

# Simulation based structural reliability assessment involving imprecise data

M. Beer

*Department of Mechanical Engineering and Materials Science, Rice University, Houston, TX, USA  
Institut für Statik und Dynamik der Tragwerke, Technische Universität Dresden, Germany*

P. D. Spanos

*Ryon Chair in Engineering, Rice University, Houston, TX, USA*

*Keywords: Simulation techniques, imprecise data, uncertainty, imprecise probabilities, fuzzy randomness*

**ABSTRACT:** In this paper a new sampling method for probabilistic safety assessment of structures involving imprecise data is presented. This method is formulated as sample-induced simulation technique and aims at combining the benefits of established simulation methods like Monte-Carlo simulation with those of models with imprecise data. In this regard an approach is pursued to overcome certain weaknesses of current methods. First, the specification of probability distributions on the basis of frequently small samples is circumvented. All information contained in a given sample is accounted for directly and without estimating probability distribution functions explicitly. Second, uncertainty in sample elements is taken into account. As no probability distributions are required, the considerable numerical cost of evaluating samples comprising imprecise data is eliminated. Examples demonstrate the capability of the sample-induced simulation technique for dealing with both real-valued data and imprecise data.

## 1 INTRODUCTION

Simulation techniques often offer the only possibility for solving problems in which random properties must be taken into account. Indeed, Monte-Carlo simulation and further developments thereof have become versatile tools for solving a variety of problems in a wide range of engineering disciplines, see Schuëller and Spanos (2001).

An essential precondition for obtaining realistic results from a simulation is the availability of statistically-validated probability distributions for the input variables. The specification of these distributions thus plays an essential role, see Schuëller (2001). For determining reliably parameters and forms of probability distributions, extensive data in the form of samples are required. This enables using well-developed and sophisticated methods of statistical estimation theory and test theory, which operate parametrically or non-parametrically.

In any case, problems may primarily occur in the following two situations. First, the available information is limited in the form of small samples. Second, the sample elements are characterized by imprecision. As a result, the required probability distributions cannot be specified to a sufficient degree

of reliability.

In this paper an attempt is made to develop a simulation technique for structural reliability assessment which yields realistic results in the aforementioned problematic situations.

## 2 PROBLEM SPECIFICATION

In the case of small samples, statistical estimations and tests may yield ambiguous results. For an appropriate level of confidence, wide intervals for the estimated values are obtained. This applies, in particular, if the distribution type is not pure, if a compound or multimodal distribution underlies, or if multi-dimensional dependencies are present. The quality of the numerical sampling result then depends, inter alia, on the degree with which the specified probabilistic model actually agrees with the small sample.

Imprecision or uncertainty of sample elements is generally either taken into account approximately by selecting "probably adverse values" from a possible value range or neglected totally. However, the actual impact of such a selection can generally not be evaluated at the pre-stage of a simulation. On the

other hand, the question arises as to how to model that uncertainty or imprecision realistically. It appears, for example, in situations in which the precision of measuring devices is strongly limited, the measuring points cannot be defined precisely (rough surfaces in thickness measurements), the expert evaluations influence the value specification, the measured values are gained under dubious conditions, and linguistic assessments are accounted for and thus generally possesses no random properties.

To perform realistic structural analysis and safety assessment in those problematic cases, an attempt is made to develop a simulation technique that operates in conjunction with a generalized uncertainty model and circumvents an explicit specification of a probabilistic model. It should be capable of attaining appropriate simulation results starting from samples of small size that consist of uncertain or imprecise data. The development starts from the basic statistical assumption that all information is contained in the sample. In the simulation a second sample of considerably larger size is numerically generated that reflects the statistical properties and uncertainty characteristics of the original small sample "as well as possible". The proposed technique is thus referred to as "sample-induced simulation".

### 3 CONCEPT

The basic concept of the new sample-induced simulation technique is to generate the sampling result directly from a given sample instead of estimating a probability distribution and performing the sampling according to this. The characteristics of a population are described by a sufficiently large sample. As the mathematical model of a distribution function is not employed herein, conventional statistical estimations are dispensed with. The concept of statistical estimation is applied in a generalized sense.

The starting point is the observed sample  $S_0$  of size  $n_0$ . A second concrete sample  $S_1$ , of a considerably larger size  $n_1 \gg n_0$  is then sought that represents the original sample  $S_0$  "as well as possible". That is, the new sample  $S_1$  is expected to exhibit statistical characteristics "comparable" to  $S_0$ . This is realized by the following iterative approach with the superscript <sup>[i]</sup> indicating the iteration step.

1. The starting point is an arbitrary estimate  $S_1^{[0]}$  for the sample  $S_1$ . This is broadly specified without consideration of the information contained in the observed sample  $S_0$ . All sample elements of  $S_1^{[0]}$  should possess the same information content. That is, they should exhibit the same probability

density in their immediate surroundings. This leads to the specification of  $S_1^{[0]}$  by uniform distribution over a sufficiently large (physically meaningful) range of values of the random variable represented by  $S_0$ .

