Performance assessment based on evidence theory and fuzzy logic: application to building and dam performance

Aurélie Talon, Corinne Curt, Daniel Boissier

Performance assessment based on evidence theory and fuzzy logic: application to building and dam performance

Aurélie Talon[1], Corinne Curt[2], Daniel Boissier[3]

Abstract

In the current economic, sociological and environmental context, the durable performance of civil engineering works is a keystone of risk management. The data available for assessing performance are often numerous, heterogeneous and imperfect. To cope with these difficulties, a four-phase method is proposed, comprising: (i) performance assessment modelling, (ii) the formalisation of heterogeneous information from different sources, (iii) the propagation of imperfections in performance assessment models, and (iv) the expression of outputs of these models in suitable formats to facilitate decision-making and improve communication. Two applications of this methodology have been developed for buildings and dams and are used as illustrations. This methodology is supported by various methods that must be chosen as a function of the case studied. Systemic approaches, knowledge-based methods, evidence theory, and fuzzy logic were used in the applications.

CE Database subject headings: data aggregation, dam, building, fuzzy logic, multi-scale data, data unification.

[1] Clermont-Universités – Polytech Clermont-Ferrand – Institut Pascal UMR 6602 – 24 avenue des Landais – BP 206 – 63174 AUBIERE Cedex – France (corresponding author). E-mail: aurelie.talon@polytech.univ-bpclermont.fr

[2] Irstea – 3275 Route de Cézanne – CS 40061 – 13182 Aix-en-Provence Cedex 5 – France. E-mail: corinne.curt@irstea.fr

[3] Clermont-Universités – Polytech Clermont-Ferrand – Institut Pascal UMR 6602 – 24 avenue des Landais – BP 206 – 63174 AUBIERE Cedex – France (corresponding author). E-mail: daniel.boissier@polytech.univ-bpclermont.fr

Introduction

In the current economic, sociological and environmental context, controlling the performance of civil engineering works is of particular interest for risk management (MRGenCi 2009; Modarres 1993) and sustainable development (United Nations 2009). The performance of a civil engineering structure mirrors its in-service, global and functional behaviour. The entire set of mechanical resistance and usage functions fulfilled by a civil engineering structure are taken into account as well as the relations between the structure and its climatic environment (for example, rain, wind, temperature) and usage characteristics (owner, third parties). The objective of controlling performance relies on the development of methods and tools to assess performance at different dates of service life. In building applications, performance assessment requires the estimated service life of building components such as windows, floors and roofs during the design phase, using the whole set of available service life data. These data are multi-scale as they can be assessed at the levels of materials, components and phenomena. In dam applications, performance assessment concerns the different functions (for example sealing, drainage) and performance linked to failure modes (for example, internal erosion through the embankment, overtopping).

The data used to assess performance can have several origins, thus a single item of information is generally insufficient when carrying out a global performance assessment (Lee and Barrett 2003). Information is frequently provided by multiple and heterogeneous sources (for example, visual observations, results from numerical models, instrumental data), with very different levels of granularity. Moreover, these data have proven to be imprecise and uncertain. It is essential to represent and propagate this imperfect data in the performance assessment model in order to better represent reality. The aim of this article is to propose a general methodology for the performance assessment of civil works for which data is numerous, heterogeneous and imperfect. This methodology is supported by various methods that must be chosen as a function of the case studied.

The current main approaches and methods for assessing the performance of a civil structure are the following:

- Physical modelling based on theoretical modelling, numerical simulation and experiments, such as those comparing long-term and short term exposure tests (Jernberg *et al*. 2004);
- Methods based on the probabilistic approach (Siemes *et al*. 1985);
- Statistical approaches (Brandt and Wittchen 1999; Foster *et al*. 2000; Kreuzer 2000);
- Artificial intelligence techniques such as feedback from practice and knowledge-based methods (Andersen *et al*. 2001; Curt *et al*. 2010), neural networks (Kim and Kim, 2008; Wu and Chau, 2006) that can be used in combination with optimisation methods such as genetic algorithms (Cheng et al., 2002);
- Systemic methods (Peyras *et al*. 2006; Serre *et al*. 2007; Lair 2000; Talon *et al*. 2008; Baroth *et al*. 2011).

Methods capable of modelling and aggregating heterogeneous and possibly multi-source data must be used. Systemic and knowledge-based methods are also relevant and they are the two types of methods we focus on in this paper.

Several approaches have been used in imprecise and uncertain situations in order to better model the behaviour of civil engineering works through time and then deduce their levels of performance. There are probability approaches, statistical approaches, approaches based on evidence and possibility theories. Probability approaches are well-adapted when the phenomena studied are well-known. Indeed, enough information is needed to clearly define the probability law and the parameters that best fit the given phenomena. Statistical approaches are relevant when a significant sample of data associated with the problem in question is available. Both statistical and probability approaches allow taking into account several random parameters and these approaches can deal with homogeneous data in terms of format and geometrical scale. However, the results obtained are difficult to transpose and extrapolate to other case studies. Approaches based on possibility and evidence theories allow representing all types of data frames and thus taking into account all the available data, whatever their frame or geometrical scale. Possibility theory also allows representing imprecise and uncertain data using fuzzy sets. Evidence theory, when used with merging, provides a consensual result if considering data from different sources. Consequently, these approaches are complementary and well-adapted when few data are available, when they come from different sources and when they are imprecise. However, these approaches require good knowledge of the conditions of acquiring the data used in order to assign a belief mass (confidence level) to them. We use possibility and evidence theories as we consider they are those most relevant to our applications.

This paper presents the use of possibility and evidence theories in the methodology applied to assessing the performance of buildings and dams. Examples based on real cases are given to illustrate the various methods. This paper does not deal with the entire set of methods used. More details concerning them can be found in Curt et al. (2010) on the formalisation and data aggregation phases for dam application, in Curt et al. (2011) on the representation and propagation of imperfections and in Talon et al. (2008) and Talon (2006) on the method used for assessing service life for building applications.

Section 2 presents the methodology proposed for the fuzzy assessment of performance while section 3 describes the models developed to assess performance in the case of buildings (using systemic methods) and dams (using knowledge-base methods). Section 4 focuses on fuzzification and data unification while Section 5 is dedicated to data aggregation and the defuzzification process. Examples are used to illustrate the methodology from an operational viewpoint.

Methodology for the fuzzy assessment of performance

The methodology proposed to assess the performance of civil engineering structures is presented in Figure 1. It comprises 4 phases. The first concerns modeling for performance assessment using heterogeneous information from different sources. As the data are imperfect and come from heterogeneous formats, the second phase is aimed at representing the entire set of data in the same format of fuzzy sets, and at performing data unification using evidence theory when necessary (data from multiple sources). The third phase concerns the propagation of fuzzy data in the performance models. The defuzzification phase ends the process; it involves the extraction of a concise and relevant result from the third phase. Indeed, human decision can be made relevantly on the basis of a limited amount of information.

These four phases are the same for the two applications presented here but the methods used differ and are adapted to the case studied (cf. Figure 1). These methods are described in the following sections.

