

Can Fuzzy Logic Via Computing with Words Bring Complex Environmental Issues into Focus?

Ashok Deshpande and Jyoti Yadav

Abstract Information on the status and changing trends in environmental quality is necessary to formulate sound public policy and efficient implementation of environmental pollution abatement programmes. In this quest, water/air quality indices are computed using US-EPA and US-NSF proposed methods for local and regional water/air quality management in many metro cities of the world. There are different types of uncertainties while adopting the procedure in vogue in the computation of these indices. However, it does not include expert's knowledge with a view to arrive at cause–effect relationship. We believe that the development of a method to quantify association between the pollutant and air/water-borne diseases is an important step before classifying air/water quality, either in numeric or linguistic terms. There exists aleatory uncertainty in the pollution parametric data and epistemic uncertainty in describing the pollutants by the domain experts in linguistic terms such as *poor*, *good*, and *very good*. Successes of probability theory have high visibility. But what is not widely recognised is that these successes mask a fundamental limitation—the inability to operate on what may be called perception-based information. In this chapter, we describe the case study 1 that relates to fuzzy description of river water quality in River Ganga for bathing purpose, while case study 2 presents fuzzy description of air quality in Pune City.

Keywords Bathing • River water quality • Fuzzy set theory • Linguistic terms • Fuzzy number • Degree of match • Fuzzy rule base system • Degree of certainty

A. Deshpande (✉)

Berkeley Initiative in Soft Computing (BISC)-Special Interest Group (SIG)-Environment Management Systems (EMS), University of California, Berkeley, CA, USA

College of Engineering Pune (COEP), Pune, Maharashtra, India

National Environmental Engineering Research Institute (NEERI), Nagpur, India

e-mail: ashok_deshpande@hotmail.com

J. Yadav

Department of Computer Science, University of Pune, Pune, Maharashtra, India

1 Introduction

You and a friend walk outside on January morning in Pune City. You announce that the weather is *mild*. Your friend declares that it is *cold*. Who is wrong? Or are you both right?

People recognise that language can be imprecise/fuzzy and that concepts such as *cold*, *hot*, or *mild* do not have well-defined boundaries. In 1965, Prof. Lotfi Zadeh introduced fuzzy sets and thereafter fuzzy logic, a means of processing data by extending classical set theory to handle partial membership. In everyday life and in fields such as environmental health, people deal with concepts that involve factors that defy classification into crisp sets—*safe*, *harmful*, *acceptable*, *unacceptable*, and so on. A classic example is a regulator carefully explaining the result of a detailed quantitative risk assessment to a community group, only to be asked over and over again, *But are we safe?* In this case, *safe* defies crisp classification because it is a multivariate state with gradations that vary among different individuals and groups.

Information on the status and changing trends in environmental quality is necessary to formulate sound public policy and efficient implementation of environmental pollution abatement programmes. One of the ways of communicating the information to the policy makers and public at large is with indices. In the computation of air/water quality index (AQI/WQI), first a numerical value is computed and then the air/water quality is described in linguistic terms. There exists aleatory uncertainty in the pollution parametric data and epistemic uncertainty in describing the pollutants by the domain experts in linguistic terms such as *poor*, *good*, and *very good*. Successes of probability theory have high visibility. But what is unrecognised is that these successes mask a fundamental limitation—the inability to operate on what may be called perception-based information. In this chapter, an attempt has been made to use fuzzy-logic-based formalism in modelling these two types of uncertainties, thereby straightway describing air/water quality in linguistic terms with a degree of certainty attached to each term.

The rest of this chapter is organised as follows: Sect. 2 is a brief account of the theoretical foundation of fuzzy-logic-based method with brief description of the other mathematical framework used. While Sect. 3 relates to a case study for describing water quality, fuzzily concluding remarks and future research efforts are covered in Sect. 4.