2. The sample  $S_1^{[0]}$  is compared with the observed sample  $S_0$ . The purpose of this comparison is to obtain a measure  $G^{[0]}$  for the statistical dissimilarity between the samples  $S_1^{[0]}$  and  $S_0$ . For this dissimilarity measure, a real valued function is selected which yields a global minimum for  $G^{[0]}$  if the samples  $S_1^{[0]}$  and  $S_0$  are "as similar as possible" in a statistical sense. That is,  $G^{[0]}$  is intended to be minimal if  $S_0$  and  $S_1^{[0]}$  originate from the same population with probability one. Due to the fact that intended application is for samples consisting of imprecise data, established statistical test methods cannot be implemented.
3. The sample  $S_1^{[0]}$  is modified in such a way that  $m_1$  elements (stipulated number with  $m_1 \ll n_1$ ) are selected by discrete uniform distribution and replaced by new elements to obtain the modified sample  $S_1^{[1]}$ . The new elements are again generated randomly over the range of values specified in Step 1. This ensures that  $S_1$  is obtained as a random sample in consistency with established sampling principles. The measure value  $G^{[1]}$  is computed for the modified sample  $S_1^{[1]}$ .
4. The measure values  $G^{[0]}$  and  $G^{[1]}$  are compared. If  $G^{[1]} \leq G^{[0]}$ , it is concluded that the modification in Step 3 has not yield an improved estimation for  $S_1$ . The modification is then nullified, and a repeat modification is carried out according to Step 3. If  $G^{[1]} > G^{[0]}$ , on the other hand, the modified sample  $S_1^{[1]}$  yields an improved estimation compared with  $S_1^{[0]}$ . The sample  $S_1^{[1]}$  is then taken as the basis for the next iteration step and modified anew according to Step 3 to produce  $S_1^{[2]}$ . Again, the result is assessed. This procedure is repeated with an iteration counter  $r$  for successful modifications until it is no longer possible to obtain an improvement of  $S_1$  beyond  $S_1^{[r]}$ . As this can only be realized with zero-probability (in the continuous case), a termination limit is defined for the probability with which an improvement can be obtained. The iteration is terminated if the average success rate of modifications attains a predefined and sufficiently small value. Finally,  $S_1$  is set equal to  $S_1^{[r]}$  obtained from the last successful modification.

The sampling result  $S_1$  directly contains the input vectors for a subsequent  $n_1$ -fold structural analysis or may be evaluated for a safety assessment. In the case

of samples comprising imprecise data, the special algorithms presented in Möller and Beer (2004) may be applied.

#### 4 PROCEDURE FOR REAL-VALUED DATA

The sample-induced simulation technique is developed first, to apply for processing real-valued samples. The samples are deemed real-valued in the sense that their elements are denoted by scalars or vectors consisting of real numbers. This enables assessing the results with the aid of established test methods. In this manner, the effectiveness of the sample-induced simulation may be evaluated.

##### 4.1 Basic aspects

The critical point of the proposed technique is to formulate an appropriate function for characterizing the statistical dissimilarity  $G^{[r]}$  between  $S_1^{[r]}$  and  $S_0$  in each iteration step  $r$  (see Step 2 in Sect. 2). This function is required to possess the following four properties.

1. The measure  $G^{[r]}$  and established statistical test methods (homogeneity tests) must lead to basically analogous propositions regarding the statistical dissimilarity between  $S_1^{[r]}$  and  $S_0$ . These propositions must be free of contradictions.
2. The mathematical formulation of the dissimilarity measure  $G^{[r]}$  must be extendable to apply for fuzzy-valued samples. That is, the mathematical operations used in the definition of  $G^{[r]}$  for the real-valued case must possess appropriate counterparts in fuzzy arithmetics.
3.  $G^{[r]}$  is required to decrease – at least tendentially – with decreasing statistical dissimilarity between  $S_0$  and  $S_1^{[r]}$ . For samples  $S_0$  and  $S_1^{[r]}$  originating from the same population the measure  $G^{[r]}$  should take its global minimum value.
4. The mathematical structure of the measure  $G^{[r]}$  should be as simple as possible to ensure a fast numerical evaluation and to keep the computational cost reasonably low.

To develop a measure  $G^{[r]}$  that satisfies these requirements the following theoretical experiment is considered. According to the statistical estimation theory it is assumed that all available information is contained in the observed sample  $S_0$ . Then, the best description of  $S_0$  is its empirical distribution function as it is a complete and unique representation of the information in  $S_0$ . Moreover, in inferential statistics, the empirical distribution function is one of the most powerful estimators. If this is taken as the basis for

sampling – to numerically generate the sample  $S_1$  –, and no smoothing is applied, the resulting sample  $S_1$  and the observed sample  $S_0$  possess identical empirical distributions (in the limit). This corresponds to two significant properties of the samples  $S_0$  and  $S_1$  with respect to each other. First, the positions of the elements of  $S_0$  and  $S_1$  coincide. Second, each element of  $S_0$  has the same number of uniquely assigned elements from the sampling result  $S_1$ . In the case of an underlying continuous random variable and an – accordingly – slightly smoothed empirical distribution, the elements of the sampling result  $S_1$  are obtained in a close neighborhood of the elements of  $S_0$  – with the same assignment property.

