Klir consistency principle:

Let $X = \{w_1, w_2, ...w_n\}$ be a finite universe, let $p_i = p_i(w_i)$ and $\pi_i = \pi_i(w_i)$. Assume that the elements of X are ordered in such a way that: $\forall i = 1, 2, ...n : p_i > 0, p_i \geq p_{i+1}$ and $\pi_i > 0, \pi_i \geq \pi_{i+1}$ with $p_{n+1} = 0$ and $\pi_{n+1} = 0$.

According to Klir the transformation from p_i to π_i must preserve some appropriate scale and the amount of information contained in each distribution (Klir 1993).The information contained in p or π can be expressed by the equality of their uncertainties. Klir has considered the principle of uncertainty preservation under two scales.

The ratio scale: This is a normalization of the probability distribution. The transformation $p \rightarrow \pi$ and $\pi \rightarrow p$ are named the normalized transformations and they are defined by

$$\pi_i = \frac{p_i}{p_1}, p_i = \frac{\pi_i}{n\sum_{i=1}^{n} \pi_i} ....(2)$$

The log-interval scales: the corresponding transformation $p \rightarrow \pi$ and $\pi \rightarrow p$ are define by:

$$\pi_i = (\frac{p_i}{p_1})^{\alpha}, p_i = \frac{\pi_i^{\frac{1}{\alpha}}}{\sum_{i=1}^{n} \pi_i^{\frac{1}{\alpha}}} ....(3)$$

These transformations are known as Klir transformation satisfying the uncertainty preservation principle defined by Klir (1993). α is a parameter that belongs to the open interval (0, 1). According to Klir any transformation should be based on the following three assumptions:

- A scaling assumption that forces each value $(\pi_i)_i$ to be a function of p_i/p_1 (where $p_1 \geq p_2 \geq ......\geq p_n$) that can be ratio scale, interval scale, log-interval scale transformation etc.
- An uncertainty invariance assumption according to which the entropy *H(p)* should be numerical equal to the measure of information $E(\pi_i)$ contained in the transformation π_i to *p*.
- The transformation should satisfy the consistency condition $\forall$ π (u) ≥ *p*(*u*), $\forall$ *i*, starting that what is probable must be possible.

Dubois and Prade gave an example to show that the scaling assumption of Klir may some time lead to violation of the consistency principle that requires $\Pi \geq P$ for all events. The second assumption is also debatable because it assumes possibilistic and probabilistic information measures are commensurate.

5.2.3: Dubois and Prade consistency Principle

The transformation $p \rightarrow \pi$ is guided by the principle of maximum specificity, which aims at finding the most informative possibility distribution. While the transformation $\pi \rightarrow p$ is guided by the principle of insufficient reason which aims at finding the possibility distribution that contains as much as uncertainty as possible but that retains the features of possibility distribution (Dubois 1993). This leads to write the consistency principle of Dubois and Prade such as:

$$\forall A \subset X : \pi(A) \geq p(A).......(4)$$

The transformation $p \rightarrow \pi$ and $\pi \rightarrow p$ are define by

$$\pi_i = \sum_{j=i}^{n} p_j; \quad p_i = \sum_{j=i}^{n} \frac{(\pi_j - \pi_{j+1})}{j} ....(5)$$

The two transformations define by (5) are not converse of each other because they are not based on same informational principle. Therefore, the transformation defined by (5) can be named as asymmetric. Dubois and Prade suggested a symmetric $p \rightarrow \pi$ transformation which is define by:

$$u_i = \sum_{j=i}^{n} \min(p_i, p_j) ....(6)$$

Dubois and Prade proved that the symmetric transformation $p \rightarrow \pi$, define by (6), is the most specific transformation which satisfies the condition of consistency of Dubois and Prade define by (4).

5.3 Probability to Possibility Transformation

Transforming probabilistic data to possibilistic data is useful when weak source of information make probabilistic data unrealistic. Also, it is useful in order to explore the advantages of possibilistic theory at combination steps, or perhaps to reduce the complexity of the solution when computing with possibility values rather than with probability values [8].