2 Fuzzy-Logic-Based Formalism

Will I suffer from water-borne diseases (WBD) if I take a bath in polluted river water? Realising the complexity in establishing cause–effect relationship between bathing in polluted river and water-borne diseases (WBDs), an attempt has been made to present a useful method to address the issue. Some of the important facets of our approach include interviewing student community (bather/non-bather) with a

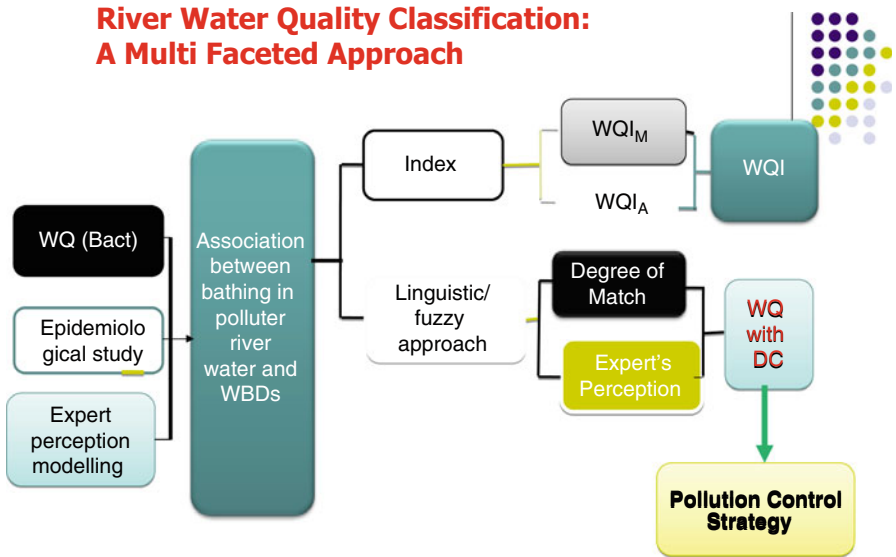


Fig. 1 Multifaceted formalism for water quality classification

structured questionnaire, collecting information on belief of the resident medical practitioners about bathing in polluted river and WBDs, and, furthermore, modelling of epistemic uncertainty in domain expert’s belief in supporting their evidence for various WBDs and the like. Figure 1 presents a novel multifaceted formalism for straightway describing river water quality in linguistic terms with degree of certainty. The technique used in estimating the possible association between bathing in polluted river water and water-borne disease includes epidemiological study including case control study, river water quality analysis, perception of the resident medical professionals regarding their belief in relation to water-borne diseases, Dempster–Shafer (DS) theory of evidence, bootstrapping along with conventional statistical techniques, and the like. Some of these methods are briefly described in this section.

2.1 Fuzzy Measures and Evidence Theory

A fuzzy measure describes the vagueness or imprecision in the assignment of an element a to two or more crisp sets. In a fuzzy measure, the concern of attention is to describe the vagueness or imprecision in assigning the point to any of the crisp sets on the power set. Shafer developed Dempster’s work and presented an important theory of evidence called Dempster–Shafer (DS) theory in which DS belief (Bel) and plausibility (Pl) are used to characterise uncertainty. A basic measure in DS theory is a basic belief assignment (BBA). The function (m) is a mapping function to express BBA for a given evidential event A , $m(A)$. BBA is a representation of partial belief not only for a single possible event but also for a set of

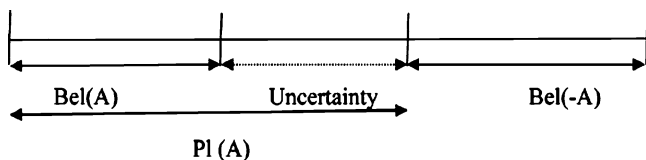


Fig. 2 Belief and uncertainty and ignorance

possible events with imprecise and incomplete knowledge and data. The main concept of evidence theory is that our knowledge of a given problem can be inherently imprecise. Hence, the bound result, which consists of both belief and plausibility, is presented (Fig. 2). BBA expresses the degree of belief in a proposition. BBA is assigned by making use of a mapping function (m) to express our belief with a number in the unit interval $[0, 1]$

$$m : 2^x \rightarrow [0, 1]. \tag{1}$$

The number $m(A)$ represents the portion of total belief assigned exactly to proposition A . The measure m , BBA function, must satisfy the following axioms:

$$m(A) \geq 0 \quad \text{for any } A \in 2^x \tag{2}$$

$$m(\varphi) = 0, \quad \sum_{A \in 2^x} mA = 1 \tag{3}$$

Though these axioms of evidence theory look similar to those of probability theory, the axioms for the BBA functions are less stringent than those for probability measure.