Sampling results generated in this manner are high quality representations of the underlying sample  $S_0$  as may be shown by applying a variety of two-sample tests of homogeneity.

The measure  $G^{[r]}$  is thus formulated based on the configuration of the sampling result  $S_1$  from the theoretical experiment. This provides two criteria for monitoring the dissimilarity  $G^{[r]}$  between  $S_1^{[r]}$  and  $S_0$ , which are defined as an assignment criterion and a distance criterion.

##### 4.2 Assignment criterion

The assignment criterion evaluates some order in the element configuration in the samples  $S_0$  and  $S_1$  with respect to each other. Each element  $\underline{s}_{0,i}$  ( $i = 1, \dots, n_0$ ) from sample  $S_0$  is supposed to have the same number  $n_{\text{ass}}(\underline{s}_{0,i})$  of uniquely assigned elements  $\underline{s}_{1,k}$  ( $k = 1, \dots, n_1$ ) from sample  $S_1^{[r]}$ . The element assignment is defined on the basis of the shortest Euclidean distance  $d(\underline{s}_{0,i}; \underline{s}_{1,k})$  between the respective elements  $\underline{s}_{0,i}$  and  $\underline{s}_{1,k}$ . For each  $\underline{s}_{1,k}$  one assigned element  $\underline{s}_{0,i}(\underline{s}_{1,k})$  is determined with

$$\underline{s}_{0,i}(\underline{s}_{1,k}) = \underline{s}_{0,i} \mid d(\underline{s}_{0,i}; \underline{s}_{1,k}) = \min_{i=1, \dots, n_0} [d(\underline{s}_{0,i}; \underline{s}_{1,k})] \quad (1)$$

The number  $n_{\text{ass}}(\underline{s}_{0,i})$  may then be obtained by means of an indicator function,

$$n_{\text{ass}}(\underline{s}_{0,i}) = \sum_{k=1}^{n_1} I(\underline{s}_{0,i}; \underline{s}_{1,k}), \quad (2)$$

$$I(\underline{s}_{0,i}; \underline{s}_{1,k}) = \begin{cases} 1 & \text{if } \underline{s}_{0,i} = \underline{s}_{0,i}(\underline{s}_{1,k}) \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

The target value for the number  $n_{\text{ass}}(\underline{s}_{0,i})$  is given by the ratio of the sample sizes  $n_1$  and  $n_0$ . The assignment criterion is then defined as the total sum of the quadratic differences between the actual numbers  $n_{\text{ass}}(\underline{s}_{0,i})$  and the target value,

$$C_1 = \sum_{i=1}^{n_0} \left( n_{\text{ass}}(\underline{s}_{0,i}) - \frac{n_1}{n_0} \right)^2 \rightarrow \text{MIN} . \quad (4)$$

The smallest possible value of  $C_1$  depends on the sample sizes  $n_1$  and  $n_0$ . With the integer parameter

$$a \in \mathbf{Z} \mid a \cdot n_0 \leq n_1 < (a + 1) \cdot n_0 \quad (5)$$

this limit is

$$\min_{C_1} = -\frac{1}{n_0} (a \cdot n_0 - n_1)^2 + n_1 - a \cdot n_0 \quad (6)$$

In the special case that the size of sample  $S_1^{[r]}$  is a whole multiple of the size of  $S_0$  the value  $\min_{C_1}$  is equal to zero.

### 4.3 Distance criterion

The distance criterion supplements the assignment criterion by additionally evaluating the particular position of the sample elements. The distances between assigned sample elements are supposed to be as small as possible. Specifically,

$$C_2 = \sum_{k=1}^{n_1} \left( d(\underline{s}_{0,i}(\underline{s}_{1,k}); \underline{s}_{1,k}) \right)^2 \rightarrow \text{MIN} , \quad (7)$$

with  $\underline{s}_{0,i}(\underline{s}_{1,k})$  specifying the assignment of  $\underline{s}_{1,k}$  to  $\underline{s}_{0,i}$  determined with Eq. (1). The smallest possible value of the distance criterion is zero.

### 4.4 Composing the dissimilarity measure

To define the dissimilarity measure  $G^{[r]}$  for real-valued samples  $S_0$  and  $S_1^{[r]}$  the assignment criterion according to Eq. (4) and the distance criterion according to Eq. (7) are combined. As a standard formulation, the quantity

$$G^{[r]} = \sqrt{C_1 + C_2} \quad (8)$$

is selected. An extension of Eq. (8) by introducing weighting factors for the criteria  $C_1$  and  $C_2$  has been investigated in several numerical tests; it has not been found particularly effective for improving the simulation results.