Performance assessment models

Input data

The data used to assess both building and dam performance comprise different kinds of data traditionally used in the field of civil engineering:

- Visual observations performed in-situ;
- Instrumental measurements. These stem from instruments located on the works ("monitoring data"): for example, flow or piezometry measurements performed on a dam. They can also be obtained during laboratory experiments (permeability, shear strength) or in situ tests during construction (penetrometer tests);
- Results of numerical models. Two types can be distinguished: "calculation data" based on theoretical models (for example, in the case of dams, hydraulic gradient) and simulation results;
- Data related to design and construction practices. These are assessed using a comparison to rules of thumb and aim at verifying that the work was correctly designed and built. For instance, the data "assessment of filter rules" allows verifying the use of a correct grading curve between the drain and the shoulder.

Some of these data can be statistically analyzed when relevant samples are available, i.e. instrumental data collected over several years. Data can be obtained from different sources, notably databases, journals and conferences, simulations, experiments, theoretical models, monitoring sensors, and expert assessments.

Model for performance assessement – Degradation scenarios – Building application

In this paper, the service life assessment of buildings (associated with a given performance) is based on the qualitative identification of the entire set of degradation scenarios (chain of degradation phenomena) that may lead to the failure of a given building component (Talon *et al*. 2008; Talon 2006). Firstly, from a methodological viewpoint, the functioning of a building component and the relations between the elements of the building component and its in-service and usage environment are modeled using structural and functional analysis. Secondly, a failure mode and effects analysis (FMEA) is performed to identify all the possible degradations of this functional model. Thirdly, a criticality analysis of the degradation scenarios is carried out and the probability of these scenarios is assessed to narrow the service life assessment down to the most critical scenarios (combination of their probability, duration and the seriousness of their consequences). The aggregation method used to obtain the duration of the scenario in the building application is the sum of the durations of the phenomena that make up this scenario. The service life of a building component is then assessed by the minimum of the scenario durations.

Figure 2 presents the different phases of this approach. An example of a scenario obtained from the Failure Mode and Effects analysis is given. It concerns the occurrence of 1 mm cracks in a concrete wall due to concrete carbonation. This is caused by the chain of 4 phenomena: biological deposit, increased porosity of external plaster coat, concrete carbonation and decreased weldmesh section. The duration of biological deposit (3 to 4 years) and increased porosity (7 to 10 years) are “by defect” data. The duration of carbonation (26 to 28 years) is obtained by merging two data (26 to 28 years – Data $D_{3\text{-}1}$ – and 25 to 30 years – Data $D_{3\text{-}2}$), as explained in section “data unification”, is obtained from an internal report (using numerical simulation) and a communication in the proceedings of the 8th Conference on the Durability of Building Materials and Components. The duration of decreased section (42 to 44 years) is obtained by combining the data of an internal report (using numerical simulation) and a communication reporting the occurrence of 1 mm cracks in the decreased section for the configuration, shown in Figure 2.

Model for performance assessement – Knowledge-based method – Dam application

In the case of dams, performance assessment concerns the different functions and performance linked to failure modes (Curt *et al*. 2010). These performances are assessed by aggregating different types of data: data from visual observations, monitoring data, calculation data, and design and construction data. All these data are formalized as “indicators”: a grid is proposed to provide a standard format for the information needed by experts. The grid is composed of six fields: name, definition, scale, references such as anchorage points on the scale, and spatial and temporal characteristics. The experts can then make good use of the indicators as they can rely on their repeatability and reproducibility. All the different types of indicators (visual, monitoring, etc.) are described with the same format. An example of slab thickness indicator is presented in Table 1. These indicators are then aggregated with arithmetical operators or IF-THEN rules proposed by a panel of five experienced engineers from Irstea (National Research Institute of Science and Technology for the Environment and Agriculture) trained and coached by an external observer. These experts had already carried out several detailed dam reviews or analyses (one to twenty-five years experience).

Figure 3 presents an example of a hierarchical model for the assessment of internal erosion through an embankment in the case of a dam with an upstream bituminous membrane. Three functions were identified: sealing, drainage and erosion defence functions. Thirteen indicators are necessary to assess dam performance regarding the mechanism of internal erosion through the embankment. Three of them are instrumental (flow/flow change (increase), flow change (decrease), piezometry) and the others are visual (visual state of the drain outlet, wetland at the downstream toe of the dam, etc.). Three indicators are involved in the performance of the sealing function; six are in that of the drainage function and four in that of the erosion defence function. The assessment of the performance function is performed using the maximum operator. For instance, in the case of the erosion protection function (μ ErDef):

[1] $$\mu \text{ ErDef} = \text{MAX}\,[\text{Sinkhole, Differential Settlements, Local and muddy seepage, Hydraulic Gradient}]$$

The assessment of phenomena or dam safety related to a failure mode is carried out using IF-THEN rules such as (case of the insufficiency of drainage capacity phenomenon – φ INSUFDC):

[2] IF "Clean water seepage" > 2 OR "Piezometry" > 2
THEN φ INSUFDC = MAX [Clean water seepage, Piezometry]

[3] IF "Clean water seepage" ≤ 2 AND "Piezometry" ≤ 2 AND μ SEAL ≤ 2
THEN $\varphi_1 = \mu$F SEAL

With µF SEAL = performance of sealing function. Clean water seepage and Piezometry are qualified as “direct indicators”, *i.e.* indicators specific to a phenomenon or a failure mode that assign a direct assessment to it. They are considered by experts as key model parameters.

Indicator definition and aggregation models were validated using simplified and complete real case studies.

Deterioration of dam safety involving a failure mode is due to the deterioration of the whole set of functions implied in this failure mode. For example, dam safety related to internal erosion through the embankment is based on the reliability of three functions: sealing, drainage and erosion protection. Function performance and dam safety are assessed on the same scale “0-10” as the indicators.

Fuzzification and data unification

Fuzzification

The aim of the fuzzification phase is to solve two kinds of problem. The first problem is modelling imperfections while the second problem is representing them in a common format for the entire set of data collected to characterize a phenomenon (duration phenomena for the building application, indicators for the dam application). The data collected can come from, for example, a visual review performed by an expert, from a sounding device, or an in-situ test. Therefore we chose a representation with a format based on the possibility theory as this theory allows transforming every type of format into fuzzy sets (cf. Figure 4).

In this fuzzification phase, the process is quite different for both applications: the possibility function is calculated by a mathematical model (building application) or declared by an expert (dam application) (cf. Figure 1).

For the building application, the transformation of all types of frame into a fuzzy frame is computed using mathematical rules and quality analysis results.

The literature shows that certain models based on evidence theory have been developed in view to transforming several frames into a single frame. For example, the works of Baudrit *et al.* (2007) allow transforming a probability frame into a fuzzy set when the median, mode, average and percentiles are available. If considering: p a probability density, I an interval, m a median, x_α the percentile of α order, $Inf(I)$ the minimum value of I, $Sup(I)$ the maximum value of I, and π a possibility density; then:

[4] $$\begin{gathered} \pi(Inf(I)) = \pi(Sup(I)) = 0 \\ \pi(x_\alpha) = \pi(x_{1-\alpha}) = 2\alpha \\ \pi(m) = 1 \end{gathered}$$

A linear interpolation of π on the $[Inf(I), x_\alpha]$, $[x_\alpha, m]$, $[m, x_{1-\alpha}]$, $[x_{1-\alpha}, Sup(I)]$ intervals is possible.