5.3.1 Normal probability distribution function to 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, 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 [9]:

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

where μ and σ represent the mean and the standard deviation respectively.

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

Hence, by normalizing the fuzzy set, we obtain the following expression for membership function of the fuzzy set (Pacheco et al, 2009):

$$\mu_A(x)=\exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) \qquad ...(7)$$

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, we consider considered a new interval, $[\mu-3\sigma, \mu+3\sigma]$, as the domain of the fuzzy number which represents approximately 99.7% of the information contained in the Gaussian and established an α -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 α -cut is placed at 0.01.

For this fuzzy number, the α -cut is defined as:

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

Where μ and σ represent the mean and the standard deviation respectively and α corresponds to the α -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 α -cut can be treated independently.

5.3.2 Triangular probability distribution function to triangular fuzzy number

A random variable X is said to be triangularly distributed with lower limit a, upper limit c and mode b such that $a < b < c$, the probability density function is given by

$$f(x)=\begin{cases}\frac{2(x-a)}{(b-a)(c-a)}, a \le x \le b\\ \frac{2(c-x)}{(c-b)(b-a)}, b \le x \le c\\ 0, otherwisw\end{cases}$$

Normalizing the distribution function (dividing the function by the maximum height i.e., by 2/(c-a)) we will have the fuzzy number whose membership function is

$$\mu_A=\begin{cases}\frac{x-a}{b-a}, a \le x \le b\\ \frac{c-x}{c-b}, b \le x \le c\end{cases}$$

5.4 Possibility to Probability Transformation

Transforming from possibility to probability may be meaningful in the case of decision making where a precise outcome is often preferred, such that, the decision maker is interested to know "what is likely to happen in future", instead of "what is possible in future" [8].

5.4.1 Gaussian fuzzy number to normal probability distribution function

As authors in [9] defined the Gaussian fuzzy number, whose membership function is

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

where μ and σ represent the mean and the standard deviation respectively.

Whenever we have mean and the standard deviation and the shape of the distribution, so in such situation where a precise outcome is often preferred, we can consider its probabilistic form as follows:

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

where μ and σ represent the mean and the standard deviation respectively.

5.4.2 Triangular fuzzy number to triangular probability distribution function

Consider a triangular fuzzy number A= [a, b, c] whose membership function is given as:

$$\mu_A=\begin{cases}\frac{x-a}{b-a}, a \le x \le b\\ \frac{c-x}{c-b}, b \le x \le c\end{cases} \qquad \ldots (8)$$

Integrating the Fuzzy set μ_A with respect to x on [a, c] we have,

$$\int_a^c \mu_A dx = \int_a^b \mu_A dx + \int_b^c \mu_A dx$$
$$= \frac{1}{b-a}\int_a^b (x-a)dx + \frac{1}{c-b}\int_b^c (c-x)dx$$
$$= \frac{1}{b-a}\left[\frac{x^2}{2} - ax\right]_a^b + \frac{1}{c-b}\left[cx - \frac{x^2}{2}\right]_b^c$$
$$= \frac{1}{b-a}\left[\frac{b^2}{2} - ab - \frac{a^2}{2} + a^2\right] + \frac{1}{c-b}\left[c^2 - \frac{c^2}{2} - bc + \frac{b^2}{2}\right]$$
$$= \frac{1}{b-a}\left[\frac{1}{2}(b^2-a^2) - a(b-a)\right] + \frac{1}{c-b}\left[\frac{1}{2}(b^2-c^2) - c(b-c)\right]$$
$$= \frac{1}{b-a}(b-a)\left[\frac{1}{2}(b+a) - a\right] - \frac{1}{c-b}(c-b)\left[\frac{1}{2}(b+c) - c\right]$$
$$= \frac{1}{2}(b-a) - \frac{1}{2}(b-c)$$
$$= \frac{1}{2}(c-a)$$

Dividing the Fuzzy set (8) by $\frac{1}{2}(c-a)$, we get the probability distribution function

$$f(x \mid a,b,c) = \begin{cases} \frac{2(x-a)}{(b-a)(c-a)}, a \le x \le b \\ \frac{2(c-x)}{(c-b)(c-a)}, b \le x \le c \end{cases}$$