### 4.5 Assembling the iteration procedure

The dissimilarity measure  $G^{[r]}$  in Eq. (8) is implemented into the numerical procedure according to Steps 1 through 4 in Sect. 2. Moreover, the number  $m_1$  of elements, see Step 3, which are modified in each iteration step, is randomly selected and varied frequently. For the random generation of the  $m_1$  new

elements in iteration step  $r$  the (slightly smoothed) empirical distribution of the sample  $S_1^{[r-1]}$  from the previous successful iteration step  $r-1$  is used. In this manner, use is made of the statistical information already gathered during the iteration.

The termination limit in Step 4 is chosen to be 2% and applied to the moving average of the recent 100 successful iteration steps.

## 5 PROCEDURE FOR IMPRECISE DATA

### 5.1 Modeling imprecise data

For dealing with imprecise data, see Viertl (1996) for relevant concepts and terminology, we must select a suitable data model that combines the benefits of the well-established probabilistic approach with an appropriate modeling of non-frequentative uncertainty or imprecision. The class of available uncertainty models in this context comprises interval analysis, convex modeling, evidence theory, random set theory, p-box, theory of rough sets, chaos models, fuzzy set theory, and the theory of fuzzy random variables. Considering the capability of these models, the concept of fuzzy random variables originally presented in Kwakernaak (1978) is selected for further investigation. This model possesses the advantage of simultaneously covering the models of real-valued random variables, fuzzy sets, intervals, random intervals, and convex models as special cases.

To define a fuzzy random variable the probability space  $[\mathbf{X}, \mathcal{E}, P]$  is extended by the dimension fuzziness. If the space of the random elementary events, as in probabilistics, is described by  $\Omega$ , a fuzzy random vector  $\underline{\tilde{X}}$  on the fundamental set  $\mathbf{X} = \mathbb{R}^n$  may be defined as the fuzzy result of the uncertain mapping

$$\Omega \rightsquigarrow \mathbf{F}(\mathbb{R}^n) \quad (9)$$

where  $\mathbf{F}(\mathbb{R}^n)$  is the set of all fuzzy numbers in  $\mathbb{R}^n$ . An ordered  $n$ -tuple of fuzzy numbers  $\tilde{x}_i$  is assigned to each (crisp) elementary event  $\omega \in \Omega$ . Every  $n$ -tuple  $\underline{\tilde{X}}(\omega) = (\tilde{x}_1, \dots, \tilde{x}_n) \subseteq \mathbf{X}$  is a realization of the fuzzy random vector  $\underline{\tilde{X}}$ . Both objective and subjective information are accounted for simultaneously. The theory of fuzzy random variables permits the modeling of uncertain structural parameters which partly exhibit randomness but which cannot be described using real-valued random variables without an element of doubt. The randomness is "disturbed" by a fuzziness component.

A comprehensive discussion on fuzzy randomness particularly with regard to engineering problems may be found in Möller and Beer (2004). In

this context the concepts of fuzzy structural analysis, see also Möller et al. (2000), and fuzzy probabilistic safety assessment, see also Möller et al. (2003), describe the processing of uncertain structural parameters with the aid of numerical procedures. This basis ensures an appropriate evaluation of the results from sample-induced simulation of fuzzy random variables within the framework of structural analysis and safety assessment.

Statistical investigations of uncertain or imprecise data and of properties of fuzzy random variables are, to a great extent, in an initial stage of development. Related research in this regard may be found in Bandemer and Näther (1992), in Viertl (1996), and in Körner (1997). These developments concern the analysis of imprecise data, the definition of statistical parameters, and the investigation of statistical laws for fuzzy random variables.

Publications discussing the simulation of fuzzy randomness are rare. An approach evaluating fuzzy probability distribution functions on a trajectory-by-trajectory basis is presented in Sickert et al. (2003). Numerical investigations of statistical properties of fuzzy random variables based on simulation are discussed in Colubi et al. (2002). However, these methods require prior knowledge about the fuzzy probability distributions or the fuzziness of the realizations to be generated. General techniques for generating fuzzy realizations of fuzzy random variables are not known at the present time.

The application of traditional sampling methods to the numerical generation of fuzzy realizations encounters considerable difficulties. For instance, the numerical effort for estimating fuzzy parameters and fuzzy probability distributions from fuzzy-valued samples (fuzzy samples) is significantly high, in particular, when interaction between the fuzzy parameters is taken into account. Moreover, the simulation of fuzzy realizations starting from fuzzy probability distribution functions is not unique. That is, different fuzzy samples may have identical empirical fuzzy probability distribution functions.

These conflicts hinder the pursuing of a traditional simulation approach and may be circumvented by implementing the uncertainty model fuzzy randomness into the sample-induced simulation technique. Due to the generalized character of this uncertainty model, the capability of processing real-valued samples is hereby preserved as a special case.