For other data, a two-step method is applied: first, the determination of the base and core of the fuzzy set and, second, the determination of its height. Indeed, the width of a fuzzy set core and base depend of the expression of the duration data. This can be expressed in four ways: "more than *x* years", "less than *y* years", "*z* years" and "between *y'* and *x'* years". Each duration datum can be described as a sure interval and a possible interval (which includes data imperfection). The sure interval is included in the possible interval. When using a representation as a fuzzy set, the core (cf. [5]) corresponds to the sure interval and the base (cf. [6]) is associated with the possible interval.

[5] $$Core(I) = \{x \in X / \pi(x) = 1\}$$

[6] $$Base(I) = \{x \in X / \pi(x) \neq 0\}$$

The fuzzy set core and base of these four cases are expressed as shown in Table 2. A linear interpolation is performed between the core values and the base values. An example of fuzzification of carbonation durations ($D_{3\text{-}1}$ and $D_{3\text{-}2}$) is provided in Figure 5.

The second step consists in determining the height of the fuzzy set. The height of the fuzzy set is equal to the belief mass deduced from a quality analysis. The belief mass expresses the confidence that can be assigned to the data source. For example, when the confidence is complete the height of the fuzzy set is equal to 1 and when the confidence is slight the height of the fuzzy set is less than 1. A quality analysis is performed and consists in estimating the reliability that can be associated with each datum collected. Here, this reliability is represented by a belief mass which is assessed by using a method derived from the NUSAP method ("Numerical Unit Spread Assessment Pedigree") developed by Funtowicz and Ravetz (1990). NUSAP is a tool for analysing and diagnosing information based on five criteria (numerical, unit, spread, assessment and pedigree) of the analysis of the imperfections (uncertainty, imprecision and incompleteness) that can be associated with the data processed. From a generic viewpoint, the different imperfections can be merged into three categories of criteria. We distinguish: imperfection associated with the intrinsic quality of the raw data, imperfection associated with the data acquisition potential, imperfection associated with the analysis of data processing (cf. Table 3). The criteria must be adapted to the types of data collected. The source criterion estimates the confidence that the scientific community has in this information, whereas the censorship criterion assesses the reliability of the observation period used to provide the raw data considered. The aim of the “type of data modelling”, “modelling hypothesis” and “type of acquisition model” criteria is to assess the theoretical model used to represent the raw data. The “consistency of data preparation” assesses the possible loss of information due to the transformation of the raw data format into a fuzzy set format. The principle of the data collection for the building application is to collect all the data fairly similar to the desired phenomena durations. Consequently, the “correspondence regarding the case study” criterion estimates the deviation (for example, climatic conditions, material, experimental conditions) between the characteristics of the desired phenomenon and the characteristics of the raw data collected. Finally, the entropy criterion assesses the quantity of information given by the data collected for solving the given duration problem. This entropy criterion is calculated using Shannon’s entropy concept (1948). The quality grid thus formulated allows obtaining a homogeneous assessment of data quality and provides a numerical scale (four markers) and the associated linguistic scale for the different data analysis criteria (cf. Table 3). The result of this phase of the methodology is a set of pairs {data; belief mass}. The belief mass corresponds to the average of the scores obtained for the set of quality criteria. Table 3 provides the values, from 0 to 1, that can be taken by the quality criteria and an example of this quality analysis for the duration data $D_{3\text{-}1}$ of carbonation (example of concrete wall in Figure 2).

For the dam application, fuzzification is currently declared directly by the experts who model each indicator by a fuzzy distribution. The experts provide an implicit analysis of the parameters of data quality and transcribe this information into fuzzy sets. For example, in Figure 4 (dam application part), the experts express scores for an indicator as a normalised fuzzy subset. The fuzzy membership function is built considering that the core (4-6 on a scale from 0 to 10) represents the most plausible values while the base (3-8 on the 0-10 scale) represents the possible values. A linear interpolation can then be performed.

From an operational viewpoint, each indicator involved in the assessment of the performance related to a failure mode has to be assessed by an expert as a possibility distribution. To this end, the experts were first coached during a training session that consisted of a brief presentation aimed at explaining the objectives of the assessment, the meanings of possibility distributions ("a distribution of possibility depicts the possibility that a variable takes value x"), the types of distribution proposed and the expected results of the session. It also allowed the expert to ask questions to clarify possibly obscure points. Next, the experts assessed the indicators as possibility distributions. During the assessment sessions, the expert had at her/his disposal the description grid for each indicator assessed (i.e. grids such as presented in Table 1) along with monitoring data collected at the dam and visual observations he/she carried out.

Figure 3 gives a real illustration of the assessment of the indicators involved in the assessment of the performance linked to internal erosion through a dam embankment, here called DAM89. The latter is a French embankment dam with an upstream membrane on an alluvium foundation, 8 m high, and impounding a lake holding 350 000 m^3. The bituminous upstream shoulder is covered with rip-rap protection up to the operating level. It has a drain located behind the membrane, a horizontal drain at the interface with the foundation and a draining toe weight. The dam is equipped with a side spillway. Reservoir filling began in 1979. Figure 3 shows the values of the various indicators declared by an expert as possibility distributions. Some of them are assessed as precise scores (Flow/Flow change, holes, hairline cracking, etc.) as no imperfection was noted by the expert. Others are expressed as triangle (“visual state of drain outlet”) or trapezoid-shaped distributions (“seepage of clean water”). They represent imperfections caused in particular by the fact the expert was unable to fully observe the drain outlet and determine the magnitude of clean water seepage due to the presence of vegetation on the downstream slope – the “vegetation present” indicator is assessed as a precise score of 6 (“poor”).

For the dam application, the height of the fuzzy set is equal to 1 as the intrinsic reliability of the source, i.e. expert reliability, is not assessed. However, the uncertainty, imprecision and incompleteness of the data, as well as the data processing (expert judgment on the basis of the visual or auscultation data and statistical processing) are integrated in the frame, core width and base of the fuzzy set. For the building application, the height of the fuzzy set integrates the reliability of the source, uncertainty and incompleteness of the data and the imprecision of the transformation into the fuzzy frame. However, the core and base widths represent the intrinsic imprecision of the data.

Data unification

In some cases, several types of data from different sources are available. This is the case for the building application with data associated with the degradation phenomena (for example, carbonation, biological deposit, cracking). When imprecise and uncertain data are modelled with fuzzy sets (possibility theory), the merging method (evidence theory) allows taking into account the confidence that can be assigned to data sources and provides a consensual result that integrates the angles of interpretation of the different sources in the same problem (for example, duration of a carbonation phenomenon). As the reliability of these sources is different, the estimations given do not have the same relative confidence as that calculated using the data quality analysis.

Merging (unification method) allows building a consensual service life obtained using estimations from these different sources. Data unification is necessary for each phenomenon in order to obtain their durations after which the data is aggregated from the level of the phenomenon to that of the scenario, in order to obtain the scenario durations. For the dam application, a single datum is associated with each performance indicator: the unification process has no meaning.