5 DOSE ASSESSMENT USING PROBABILITY-POSSIBILITY TRANSFORMATION

5.1 Problem definition

Here, we have considered a case of soil contamination by lead on an ironworks Brownfield in the south of France. Following an on-site investigation revealing the presence of lead in the superficial soil at levels on the order of tens of grams per *kg* of dry soil, a cleanup objective of 300 *mg*/*kg* was established by a consulting company, based on a potential risk assessment, taking into account the most significant exposure pathway and the most sensitive target (direct soil ingestion by children).

5.2 Dose Assessment Model

The mathematical model calculating the quantity D_{lead} absorbed by a child living on the site and exposed via soil ingestion is given by EPA, (1989) [6].

$$D_{lead} = \frac{C_{soil} \times IR_{soil} \times (Fi_{inside} + Fi_{outside}) \times EF \times ED}{Bw \times AT \times 10^6} \qquad ...(9)$$

where, D_{lead} is the absorbed lead dose related to the ingestion of soil (*mg*/[*kg*:*day*]), C_{soil} is the lead concentration in Soil (*mg*/*kg*),IR_{soil} is the ingestion Rate *mg* soil/*day*, Fi_{indoor} is the indoor Fraction of contaminated soil ingestion (unitless), $Fi_{outdoor}$ is the outdoor Fraction of contaminated soil ingestion (unitless), *E F* is the exposure Frequency (*days*/*year*), *ED*is the exposure Duration (*years*), *Bw* is the body Weight (*kg*), *AT* is the Averaging time (period over which exposure is averaged–days)

5.3 Results and discussion

Here, we have considered three scenarios. In scenario1, representations of some model parameters are probabilistic while some are probabilistic. In scenario2, possibilistic model parameters are transformed into probabilistic mode while in scenario3, probabilistic model parameters are transformed into possibilistic mode.

5.3.1 ***Scenario1***

Here, representations of the parameters C_{soil} as well as Fi_{indoor} are possibilistic and IR_{soil} and *Bw* are probabilistic and other parameters are deterministic.

Table 1 Parameter values used in the dose assessment

Variable	Mode of Representation	Value
C_{soil}	Fuzzy	TFN(40,300,500)
IR_{soil}	Probabilistic	Triangular(20,160,300)
Fi_{indoor}	Fuzzy	TFN(0.2,0.55,0.9)
$Fi_{outdoor}$	Deterministic	1.0
EF	Deterministic	273.75
ED	Deterministic	6
AT	Deterministic	6×365 = 2190
Bw	Probabilistic	Normal(17.2,2.57)

The graphical representation of the result of the calculation D_{lead} using hybrid method [4] is given in figure 1.

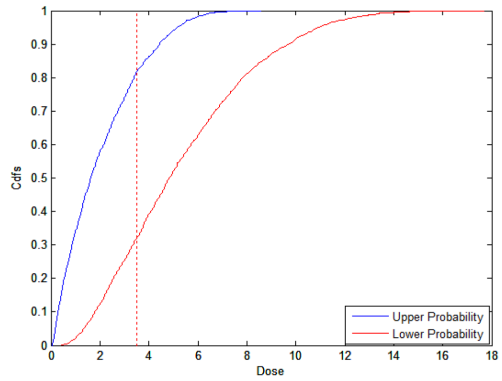

Fig. 1: Plot of upper and lower probability function for estimating *dose*

According to the World Health Organization prescribed the acceptable lead dose related to the ingestion of polluted soil to be equal to 3.5 μg/[*kg*:*day*]. That means that after the cleanup objective of 300 *mg*/*kg* on the site (ironworks Brownfield), calculated doses D_{lead} should not be larger than 3.5 μg/[*kg*:*day*]. From the figure1 it is clear that probability that for *dose value less than 3.5 μg/[kg:day]*, we have probability comprised between 0.31 (lower probability or belief) and 0.81 (upper probability or plausibility). So, consideration of imprecision regarding input parameters leads to a clearer rejection of the proposition D_{lead} < 3.5 µg/[*kg*:*day*]. The gap between the lower probability and upper probability reflects the imprecision of some model parameters.