Dempster’s Rule of Combining. The information from different sources can be aggregated by Dempster’s rule of combining to make a new combined BBA structure as given in the following equation:

$$M_{12}(A) = \frac{\sum_{c_i \cap c_j = A} m_1(c_i)m_2(c_j)}{1 - \sum_{c_i \cap c_j = \phi} m_1(c_i)m_2(c_j)}, \quad A \neq \phi \tag{4}$$

where C_i and C_j are propositions from each sources (m_1 and m_2). In Eq. (4), $\sum_{c_i \cap c_j = \phi} m_1(c_i)m_2(c_j)$ can be viewed as contradiction or conflict among the information given by the independent knowledge sources (Ross 1997). Even when some conflict is found among the information, Dempster’s rule disregards every contradiction by normalising with the complementary degree of contradiction to consider only consistent information. However, this normalisation can cause a counterintuitive and numerically unstable combination of information when the given information from different sources has significant contradiction or conflict. If there is a serious conflict, it is recommended to investigate the given information or to collect more information.

2.2 Belief and Plausibility Function

Owing to lack of information and various possibilities in constructing BBA structure, it is more reasonable to present a bound of the total degree of belief in a proposition, as opposed to a single value of probability given as a final result in probability theory. The total degree of belief in a proposition A is expressed within bound $[\text{bel}(A), \text{pl}(A)]$, which lies in the unit interval $[0, 1]$ as shown in Fig. 1, where $\text{Bel}(A)$ and $\text{Pl}(A)$ are given as

$$\begin{aligned} \text{Bel}(A) &= \sum_{c \subset A} m(C) : \text{Belief function} \\ \text{Pl}(A) &= \sum_{c \subset A \neq \phi} m(c) : \text{Plausibility function.} \end{aligned} \tag{5}$$

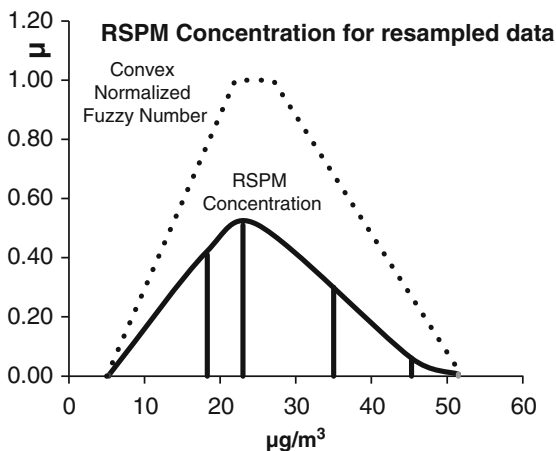
$\text{Bel}(A)$ is obtained by the summation of BBAs for proposition, which is included in proposition A fully. $\text{Bel}(A)$ is the total degree of belief. The degree of plausibility $\text{Pl}(A)$ is calculated by adding BBAs of propositions whose intersection with proposition A is not an empty set. That is, every proposition consistent with proposition A at least partially is considered to imply proposition A because BBA in a proposition is not divided into its subsets. Briefly, $\text{Bel}(A)$ is obtained by adding the BBAs of propositions that totally agree with the proposition A as a measure of belief, whereas $\text{Pl}(A)$ plausibility is calculated by adding BBAs of propositions that correspond to the proposition A totally or partially. In a sense, these two measures consist of lower and upper probability bounds.