The unification is based on evidence theory. The principle of merging *n* data, defined by an interval of values (or a singleton) and a belief mass, consists in: (1) taking two data and calculating the mass assigned to the intersection of their intervals, union, frame of discernment (depending on the data merging method used), (2) repeating this calculation by combining the resulting data and one of the last *n*-2 data until all the *n* data have been combined. Several merging methods are available, such as the Dempster-Shafer intersection (1976), Dubois' and Prade's union (1992), Yager's report on ignorance (1994) and the intersection of low consistency from Lair (2000). These merging methods make it possible to manage, in varying degrees, the characteristics of the available data, the number of data, and the consistency and conflicts between them. J. Lair (2000) proposed a choice algorithm to select the best merging method. However, in our work, we used the Dempster-Shafer method as the data used are non-conflicting. Table 4 presents this method considering two data (D_1 and D_2) that point to two intervals of values (I_1 and I_2). Rules of belief mass allocation, according to Dempster-Shafer, are presented in this table: (i) for non-conflicting data (intersection $I_1 \cap I_2$ non-nil), (ii) for conflicting data without re-allocation of belief masses, and (iii) for conflicting data with re-allocation. Figure 5 gives an example of merging two carbonation duration data ($D_{3\text{-}1}$ and $D_{3\text{-}2}$ of section "model for the performance assessement – case of buildings"). The top left of the figure shows the two data on the time axis. The tables at the bottom correspond to the method of merging data 1 and 2 and the result of this merger. The top right of the figure shows the result of the merger, i.e. the consensus curve. The interpretation of this consensus curve is provided in the defuzzification section (cf. section "defuzzification and decision-making").

Imperfection propagation and defuzzification

Imperfection propagation

Once the imperfect data are represented, they have to be propagated in the performance models, as in the model presented in Figure 3.

The propagation of possibility distributions via operation *f* obeys Zadeh's extension principle (Zadeh, 1978):

[7] $$\pi_{Fi}(s_{Fi}) = \sup_{(S_1,\ldots S_n)/f(S_1,\ldots S_n)=S_{Fi}} (\min(\pi_{S1}(S_1),\ldots,\pi_{Sn}(S_n)))$$

with $(S_1,\ldots,S_n)$ being the input data scores and S_{Fi} the output score.

Function *f* can be either a direct mathematical operation (addition, subtraction, product, division, maximum, minimum, mean), or a function stemming from IF-THEN fuzzy rules.

The case of DAM89 is used to illustrate the process from an operational viewpoint (cf. Figure 3). The values given by the indicators are first bottom-up aggregated to obtain the degradation of the technical functions using Equation 7, with *f* being the maximum operator. Two of them are assessed as precise scores (the sealing and the erosion protection functions) and the other (drainage) is assessed in a trapezoid distribution. Then, the assessment of phenomena and dam safety related to internal erosion through the embankment is performed using fuzzy rules (see Equations 2 and 3). Here the result of the fuzzy aggregation is the score of the deterioration of dam safety related to internal erosion through the embankment. The score obtained is precise and equal to 3 ("passable"). The expert recommended scrapping the downstream slope to enable: (1) better inspection of this slope, which would lead to improving the quality of the "vegetation present", "seepage of clean water" and "visual state of drain outlet" indicators, and (2) reducing the score of the "presence of trees" indicator.

For the building application, scenario duration is assessed by the sum of the phenomena durations that composed this scenario. Moreover, the unification process (cf. section "data unification") provides a consensus curve for each phenomenon. A consensual duration and Smet's probability

($P(\theta)=\sum_{\theta\in R_l}\frac{m_c(R_l)}{|R_l|}$) are deduced from this consensus curve (cf. section "defuzzification and decision-making"). Consequently, the reliability associated with the scenario duration corresponds to the minimum Smet's probability (or belief mass when no unification is carried on) of phenomenona. By way of example, the belief mass or Smet's probability of phenomena durations shown in Figure 5 are: 0.6 for D_1, 0.4 for D_2, 0.81 for D_3 (cf. section "defuzzification and decision-making") and 0.7 for D_4. Consequently, the scenario duration of 1 mm crack occurrence is [76; 86] years and its reliability is 0.4.

Defuzzification and decision-making

The result is expressed as a fuzzy frame at the end of the imperfection propagation phase. Defuzzification consists in extracting only the information required for the decision-making associated with the problem in question so that this information is the best representation of the data unification and aggregation results. The defuzzification results must integrate the reliability of the result obtained, which depends on the data collected and the methodology.

Several methods are available, such as those providing real values (Leeckwijck *et al*. 2001; Liu 2007), methods providing intervals of values (Dubois *et al*. 1987; Chanas 2001), methods classifying fuzzy sets in relation to the others (Asady and Zendehnam 2007; Fortemps *et al*. 1996), plausibility and belief (Shafer 1976), Smets' probability (Dubois 1990), etc. As these methods can lead to different results, it is important to make a relevant choice of method. Indeed, the results obtained comprise the parameter on which the final decision is made. The defuzzification phase provides consensual data and information on the quality of this data for both applications.

In the building application, the consensual data represents a value or an interval of values of service-life of building components or phenomena durations (cf. Figure 5). This data is deduced from the consensual function; it corresponds to the data associated with the *k*% fractal of the consensus. The quality indicators are belief ($Bel(\theta)=\sum_{R_l\subseteq\theta}m_c(R_l)$), plausibility ($Pl(\theta)=\sum_{\theta\cap R_l\neq\varnothing}m_c(R_l)$) and Smets's probability associated with service-life. $m_c(R_l)$ represents the belief mass assigned to interval R_l

resulting from data merging. For the example of carbonation duration (cf. Figure 5), the duration that groups 100% of consensuses is the [26; 28] year interval. The belief is equal to 0.77 (belief mass of [26; 28] year interval). The plausibility is equal to the sum of belief masses of [0; 36] year, [25; 30] year and [26; 28] year intervals ($Pl([26;28]) = 1$), as these intervals have a non-nil intersection with a [26; 28] year interval. Smet's probability is calculated as follows:

[8] $$P([26;28]) = 0.77 + 2/5 \times 0.08 + 2/36 \times 0.15 = 0.81$$

The belief indicator provides a pessimistic boundary of the quality interval while the plausibility indicator corresponds to an optimistic boundary of this interval. Smet's probability offers a balance between both these indicators. Indeed, it balances the belief masses of the intervals considered ($\theta \in R_l$), by considering the common length between them and the θ interval ($Bel(\theta) \leq P(\theta) \leq Pl(\theta)$).

The width of the uncertainty interval depends on the reliability of the data collected, their informativeness and the aggregation model.

From the operational viewpoint, this defuzzification result also provides feedback on the data collected. Indeed, the closer Smets's probability is to 1, the narrower the belief-plausibility interval, and the more reliable the service life obtained will be. Therefore, the data indicating decreased reliability can be identified and more reliable data can be sought on the basis of the intermediate data unification and data aggregation results.

In the dam application, consensual data represents a value on the assessment scale, i.e. the performance of a dam related to failure modes. Dam experts require several kinds of information to interpret dam performance in a fuzzy frame for a failure mode:

- For decision-making: fuzzy set comparison methods;
- For transmitting information to dam managers: synthetic information that integrates all the information of the fuzzy sets.

The defuzzification methods were selected by experts during group sessions. They are presented in Table 5.