#### 4.2 Extension of criteria $C_1$ and $C_2$

To extend the sample-induced simulation to apply for fuzzy samples  $\tilde{S}_0$  and  $\tilde{S}_1$ , a suitable replacement for the Euclidean distance  $d(\underline{s}_{0,i}, \underline{s}_{1,k})$  must be

introduced to enable criteria  $C_1$  and  $C_2$  to deal with fuzzy vectors  $\tilde{s}_{0,i}$  and  $\tilde{s}_{1,k}$ . For this purpose the fuzzy vectors  $\tilde{s}_{0,i}$  and  $\tilde{s}_{1,k}$  are represented with the aid of  $\alpha$ -discretization, see Fig. 1. For a sufficiently high number of  $\alpha$ -levels the fuzzy vectors  $\tilde{s}_{0,i}$  and  $\tilde{s}_{1,k}$  are completely described by the set of their  $\alpha$ -level sets  $\underline{s}_{0,i,\alpha}$  and  $\underline{s}_{1,k,\alpha}$ , respectively.

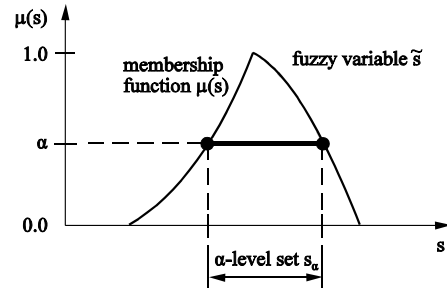


Figure 1.  $\alpha$ -discretization of a fuzzy variable

On this basis, the distance  $d_F(\tilde{s}_{0,i}, \tilde{s}_{1,k})$  between the fuzzy vectors  $\tilde{s}_{0,i}$  and  $\tilde{s}_{1,k}$  may be defined by recombining the distances  $d_H(\underline{s}_{0,i,\alpha}, \underline{s}_{1,k,\alpha})$  between the associated  $\alpha$ -level sets  $\underline{s}_{0,i,\alpha}$  and  $\underline{s}_{1,k,\alpha}$  (for the same  $\alpha$ -level). Specifically, the metric

$$d_F(\tilde{s}_{0,i}, \tilde{s}_{1,k}) = \int_{\alpha=+0}^{\alpha=1} d_H(\underline{s}_{0,i,\alpha}, \underline{s}_{1,k,\alpha}) d\alpha \quad (10)$$

is applied, see Körner (1997), which makes use of the Hausdorff metric

$$d_H(\underline{s}_{0,i,\alpha}, \underline{s}_{1,k,\alpha}) = \max [d_{H,1}, d_{H,2}], \quad (11)$$

$$d_{H,1} = \sup_{\underline{s}_0 \in \underline{s}_{0,i,\alpha}} \inf_{\underline{s}_1 \in \underline{s}_{1,k,\alpha}} [d(\underline{s}_0, \underline{s}_1)],$$

$$d_{H,2} = \sup_{\underline{s}_1 \in \underline{s}_{1,k,\alpha}} \inf_{\underline{s}_0 \in \underline{s}_{0,i,\alpha}} [d(\underline{s}_0, \underline{s}_1)],$$

between the associated  $\alpha$ -level sets  $\underline{s}_{0,i,\alpha}$  and  $\underline{s}_{1,k,\alpha}$ , see Fig. 2. The outcome of Eq. (11) and hence the distance  $d_F(\tilde{s}_{0,i}, \tilde{s}_{1,k})$  from Eq. (10) are crisp values, which are directly applied in Eqs. (1) and (7) to eventually compute criteria  $C_1$  and  $C_2$ .

The application of criteria  $C_1$  and  $C_2$  to evaluate the dissimilarity of fuzzy-valued samples enables a consideration of the order in the element configuration and the distance between the respective sample elements. Dissimilarities in the fuzziness of the elements  $\tilde{s}_{0,i}$  and  $\tilde{s}_{1,k}$ , however, are taken into account only to a partial degree. In addition to the criteria  $C_1$

and  $C_2$ , the fuzziness of the realizations provides a basis for a third dissimilarity criterion.

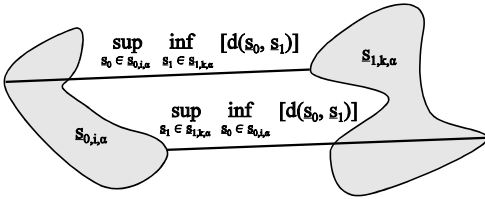


Figure 2. Hausdorff metric applied to  $\alpha$ -level sets

### 4.3 Fuzziness criterion

The fuzziness criterion evaluates the matching in the fuzziness of the respective fuzzy sample elements  $\tilde{s}_{0,i}$  and  $\tilde{s}_{1,k}$ . Fuzzy sample elements that are assigned to each other according to the assignment rule Eq. (1) are supposed to exhibit the same fuzziness. For this purpose, the fuzziness of the sample elements is computed with an analog to Shannon's entropy applied to the membership functions  $\mu(\underline{s}_{0,i})$  and  $\mu(\underline{s}_{1,k})$  of  $\tilde{s}_{0,i}$  and  $\tilde{s}_{1,k}$ , respectively. For the fuzzy vector  $\tilde{s}$ , this uncertainty measure is defined as

$$H_u(\tilde{s}) = -k \cdot \int_{\underline{s}} g(\mu(\underline{s})) d\underline{s} \quad (12)$$

$$g(\mu(\underline{s})) = \mu(\underline{s}) \cdot \ln(\mu(\underline{s})) + (1 - \mu(\underline{s})) \cdot \ln(1 - \mu(\underline{s})) .$$

And the fuzziness criterion is

$$C_3 = \sum_{k=1}^{n_1} (H_u(\tilde{s}_{0,i}(\tilde{s}_{1,k})) - H_u(\tilde{s}_{1,k}))^2 \Rightarrow \text{MIN} . \quad (13)$$

For a "perfect matching", the fuzziness criterion  $C_3$  becomes zero.