2.3 Fuzzy Inference System

Firstly, water/air quality experts are identified, and relevant field data is collected. Additional data generation is a logical step if the available data is inadequate for analysis. Perception of experts about the linguistic description of river water quality for bathing is obtained on interviewing or through a questionnaire. Modelling of uncertainty in the expert’s perception by constructing fuzzy sets/fuzzy numbers and the uncertainty in the field data of water quality parameters using the concept of convex normalised fuzzy number is the next step. The parameters identified for defining bathing, say water quality by the experts, are faecal coliforms (FC), dissolved oxygen (DO), biochemical oxygen demand (BOD), pH , and *turbidity*. The relevant parameters could be considered while describing the overall air quality of a city/region.

Randomness in the air/water quality data can be transformed into a convex normalised fuzzy number A with membership grade function $\mu_A(x)$, thereby characterising the dynamic behaviour of the water quality parameters. We refer to

Fig. 3 Probability and possibility distribution for respirable suspended particulate matter (RSPM)



air quality parameter for illustration. If x_i is some point on the parametric domain for which $p(x_i)$ is maximum, then define function $\mu_A(x)$ (Fig. 3):

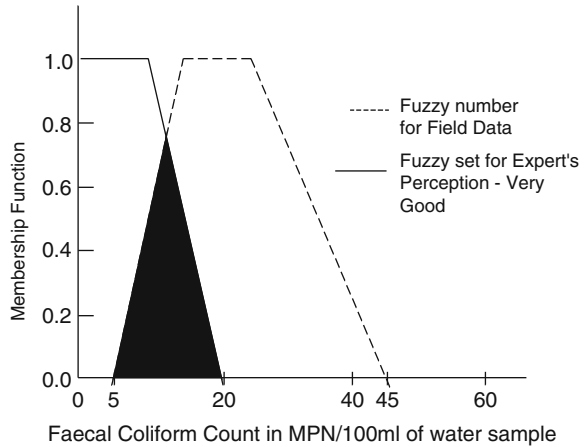
$$\mu_A(x) = p(x)/p(x_i) \tag{6}$$

Construction of fuzzy number or fuzzy sets for modelling the perception of the experts in classifying each parametric domain in linguistic terms such as *very good and good* allows for referencing all possible parametric values to be described. This transforms a random variable into a convex normalised fuzzy number A with membership grade function $\mu_A(x)$, thereby characterising the dynamic behaviour of the water quality parameter. The construction of fuzzy number or fuzzy sets for modelling the perception of the experts in classifying each parametric domain linguistically involves selection of linguistic terms such as *very good and good*, which allows for referencing all possible parametric values to be described; classification of the parametric domain and assigning linguistic terms to each class linearly by the experts reflecting the imprecision in their perception; the set of values for which all the experts assign the same linguistic term is given $\mu = 1.0$, while none of the experts assigning that term are given $\mu = 0.0$. The breakeven point membership grades 0.0 and 1.0 are connected by continuous monotonic function which presupposes that the degree of consensus among the experts goes on increasing as the parametric values approach the core of fuzzy number for the specified linguistic term.

2.4 Matching Between Two Fuzzy Values

The fuzzy number for field data (A) on parameters and the fuzzy number characterising linguistic terms (A') are matched together to arrive at a measure called degree of match (DM) defined by

Fig. 4 Fuzzy numbers for very good faecal coliforms



$$DM_{ff}(A, A') = \int \mu_{A \cap A'}(x) dx / \int \mu_A(x) dx, x \subset X \tag{7}$$

in which X denotes the universe and $\mu_{\zeta \cap \zeta'}(x)$ is membership grade for $\zeta \cap \zeta'$. Furthermore, if A and A' are the discrete possibility distributions, the measure is defined as

$$DM_{ff}(A, A') = \sum \mu_{A \cap A'}(x) / \sum \mu_A(x), x \subset X \tag{8}$$

Figure 4 shows the fuzzy number for *very good faecal coliforms* reveals that almost all the experts agree that the *faecal coliforms* count between 0 and 10 MPN/100 ml of water. The level of presumption or membership function decreases with the increasing faecal coliforms count. When the count exceeds 20, none of the experts define the parameters as *very good* for bathing purpose. This is indicated by the level of presumption $\lambda = 0$.