The advantage of performing a dam performance assessment at a moment in its service life is that it permits defining the repair actions to be carried out. From the operational standpoint, this defuzzification result also provides feedback on the data collected. Indeed, analysis of the values of function performances and indicators allows identifying the data that lead to reduced reliability and thus propose actions to collect more reliable data.

Conclusion

This paper addresses specific domains where data is numerous, heterogeneous and imperfect, which is often the case in civil engineering. To cope with these difficulties, a four-phase method was proposed consisting of: (i) performance assessment modelling, (ii) the formalisation of heterogeneous information from different sources, (iii) the propagation of imperfections in performance assessment models, and (iv) the expression of outputs of the models in suitable formats to facilitate decision-making and improve communication. The methodological choices of each of these steps have to be adapted to the specificity of the application. For example:

- In the case of uncertain and imprecise data, the theory of possibilities is an adequate methodological frame, while in the case of multiple data sources, belief function theories are particularly relevant;
- Fuzzification is performed by using mathematical models for the building application while it is assessed by experts for the dam application;
- The defuzzification process is adapted to both the context and decision-making. The methodology results in a performance assessment and an estimation of its reliability. It provides an indicator of the quality of the result obtained that integrates the reliability and the informativeness of the data collected on the one hand, and the reliability of the data unification and data aggregation models on the other.

The methodology was illustrated with operational viewpoints, using two civil engineering structures, namely a dam and a building, but the methodology should be transposable to other applications.

Several difficulties can be encountered:

- Performance assessment modelling: the main drawback of the knowledge-based method is the possibly long duration of development. The development of such models obviously requires the existence and availability of knowledge satisfying the project's objectives. A key requirement in this approach is to identify the appropriate level of abstraction;
- Fuzzification and the propagation of imperfection: assumptions stated during the fuzzification phase, determination of transition thresholds between phenomenon durations when assessing scenario duration. Sensitivity analysis of these assumptions is a way to assess their impacts on the results. Concerning merging method, it can be performed using the method proposed by Lair (2000);
- Defuzzification: a plethora of defuzzification methods exist. The choice is not obvious and requires in-depth analysis of the problem;
- Validation: it is necessary to propose specific validation protocols. Before performing full-scale validations, an initial validation stage was performed on the basis of simulations. In the case of dams, simplified cases were built from completed dam reports written at the end of detailed dam reviews performed by Irstea experts.

Two future directions seem interesting to us. The first concerns the addition of a fifth phase to the methodology with the automated proposal of corrective actions based on the results of the different function performances and the defuzzification process. The second deals with handing on knowledge, something which should be facilitated by the methodology proposed. The method and associated tool can be used as a training aid by novice engineers only recently involved in dam reviews.

References

Andersen, G. R., Chouinard, L. E., Hover, W. H., and Cox, C. W. (2001). “Risk indexing tool to assist in prioritizing impovement embankment dam inventories.” Journal of Geotechnical and Geoenvironnemental Engineering, 127, 325-334.

Asady, B., and Zendehnam, A. (2007). "Ranking fuzzy numbers by distance minimization." Applied Mathematical Modelling, 31, 2589–2598.

Baroth, J., Breysse, D. and Schoefs, F. (2011). "Construction Reliability." ISTE-Wiley, Hoboken.

Baudrit, C., Couso, I., and Dubois, D. (2007). "Joint propagation of probability and possibility in risk analysis: Towards a formal framework." International Journal of Approximate Reasoning, 45, 82-105.

Brandt, E. and Wittchen, K.B. (1999). "EPIQR – A new surveying tool for maintenance and refurbishment." Proceedings of the Durability of Building Materials and Components 8, 30 May – 3 June 1999, Vancouver. National Research Council of Canada, Ottawa.

Chanas, S. (2001). "On the interval approximation of a fuzzy number." Fuzzy Sets and Systems, 122, 353-356.

Cheng, C.T., Ou, C.P. and Chau, K.W., (2002). "Combining a fuzzy optimal model with a genetic algorithm to solve multiobjective rainfall-runoff model calibration." Journal of Hydrology, 268 (1-4), 72-86.

Curt, C., Peyras, L., and Boissier, D. (2010). "A knowledge formalisation and aggregation-based method for the assessment of dam performance." Computer-aided Civil and Infrastructure Engineering, 25, 171-183.

Curt, C., Talon, A., and Mauris, G. (2011). "A dam assessment support system based on physical measurements, sensory evaluations and expert judgements." Measurement, 44, 192-201.

Dempster, A. P. (1976). "Upper and lower probabilities induced by a multivalued mapping." Annals of Mathematical Statistics, 38, 325-339.

Dubois, D., and Prade, H. (1992). "On the relevance of non-standard theories of uncertainty in modeling and pooling expert opinions." Reliability Engineering and System Safety, 36, 95-107. Dubois, D., and Prade, H. (1990). "Consonant approximations of belief functions." International Journal of Approximate Reasoning, 4, 244-264.

Dubois, D., and Prade, H. (1987). "The mean value of a fuzzy number." Fuzzy Sets and Systems, 24, 179–300.

Fortemps, P., and Roubens, M. (1996). “Ranking and defuzzification methods based on area compensation.” Fuzzy Sets and Systems, 82, 319-330.

Foster, M., Fell R., and Spannagle M. (2000). “A method for assessing the relative likelihood of failure of embankment dams by piping.” Canadian Geotechnical Journal, 37, 1025-1061.

Funtowicz, S. O., and Ravetz J. R. (1990). “Uncertainty and Quality in Science for Policy.” Kluwer Academic Publishers, Dordrecht.

Jernberg, P., Lacasse, M. A., Haagenrud, S. E., and Sjöström, C. (2004). “Guide and Bibliography to Service Life and Durability Research for Building Materials and Components.” CIB Report, Publication n°295.

Kim, Y.S. and Kim, B.T. (2008) “Prediction of relative crest settlement of concrete-faced rockfill dams analyzed using an articifical neural network model”. Computers and Geotechnics, 35, 313-322.

Kreuzer, H. (2000). “The use of risk analysis to support dam safety decisions and management.” In Proceedings of 20th International Congress on Large Dams, Beijing, China.

Lair, J. (2000). “Evaluation de la durabilité des systèmes constructifs du bâtiment.” Ph.D. thesis. Centre Scientifique et Technique du Bâtiment and Laboratoire Génie Civil, Clermont-Ferrand University, France. [In French.]

Lee, A., and Barrett, P. (2003). “Performance based building: First International State-of-the Art Report [online].” Available from: http://www.pebbu.nl [cited 11 february 2011].

Leekwijck, W. V., and Kerre, E. E. (2001). “Continuity focused choice of maxima: Yet another defuzzification method.” Fuzzy Sets and Systems, 122, (2) 303–314.

Liu, X. (2007). “Parameterized defuzzification with maximum entropy weighting function – Another view of the weighting function expectation method.” Mathematical and computer modeling, 45, 177-188.

Modarres, M. (1993). “What every engineer should know about reliability and risk analysis.” Marcel Dekker INC, New York.

MRGenCi. 2009. “GIS Maîtrise des Risques en Génie Civil **[On line].”** Available on: http://www.mrgenci.u-bordeaux1.fr/ [cited 30 March 2009].