### 4.4 Procedure features for imprecise data

The generation and the iterative modification of a fuzzy-valued sample  $\tilde{S}_1^{[r]}$  require not only determining the position of the sample elements  $\tilde{s}_{1,k}$  but also specifying their membership functions  $\mu(\underline{s}_{1,k})$ . A new fuzzy realization  $\tilde{s}_{1,k}$  is generated in the following three steps. First, the mean value  $\underline{s}_{1,k} \in \tilde{s}_{1,k} \mid \mu(\underline{s}_{1,k}) = 1$  is drawn from the current smoothed empirical distribution of the mean values in the fuzzy sample  $\tilde{S}_1^{[r-1]}$ . Then, the fuzziness  $H_u(\tilde{s}_{1,k})$  is determined by means of a logarithmic normal distribution estimated from the fuzziness  $H_u(\tilde{s}_{0,i})$  of the fuzzy sample elements  $\tilde{s}_{0,i}$  in the observed fuzzy sample  $\tilde{S}_0$ . In a third step, the shape of the membership function  $\mu(\underline{s}_{1,k})$  is

also randomly specified according to the distribution of the shape in  $\tilde{S}_0$ .

The consideration of fuzzy samples requires incorporating the criterion  $C_3$  into the iterative procedure. Tests have shown that it is effective to perform the iteration for fuzzy samples in two steps. In the first step, only the criteria  $C_1$  and  $C_2$  are satisfied. Subsequently, the obtained element assignment and the mean value positions are frozen. In the second step, criterion  $C_3$  is applied in a separate fuzziness iteration. That is, in the second iteration step, only the  $H_u(\tilde{s}_{1,k})$  and the shape of the membership functions of the fuzzy sample elements  $\tilde{s}_{1,k}$  are adjusted. The iteration termination criterion is also applied separately in both iteration steps.

## 6 RELIABILITY ASSESSMENT

Structural reliability assessment based on sample-induced simulation is realized as a straightforward extension to traditional methods. The sampling result  $\tilde{S}_1$  is directly evaluated with regards to the limit state surface in the original space of the fuzzy-probabilistic basic variables. That is, the structural reliability is determined by counting the fuzzy sample elements that lead to failure; see Möller and Beer (2004). Due to the fuzziness, however, some fuzzy sample elements lie only partly in the failure region  $S_f$ . This leads to a fuzzy failure probability  $\tilde{P}_f$ . For computing  $\tilde{P}_f$   $\alpha$ -discretization is applied again, see Sect. 4.2. Specifically,

$$\tilde{P}_f = \left\{ \left( P_{f,\alpha}, \mu(P_{f,\alpha}) \right) \right\} , \quad (14)$$

$$P_{f,\alpha} = [P_{f,\alpha 1}, P_{f,\alpha r}] ,$$

$$\mu(P_{f,\alpha}) = \alpha \quad \forall \alpha \in (0, 1] .$$

The interval bounds  $P_{f,\alpha 1}$  and  $P_{f,\alpha r}$  are calculated with the aid of indicator functions and particular conditions for evaluating fuzzy realizations; see Möller and Beer (2004). Specifically,

$$P_{f,\alpha 1} = \frac{1}{n_1} \cdot \sum_{k=1}^{n_1} I_{\alpha 1}(\tilde{s}_{1,k}) , \quad (15)$$

$$I_{\alpha 1}(\tilde{s}_{1,k}) = \begin{cases} 1 & \text{if } \underline{s}_{1,k,\alpha} \in S_f \\ 0 & \text{otherwise} \end{cases} ,$$

and

$$P_{f,ar} = \frac{1}{n_1} \cdot \sum_{k=1}^{n_1} I_{ar}(\tilde{s}_{1,k}), \quad (16)$$

$$I_{ar}(\tilde{s}_{1,k}) = \begin{cases} 1 & \text{if } \underline{s}_{1,k,\alpha} \cap S_f \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

## 7 NUMERICAL EXAMPLES

### 7.1 Real-valued data

#### 7.1.1 Sampling

A one-dimensional real-valued sample  $S_0$  of size  $n_0 = 200$  is taken as the basis, see Fig. 3. This is numerically generated from a compound distribution consisting of two extreme value distributions of Ex-Max type I. The extreme values of the sample  $S_0$  are  $\min_{S_0} = 5.1$  and  $\max_{S_0} = 21.55$ .