A set of rules is constructed for classifying air/water quality as highly acceptable, just acceptable, or not acceptable (rejected) in order to aggregate the set of attributes. Each rule has antecedent propositions connected together using AND operator, resulting in some consequences. The assertions related to its antecedent part are obtained from the users, which are imprecise or fuzzy. Thus, a fuzzy rule-based system can be developed for the knowledge representation or reasoning process. Here, the partial matching is allowed, and the analyst can estimate the extent to which the assertion satisfies the antecedent part of the rule contrary to the rule-based system which examines as to whether the antecedent part is satisfied or not [1]. A hierarchical structure for water classification resulting in a set of rules can be constructed (Fig. 5). The chemical status of water is judged in the first hierarchical level of knowledge base. The second hierarchical level characterises bacteriological, chemical, and physical status of water to arrive at the ultimate acceptable strategy of water quality for bathing purpose. If need be, a similar structure can be developed for air quality classification.

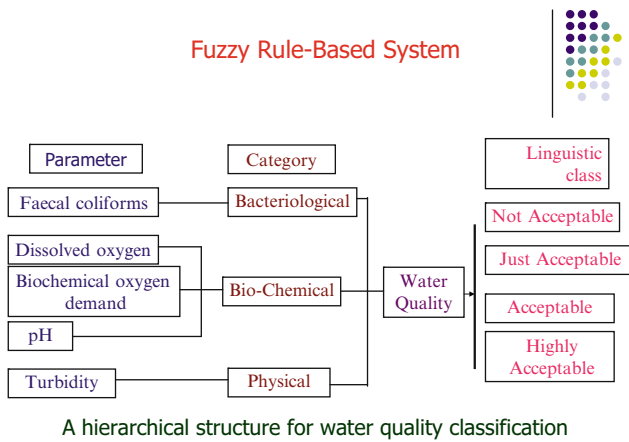


Fig. 5 Type 1 fuzzy inference system/fuzzy expert system

Following are the sample rules stored at two different hierarchical levels of the knowledge base:

Rule 1

*If DO is <fair> and
BOD is <good> and
pH is <very good>,
then chemical status of water is <good>.*

3 Fuzzy Description of Environment Quality

With the approach presented in this chapter, it is possible to describe any environmental quality fuzzily. We present herein water and air quality.

3.1 Case Study 1: Fuzzy Description of River Water Quality

Around 21% of communicable diseases in India are water related, and out of these, diarrheal diseases alone killed over 700,000 Indians in the year 1999. Since time immemorial, sacred bathing on the holy riverbanks (*ghats*) is practised in India and Nepal. It has been a religious belief of some of the bathers that gulping water at these holy places washes away their sins! The first study, therefore, relates to establishing the cause–effect relationship between bathing in polluted river water and water-borne diseases carried out at Alandi town near Pune situated on the

Table 1 Fuzzy description of river water quality with degree of certainty

Degree of certainty		
Water quality description	Rishikesh	Varanasi
Highly accepted	0.21	0
Accepted	0.41	0
Just accepted	0.36	0.01
Not accepted	0.1	0.98

banks of river Indrayani. We very briefly discuss the final outcome of such an investigation [2]. The analysis of the results reveals result that the combined belief of the two domain experts in identifying for the single disease *diarrhoea is 0.58*. In sum, bathing in polluted water can be a cause for the incidence of water-borne diseases in bathers [2].

Describing Ganga water quality straightway in linguistic terms, with some degree of certainty attached to each linguistic term, concludes that the Government of India should continue the efforts on the reduction of pollution levels, especially from bacteriological standpoint. Aerobic treatment is the option that should be attached great importance in future planning of GAP 2. The results depicted after Ganga Action Plan 1 infers that even at Rishikesh, the water quality of Ganga is not very good for bathing and still the pollution persists as the degree of certainty attached to fair is 0.36 while it is 0.41 for the linguistic description of water as good. Varanasi has been a serious cause of concern from the standpoint of bathing in the holy river (Table 1).