5.3.2 *Scenario2*

In this scenario, possibilistic model parameters are transformed to probabilistic mode (section 4). Parameter values used in this calculation are given in the following table 2.

Table 2: Parameter values used in the dose assessment

Variable	Mode of Representation	Value
C_{soil}	Probabilistic	Triangular(40,300,500)
IR_{soil}	Probabilistic	Triangular(20,160,300)
Fi_{indoor}	Probabilistic	Triangular(0.2,0.55,0.9)
$Fi_{outdoor}$	Deterministic	1.0
EF	Deterministic	273.75
ED	Deterministic	6
AT	Deterministic	6×365 = 2190
Bw	Probabilistic	Normal(17.2,2.57)

The graphical representation of the result of the calculation D_{lead} using proposed method (previous chapter) of scenario 2 is given in figure 2.

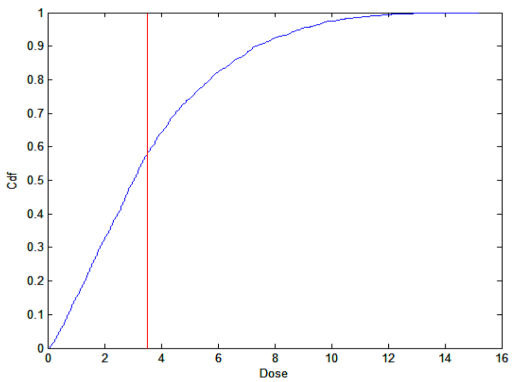

Fig. 2: Plot of probability function for estimating *dose*

Here, we have seen from the figure 2 that probability of *dose value less than 3.5 µg/[kg:day]*, is 0.576 which is not enough to consider to be a low dose i.e., to accept the proposition D_{lead} < 3.5 µg/[*kg*:*day*].

5.3.3 *Scenario3*

In this scenario, probabilistic model parameters are transformed to possibilistic mode (section 3). Representation the parameter body weight (Bw) is a Gaussian fuzzy number with mean 17.2 and standard deviation is 2.57. Parameter values used in this calculation are given in the following table 3.

Table 3: Parameter values used in the dose assessment

Variable	Mode of Representation	Value
C_{soil}	Fuzzy	Triangular(40,300,500)
IR_{soil}	Fuzzy	Triangular(20,160,300)
Fi_{indoor}	Fuzzy	Triangular(0.2,0.55,0.9)
$Fi_{outdoor}$	Deterministic	1.0
EF	Deterministic	273.75
ED	Deterministic	6
AT	Deterministic	6×365 = 2190
Bw	Fuzzy	Gaussian(17.2,2.57)

Here, we will relax the convexity condition to combine Gaussian fuzzy number and triangular fuzzy number using α -cut by considering that α corresponds to the α -cut defined in the interval [0.01, 1]. It will not affect the uncertainty involved in the fuzzy number.

The result of the dose assessment of scenario 3 is depicted in figure 3.

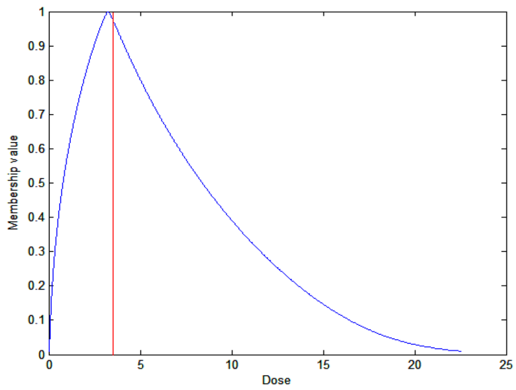

Fig. 3: Plot of membership function for estimating *dose*

The resulting dose is also obtained in the form of a fuzzy number as some input parameters are available in the form of fuzzy number. 3.244 is the core of the output fuzzy dose while the range is [0.03292, 22.5]. Here, the threshold value 3.5 µg/[*kg*:*day*] slightly exceeds the core value with membership value 0.976 and it belongs to the range [0.03292, 22.5]. Here, also consideration of imprecision regarding input parameters leads to a clearer rejection of the proposition D_{lead} < 3.5 µg/[*kg*:*day*].