An initial estimate  $S_1^{[0]}$  is numerically generated by uniformly distributing  $n_1 = 10,000$  sample elements  $s_{1,k}$  over the interval  $[0, 25]$ , see Fig. 3. Then, the iteration to improve the generalized estimation  $S_1^{[r]}$  is started. The number  $m_1$  of modified elements is randomly selected from the interval  $[5, 30]$  and frequently changed during the iteration. After about  $r = 4,000$  iteration steps the average success rate starts decreasing distinctly and attains the termination limit in iteration step  $r = 4,710$ , see Fig. 3.

Clearly, there is no visible difference between the empirical distribution functions of the samples  $S_0$  and  $S_1 = S_1^{[4,710]}$ , see Fig. 3. The sampling result shows no clumping of the generated sample elements  $s_{1,k}$  around the original sample elements  $s_{0,i}$ . Homogeneity tests (Kolmogorov-Smirnov and chi-squared) yield rejection probabilities of  $P < 0.001$  for the  $H_0$  hypothesis that both samples originate from the same population. The tails of the generated sample  $S_1$  run beyond the extreme values of  $S_0$  with  $\min_{S_1} = 3.26$  and  $\max_{S_1} = 24.01$ . A total of 39 elements  $s_{1,k}$  are smaller than  $\min_{S_0} = 5.1$ , and 48 elements  $s_{1,k}$  are bigger than  $\max_{S_0} = 21.55$ . The proportions of  $S_1$  therewith correspond to an extreme value distribution with a thicker tail on the right side than on the left side. Fishers exact probability test yields a probability of  $P = 0.386$  with which the  $H_0$  hypothesis is not rejected.

Results generated via traditionally estimated probability distributions did not attain the quality level of the present sample  $S_1$ . Kernel based estimation methods led to samples showing test results comparable to the present approach. Their tails, however, did not run significantly beyond  $\min_{S_0}$  and  $\max_{S_0}$ .

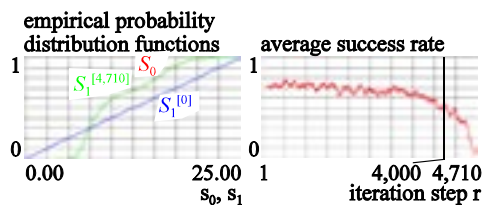


Figure 3. Empirical distribution functions of  $S_0$ ,  $S_1^{[0]}$ , and  $S_1^{[4,710]}$ ; average success rate (last 100 steps) during iteration

#### 7.1.2 Reliability assessment

The reliability level assessment is pursued by directly evaluating the sampling result  $S_1$  with respect to a given limit state surface. Since the related procedures are well-known, these are not highlighted in the example. Herein, it is focused on the dependency of the assessment result on the quality of the sampling result.

As an example, the observed sample  $S_0$  is interpreted as a possible record of a live load  $s$  resulting from road traffic and acting on a structural member of a road bridge. The sampling result  $S_1$  then represents a statistical loading prognosis for future traffic. For defining a limit state surface, the serviceability requirement  $s \leq 22$  is defined.

The empirical failure probability obtained from sample  $S_0$  is  $P_f = 0$ , whereas the sampling result  $S_1$  yields  $P_f = 3.4 \cdot 10^{-3}$ . A compound probability distribution estimated from  $S_0$  without additional prior knowledge leads to  $P_f = 1.7 \cdot 10^{-3}$ . According to the underlying extreme value distribution  $P_f = 8.9 \cdot 10^{-3}$  is obtained. These results indicate a good agreement between the prognoses from traditional approaches and from sample-induced simulation.

### 7.2 Imprecise data

#### 7.2.1 Sampling

As a starting point the sample  $S_0$  from Sect. 7.1.1 is "fuzzified" to represent an uncertain measurement series, for example, of a live load, see Sect. 7.1.2. The resulting fuzzy sample  $\tilde{S}_0$  consists of  $n_0 = 200$  fuzzy triangular numbers with fluctuating fuzziness  $H_{\alpha}(\tilde{s}_0)$  over the sample elements; for relevant concepts and terminology see Bandemer and Näther (1992) and Möller and Beer (2004). An initial estimate  $\tilde{S}_1^{[0]}$  of size  $n_1 = 10,000$  is generated in compliance with Sect. 3.2.3 starting from uniformly distributed mean values and restricting the fuzzy sample elements completely to  $[0, 25]$ , see Fig. 4. Again, the iteration is carried out with a randomly selected number  $m_1 \in [5, 30]$  of modified elements. First, the dissimilarity measure  $G^{[r]}(C_1, C_2)$  is

minimized in 5,990 iteration steps. The empirical distributions of  $\tilde{S}_0$  and  $\tilde{S}_1^{[5,990]}$  agree very well. However, there is almost no correspondence between the fuzziness  $H_u(\tilde{s}_0)$  and  $H_u(\tilde{s}_1)^{[5,990]}$  of the respective fuzzy sample elements, see Fig. 4. The subsequent fuzziness iteration (minimization of criterion  $C_3$  up to iteration step  $r = 16,150$ ) almost does not affect the empirical distribution, but improves considerably the fuzziness agreement, see Fig. 4.