3.2 Case Study 2: Fuzzy Description of River Air Quality

The case study relates to fuzzy air quality description with the available air quality data from five monitoring stations in Pune City. These are Pimpri-Chinchwad Municipal Corporation, Karve Road, Swargate, Bhosari, and Nal Stop. In order to bring down progressive deterioration in air quality, the Government of India has enacted Air (Prevention and Control of Pollution) Act in 1981, and further stringent guidelines are promulgated in Environment (Protection) Act, 1986. The need for ambient air quality monitoring programme that is needed to determine the existing quality of air and evaluation of the effectiveness of control programme and to develop new programme was recognised. As a follow-up action, the Central Pollution Control Board (CPCB) initiated National Air Quality Monitoring (NAQM) Programme during 1984–1985 at the national level.

A well-structured air quality monitoring network involves selection of pollutants, selection of locations, frequency, duration of sampling, sampling techniques, infrastructural facilities, man power and operation, maintenance costs, and the like. The network design also depends upon the type of pollutants in the atmosphere through various common sources, called common urban air

Table 2 Comparison of conventional AQI and fuzzy description of air quality with degree of certainty

Monitoring station	Conventional AQI	AQI with degree of certainty
Karve Road	206 very poor	0.45 very poor
Bhosari	145.89 poor	0.43 very poor
Swargate	147.13 poor	0.54 poor
Nal Stop	120.2 poor	0.67 fair
PCMC	151.13 very poor	0.91 poor

pollutants, such as suspended particulate matter (SPM), respirable suspended particulate matter (RSPM), sulphur dioxide (SO₂), oxides of nitrogen (NO_x), and carbon monoxide (CO). The areas to be chosen primarily are such areas which represent high traffic density, industrial growth, human population and its distribution, emission source, public complaints if any, the land use pattern, etc. Generally, most of the times the basis of a network design is the pollution source and the pollution present.

Generation of fuzzy numbers for different linguistic hedges (*very good*, *good*, *fair*, *poor*, *very poor*) of RSPM concentration is an important issue in any FIS. According to the expert, RSPM count between 18–57 µg/m³ is very good, 42–75 µg/m³ is good, 67–100 µg/m³ is fair, 90–130 µg/m³ is poor µg/m³, and above 120 µg/m³ it is very poor. The level of membership function decreases with the increasing RSPM level. When it exceeds 57, the expert does not define the parameter as very good. Table 2 is the comparison between the computed AQI and the proposed fuzzy-logic-based method. It can be revealed that AQI based on the traditional method does not attach any certainty while describing the air quality. In addition, the method does not consider the aggregated effect, and the highest computed AQI is considered as the final decision on the air quality which, in our opinion, is the departure from human thinking. Alternately, using type 1 fuzzy inference system, we can describe the air quality straightway in linguistic terms with some degree of certainty attached to each term.

4 Outlook in Environmental Policy [3]

Over the past few decades, soft computing tools such as fuzzy-logic-based methods, neural networks, and genetic algorithms have had significant and growing impacts. But we have seen only limited use of these methods in environmental fields, such as risk assessment, cost-benefit analysis, and life-cycle impact assessment. Because fuzzy methods offer both new opportunities and unforeseen problems relative to current methods, it is difficult to determine how much impact such methods will have on environmental policies in the coming decades. Here, we consider some obvious advantages and limitations.