Peyras, L., Royet, P. and Boissier, D. (2006). "Dam ageing diagnosis and risk analysis: Development of methods to support expert judgement", Canadian Geotechnical Journal, 43, 169-186.

Serre, D., Peyras, L., Curt, C., Boissier, D., and Diab, Y. (2007). "Evaluation des ouvrages hydrauliques de génie civil." Canadian Geotechnical Review, 44, 1298-1313. [In French.]

Shafer, G. (1976). "A mathematical theory of evidence." Princeton University Press, Chichester.

Shannon, C. E. (1948). "A Mathematical Theory of Communication." The Bell System Technical Journal, 27, 379-423 and 623-656.

Siemes, A., Vrouwenvelder, A. and Van Der Beukel, A. (1985). "Stochastic modelling of building materials performance in durability." Proceedings of the Problems in Service Life Prediction of Building and Construction Materials, NATO ASI Series. Martinus Nijhoff Publishers, Dortrecht.

Talon, A., Boissier, D., and Lair, J. (2008). "Service life assessment of building components: Application of evidence theory." Canadian Journal of Civil Engineering, 35(3), 66-70.

Talon, A. (2006). "Evaluation des scénarii de dégradation des produits de construction." Ph.D. thesis. Centre Scientifique et Technique du Bâtiment and Laboratoire Génie Civil, Clermont-Ferrand University, France. [In French.]

United Nations. (2009). "Division for Sustainable Development [**On line**]". Available on: http://www.un.org/esa/dsd/index.shtml [cited 30 March 2009].

Wu, C.L. and Chau, K.W., (2006). "A flood forecasting neural network model with genetic algorithm." International Journal of Environment and Pollution, 28 (3-4), 261-273.

Yagger, R. R., Kacprzyk, J., and Fedrizzi, M. (1994). "Advances in the Dempster-Shafer Theory of Evidence." John Wiley & Sons, New-York.

Zadeh L.A. (1978). Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems, 1, 3-28.

Figure caption list:

Figure 1: Principle of the methodology proposed for building and dam applications and associated methods

Figure 2: Methodology for assessing building service life – Case of the crack occurrence scenario **Figure 3:** Hierarchical model of internal erosion through the embankment – Case of a dam with a upstream bituminous membrane – Assessment of DAM89 using possibility distributions – µF: function performance – φ: phenomenon – µMR: performance related to failure mode

Figure 4: Formalization and signification of fuzzy sets for building and dam applications

Figure 5: Example of the merging of $D_{3\text{-}1}$ and $D_{3\text{-}2}$ carbonation durations by using the intersection method

Table 1. Slab thickness indicator

Name	Slab thickness (sectional view) – Concrete upstream facing
Definition	Concrete slabs constitute the sealing component for embankment dams with concrete facings. Sufficient thickness is necessary to satisfy sealing requirements
Scale and references 0: Excellent 1-2: Good 3-4: Passable 5-6: Poor 7-9: Bad 10: Unacceptable	0-2: slabs 40 cm thick 3-4: slabs 30 cm thick 7-10: slabs less than 30 cm thick
Location	Concrete upstream facing
Time characteristics	Evaluation carried out when emptying the reservoir

Table 2. Building data fuzzification

Type	Core	Base
More than x years	*x*	[*x*; 2*x*]
Less than y years	*y*	[0; *y*]
z years	*z*	[0.9*z*; 1.1*z*]
Between *y'* and *x'* years	[*y'*; *x'*]	[*y'*; *x'*]
Representation		

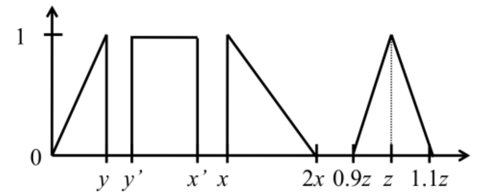

Table 3. Quality criteria in the case of a building

Categories	Criteria	Possible values	Example for *D*3-1 (carbonation data)
Quality of raw data	1. Source	0 (isolated) to 1 (referenced)	2/3
	2. Censorship	0 (not taken into account) to 1 (no censured)	1
Potential of data acquisition	3. Type of data modeling	0 (non modeled) to 1 (established theory)	1
	4. Hypothesis of modeling	0 (very high) to 1 (poor)	2/3
	5. Type of acquisition method	0 (poor estimation) to 1 (experience plan)	1
Analysis of data treatment	6. Consistency of data preparation	0 (poor) to 1 (excellent)	2/3
	7. Correspondence regarding the case study	0 (poor) to 1 (excellent)	1
	8. Entropy	$\frac{\sum_{x \in X} f_{mc}(x)\ln(f_{mc}(x))}{\ln(1/\Theta)}$ where X represents the interval of duration, x is a discrete value of the interval X, Θ is the "frame of discernment" (i.e. observation period and $f_{mc}(x)$ is the belief function associated with the value x.	0,19 with Θ = (0; 36) years
Result			Belief mass = 0.77

Table 4. Rules of belief mass allocation according to Dempster-Shafer

		I_1	$I_1 \cap I_2$	I_2	Ignorance	
D_1	Non conflicting	$m_c(I_1)$	$m_c(I_1)$	$1-m_c(I_1)$	$1-m_c(I_1)$	
	Conflicting	$m_c(I_1)$		$1-m_c(I_1)$	$1-m_c(I_1)$	
D_2	Non conflicting	$1-m_c(I_2)$	$m_c(I_2)$	$m_c(I_2)$	$1-m_c(I_2)$	
	Conflicting	$1-m_c(I_2)$		$m_c(I_2)$	$1-m_c(I_2)$	
Consensus	Non conflicting	$m_c(I_1)\cdot(1-m_c(I_2))$	$m_c(I_1)\cdot m_c(I_2)$	$(1-m_c(I_1))\cdot m_c(I_2)$	$(1-m_c(I_1))(1-m_c(I_2))$	$\sum = 1$
	Conflicting without re-allocation	$m_c(I_1)\cdot(1-m_c(I_2))$		$(1-m_c(I_1))\cdot m_c(I_2)$	$(1-m_c(I_1))\cdot(1-m_c(I_2))$	$\sum = 1- m_c(S_1)\ m_c(S_2)$
	Conflicting with re-allocation	$\frac{m_c(I_1)\cdot(1-m_c(I_2))}{1-m_c(S_1)\cdot m_c(S_2)}$		$\frac{(1-m_c(I_1))\cdot m_c(I_2)}{1-m_c(S_1)\cdot m_c(S_2)}$	$\frac{(1-m_c(I_1))\cdot(1-m_c(I_2))}{1-m_c(S_1)\cdot m_c(S_2)}$	$\sum = 1$

Table 5. Defuzzification methods selected for dam application

Objective	Defuzzification method
Communication	1. Defuzzification value: upper limit of the α-cut at 0.8 2. Qualitative description: percentage of correspondance between the given fuzzy set and reference intervals 3. Qualitative interpretation of the dispersion from the relative specificity
Decision making	Classification of fuzzy sets in comparison with the average value of the α-cut at 0.8

Figure 1. Principle of the methodology proposed for building and dam applications and associated methods