6. CONCLUSION

Basically, transforming probabilistic data to possibilistic data is useful when weak source of information make probabilistic data unrealistic. Also, it is useful in order to explore the advantages of possibilistic theory at combination steps, or perhaps to reduce the complexity of the solution when computing with possibility values rather than with probability values. Transforming from possibility to probability may be meaningful in the case of decision making where a precise outcome is often preferred, such that, the decision maker is interested to know "what is likely to happen in future", instead of "what is possible in future". The motivation for study of probability-possibility transformations arises not only from a desire to comprehend the relationship between the two theories of uncertainty, but also for some practical problems. For **e**xample: to construct a membership grade function of a fuzzy set from statistical data, to construct a probability measure from a given possibility measure in the context of decision making or system modeling, to combine probabilistic and possibilistic information in expert systems, or to transform probabilities to possibilities to reduce computational complexity. To deal with these problems, various probability-possibility transformations satisfying different consistency principles have been suggested in the literature. Here, we have considered a case of soil contamination by lead on an ironworks Brownfield in the south of France. Following an on-site investigation revealing the presence of lead in the superficial soil at levels on the order of tens of grams per *kg* of dry soil, a cleanup objective of 300 *mg*/*kg* was established by a consulting company, based on a potential risk assessment, taking into account the most significant exposure pathway and the most sensitive target (direct soil ingestion by children). The assessment is carried out by considering there scenarios. In scenario1, representations of some model parameters are probabilistic while some are probabilistic. In scenario2, possibilistic model parameters are transformed into probabilistic mode while in scenario3, probabilistic model parameters are transformed into possibilistic mode. In each scenario, we have seen the clear rejection of the World Health Organization prescribed the acceptable lead dose related to the ingestion of polluted soil to be equal to 3.5 µ*g*/[*kg*:*day*].

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.

REFERENCES

[1] Ali T., Dutta P., Boruah H., (2011), Comparison of a Case Study of Uncertainty Propagation using Possibility - Probability Transformation, International Journal of Computer application, 35:47-55.

[2] Dubois D., Prade H., Sandri S. (1993): On possibility/ probability transformations, in Fuzzy Logic: State of the Art,R. Lowen and M. Lowen, Eds. Boston, MA: Kluwer,pp. 103-112.

[3] Dubois D., Foulloy L., Mauris G., Prade H. (2004): Probability-Possibility Transformation, Triangular Fuzzy Sets and Probabilistic Inequalities, in ReliableComput., 10:273-297

[4] Dutta P., Ali T., (2012), A Hybrid Method to deal with aleatory and epistemic uncertainty in risk assessment, International Journal of Computer Applications, 42 :37-44.

[5] Elena Castineira, Susana Cubillo, Enric Trillas (2007): On the Coherence between Probability and Possibility Measures, International Journal “Information & Applications” 14:303-310.

[6]. EPA, U.S. (1989) Risk Assessment Guidance for Superfund. Volume I: Human Health Evaluation Manual (Part A). (EPA/540/1-89/002).O_ce of Emergency and Remedial Response. U.S EPA., Washington D.C.

[7] Mouchaweh M.S., Bouguelid M.S., Billaudel P., Riera B. (2006): Variable Probability-possibility Transformation, 25th European Annual Conference on Human Decision-Making and Manual Control (EAM'06), September 27-29, Valenciennes, France.

[8] Oussalah, M. (2000): On the probability /possibility transformations: a comparative analysis, Int. journal of General system, 29:671-718

[9] Pacheco M. A.C., Vellasco M .B. R. (2009): Intelligent Systems in Oil Field Development under Uncertainty. (Springer-Verlag, Berlin Heidelberg.)

[10]Yamada K. (2001): Probability –Possibility Transformation based on Evidence Theory. IEEE, pp (70-76)