Quantitative models with explicit and crisp delineations of systems have long been the currency of discourse in engineering and the physical sciences, where basic physical laws form the foundations of analyses. These fields place high value on the causal linkages implicit in model structure and parameterization. But for problems that involve human values, language, control theory, biology, and even environmental systems, researchers have had to rely more on descriptive and empirical approaches. When the goal is to summarise the observations in an efficient and useful manner, fuzzy-logic-based methods should be further investigated as alternative—and perhaps more appropriate—methods for addressing uncertain and complex systems. For the types of complex and imprecise problems that arise in environmental policy, the ability to model complex behaviours as a collection of simple if–then rules makes fuzzy logic an appropriate modelling tool. Because fuzzy arithmetic works well for addressing linguistic variables and poorly characterised parameters, fuzzy methods offer the opportunity to evaluate and communicate assessments on the basis of linguistic terms that could possibly match those of decision makers and the public. Moreover, approximate reasoning methods such as fuzzy arithmetic do not require well-characterised statistical distributions as inputs. Another key advantage of fuzzy logic in risk assessment is the ability to merge multiple objectives with different values and meanings, for example, combining health objectives with aesthetic objectives. It also provides rules for combining qualitative and quantitative objectives [3].

But fuzzy logic has at least two limitations for expressing health risks and other environmental impacts. One problem is its strong reliance on subjective inputs. Although this is a problem in any type of assessment, fuzzy methods might provide more opportunity for the misuse of subjective inputs. Although probabilistic assessments based on tools such as Monte Carlo methods are analogous to assessments based on fuzzy logic, these two techniques differ significantly both in approach and in interpretation of results. Fuzzy logic confronts linguistic variables such as ‘safe’, ‘hazardous’, ‘acceptable’, and ‘unacceptable’, whereas Monte Carlo methods are forced to fit linguistic variables for probabilistic assessments. Fuzzy arithmetic combines outcomes from different sets in a way that is analogous to but still different from Monte Carlo methods. Possibility theory can be used as an alternative to probabilistic analysis, but this strategy creates the potential for misuse if membership functions are interpreted as probability distributions.

4.1 No More Crisp Lines?

Fuzzy logic represents a significant change in both the approach to and the outcome of environmental evaluations. Currently, risk assessment implicitly assumes that probability theory provides the necessary and sufficient tools for dealing with uncertainty and variability. The key advantage of fuzzy methods is how they reflect the human mind in its remarkable ability to store and process information that is consistently imprecise, uncertain, and resistant to classification. Our case

study illustrates the ability of fuzzy logic to integrate statistical measurements with imprecise health goals. But we submit that fuzzy logic and probability theory are complementary and not competitive. In the world of soft computing, fuzzy logic has been widely used and has often been the ‘smart’ behind smart machines. But more effort and further case studies will be required to establish its niche in risk assessment and other types of impact assessment. Could we adapt to a system that relaxes ‘crisp lines’ and sharp demarcations to fuzzy gradations? Would decision makers and the public accept expressions of water- or air-quality goals in linguistic terms with computed degrees of certainty? Resistance is likely. In many regions, such as the United States and EU, both decision makers and members of the public seem more comfortable with the current system—in which government agencies avoid confronting uncertainties by setting guidelines that are crisp and often fail to communicate uncertainty. Perhaps someday, a more comprehensive approach that includes exposure surveys, toxicological data, and epidemiological studies coupled with fuzzy modelling will go a long way towards resolving some of the conflict, divisiveness, and controversy in the current regulatory paradigm.

Acknowledgement The wholehearted assistance received from Dr. D. V. Raje and Dr. Kedar Rijal for the implementation of the concept is gratefully acknowledged. My special thanks to Professor Thomas McKone who helped the author in many fuzzy ways!

References

1. Deshpande AW, Raje DV, Khanna P (1996). Fuzzy description of river water quality. *Eufit* 96, 2–5 September. pp 1795–1801
2. Rijal K, Deshpande A, Ghole V (2009) Bathing in polluted rivers, water-borne diseases, and fuzzy measures: a case study in India. *Int. J. Environ Waste Manage* 6(3–4):255–263
3. Mckone TE, Deshpande AW (2005) Can fuzzy logic bring complex environmental problems into focus? *Int J Environ Sci Technol* 39(2):42A–47A