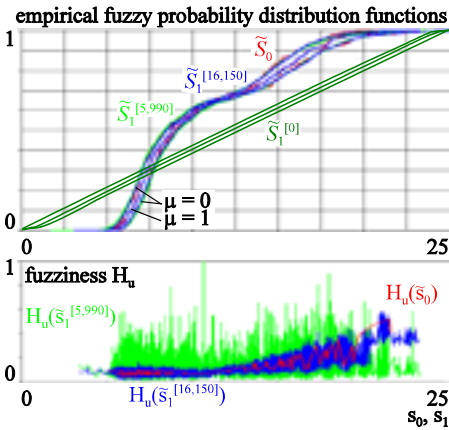


Figure 4. Empirical fuzzy probability distribution functions of  $\tilde{S}_0$ ,  $\tilde{S}_1^{[0]}$ ,  $\tilde{S}_1^{[5,990]}$ , and  $\tilde{S}_1^{[16,150]}$  and fuzziness of the associated fuzzy sample elements

### 7.2.2 Reliability assessment

The serviceability requirement  $s \leq 22$  specified in Sect. 7.1.2 is evaluated with the fuzzy samples  $\tilde{S}_0$  and  $\tilde{S}_1 = \tilde{S}_1^{[16,150]}$ . The fuzzy failure probability  $\tilde{P}_f$  is computed according to Eqs. (14), (15), and (16) with eleven  $\alpha$ -levels, see Fig. 5. Whereas sample  $\tilde{S}_0$  yields an almost useless result with an overestimated fuzziness, sample  $\tilde{S}_1$  leads to a more meaningful result. The probability values covered by  $\tilde{P}_f$  from  $\tilde{S}_1$  again comprise a reasonable range with respect to the results from traditional estimations and from the underlying distribution for the mean values ( $\mu = 1$ ) presented in Sect. 7.1.2.

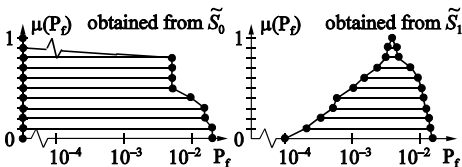


Figure 5. Empirical fuzzy failure probability obtained from  $\tilde{S}_0$  and from  $\tilde{S}_1$

## 6 CONCLUSIONS

The presented sample-induced simulation technique may be useful if the data bank comprises, solely, a small sample with uncertain or imprecise elements. It operates free of a probability model and is capable of considering randomness and non-stochastic impreciseness simultaneously. Future developments may focus on multidimensional problems including dependencies in terms of probability and fuzzy set theories.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of the German Research Foundation (DFG), of the Alexander von Humboldt-Foundation (AvH), and of the Office of Naval Research (ONR), USA.

## REFERENCES

Bandemer, H. & Näther, W. 1992. *Fuzzy Data Analysis*. Dordrecht: Kluwer Academic Publishers.

Colubi, A., Fernández-García, C. & Gil, M.A. 2002. Simulation of random fuzzy variables: an empirical approach to statistical/probabilistic studies with fuzzy experimental data. *IEEE Transactions on Fuzzy Systems* Vol. 10: 384–390.

Kwakernaak, H. 1978. Fuzzy random variables – I. Definitions and Theorems. *Information Sciences* Vol. 15: 1–29.

Körner, R. 1997. *Linear Models with Random Fuzzy Variables*. Diss., Bergakademie Freiberg, Fakultät für Mathematik und Informatik.

Möller, B. & Beer, M. 2004. *Fuzzy-Randomness – Uncertainty in Civil Engineering and Computational Mechanics*. Berlin, Heidelberg, New York: Springer.

Möller, B., Graf, W. & Beer, M. 2000. Fuzzy structural analysis using  $\alpha$ -level optimization. *Computational Mechanics* Vol. 26 (No. 6): 547–565.

Möller, B., Graf, W. & Beer, M. 2003. Safety assessment of structures in view of fuzzy randomness. *Computers and Structures* Vol. 81 (No. 15): 1567–1582.

Schuëller, G.I. 2001. On Computational Procedures for Processing Uncertainties in Structural Mechanics. *Proceedings 2nd European Conference on Computational Mechanics*. Cracow, June 26–29, 2001, CD-ROM, Doc. 608: 1–24.

Schuëller, G.I. & Spanos, P.D. (eds) 2001. *Monte Carlo simulation: proceedings of the International Conference on Monte Carlo Simulation*. Monaco, June 18–21, 2000. Lisse, Exton, PA: A.A. Balkema.

Sickert, J.-U., Beer, M., Graf, W. & Möller, B. 2003. Fuzzy probabilistic structural analysis considering fuzzy random functions. *Proceedings 9th International Conference on Applications of Statistics and Probability in Civil Engineering*. San Francisco, July 6–9, 2003, Rotterdam: Millpress, Vol. 1: 379–386.

Viertl, R. 1996. *Statistical Methods for Non-Precise Data*. Boca Raton, New York, London, Tokyo: CRC Press.