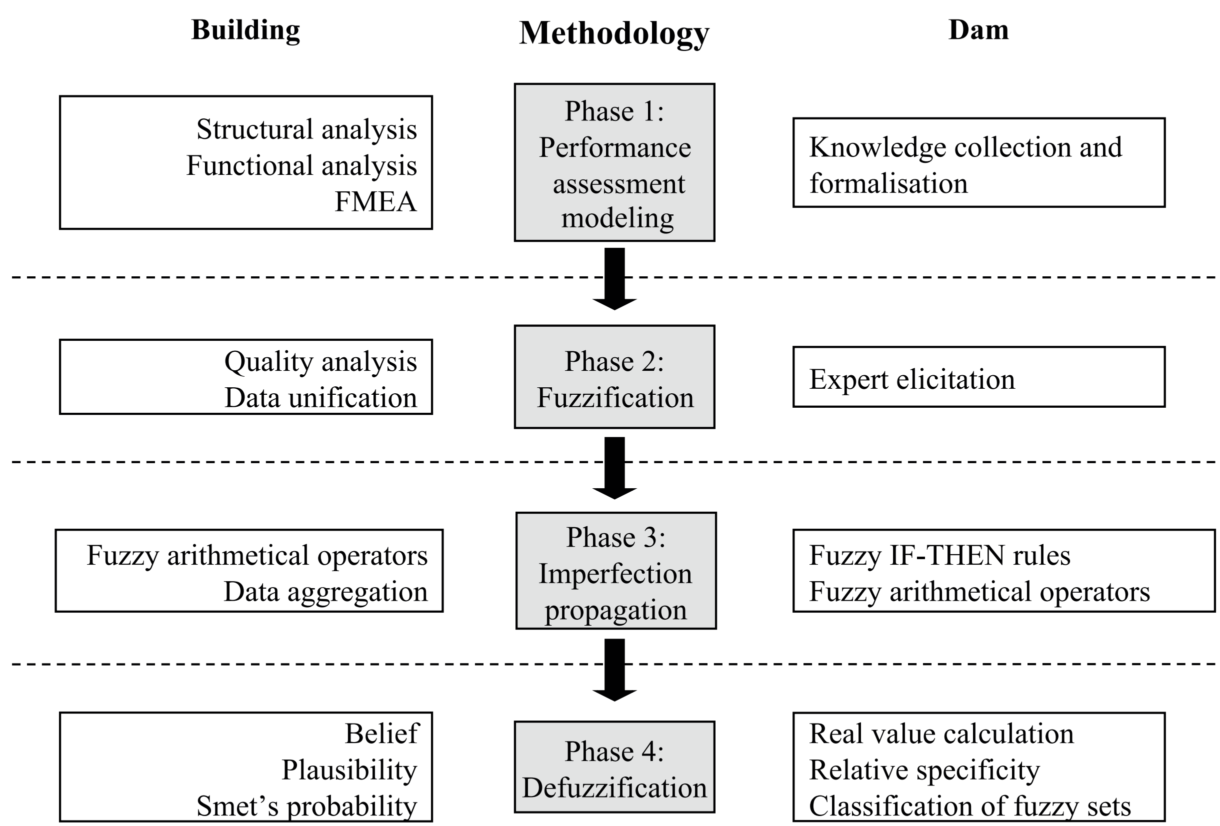

Figure 2. Methodology for assessing building service life – Case of the crack occurrence scenario

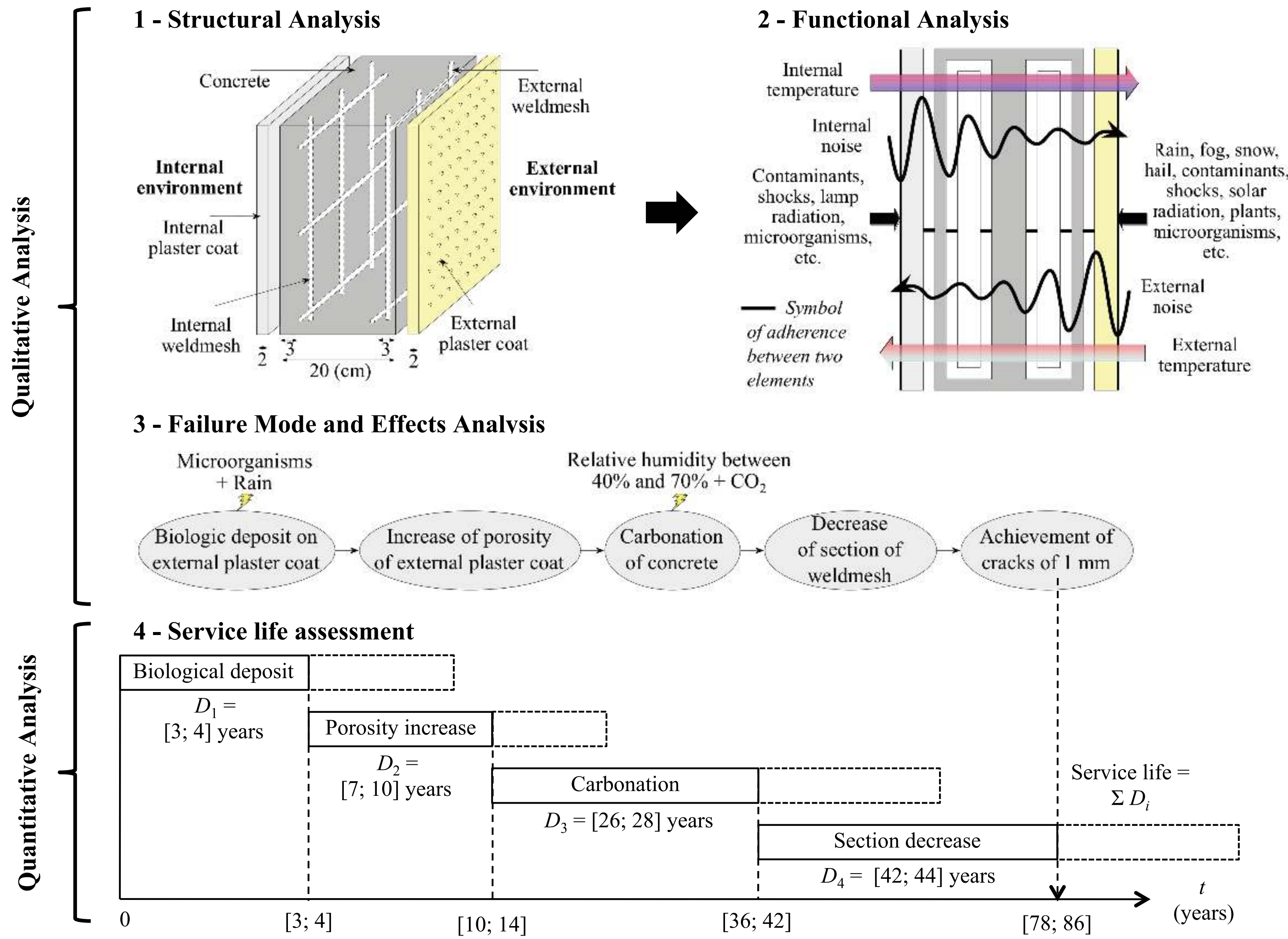

Figure 3. Hierarchical model of internal erosion through the embankment – Case of a dam with a upstream bituminous membrane – Assessment of DAM89 using possibility distributions – µF: function performance – φ: phenomenon – µMR: performance related to failure mode

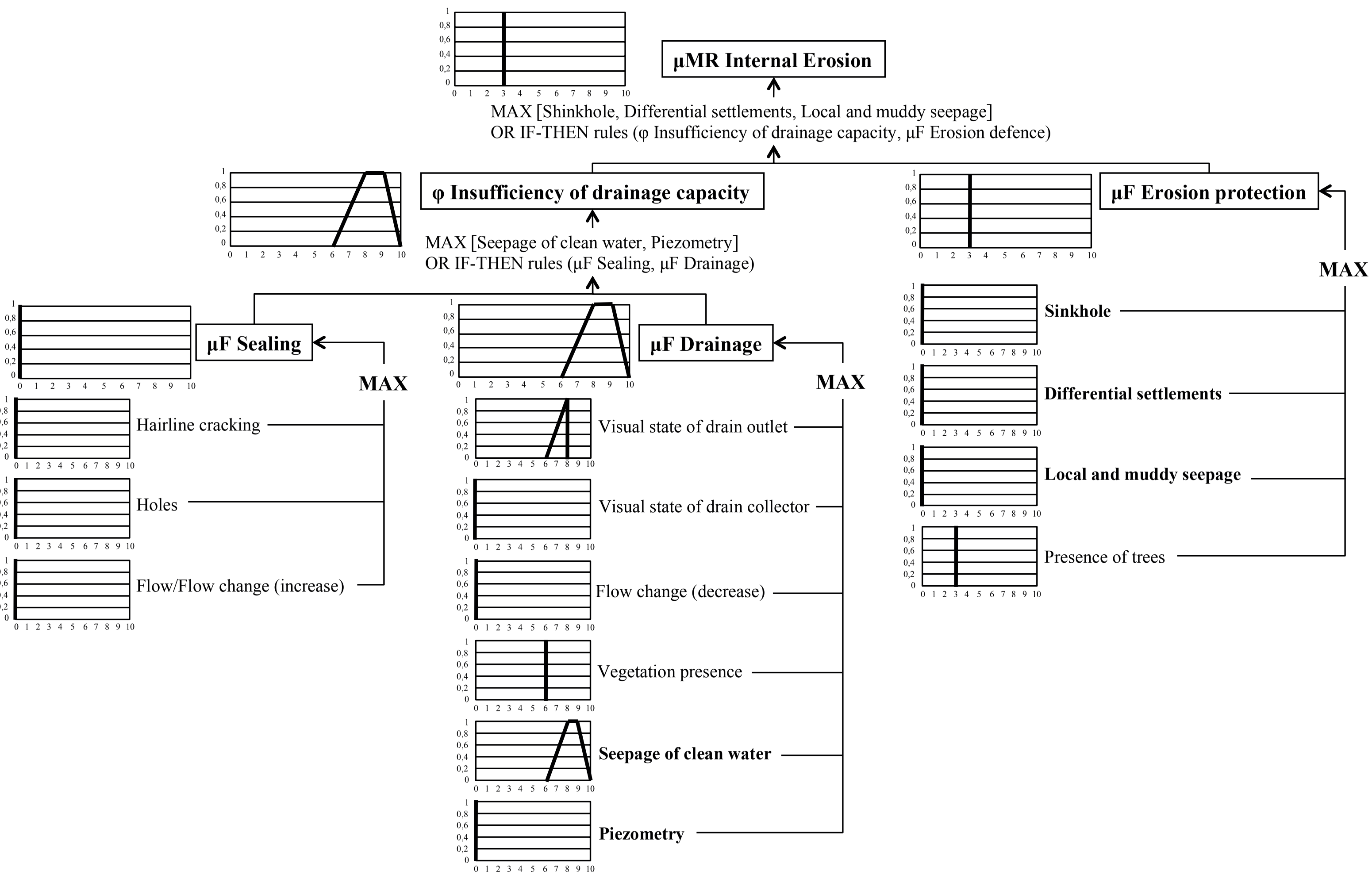

Figure 4. Formalization and signification of fuzzy sets for building and dam applications

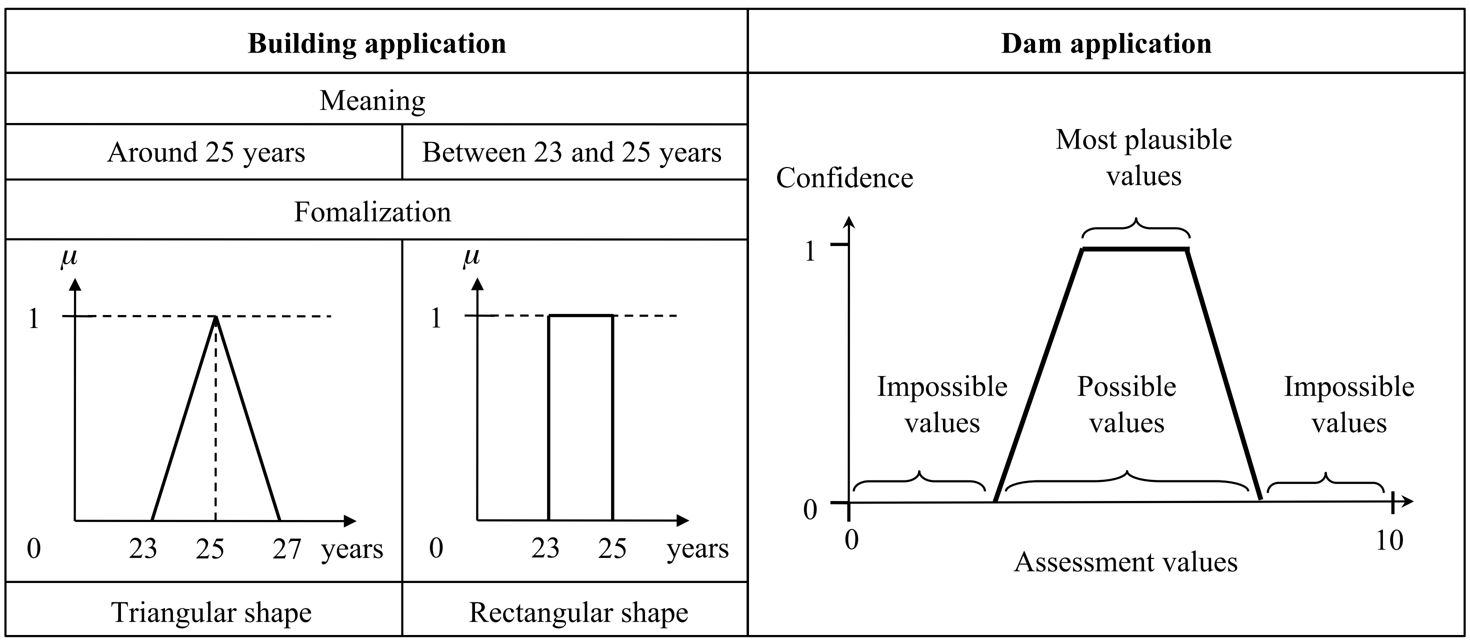

Figure 5. Example of the merging of $D_{3\text{-}1}$ and $D_{3\text{-}2}$ carbonation durations by using the intersection method