Evaluation of Inconsistent Engineering data

Michael Beer
Department of Civil Engineering
National University of Singapore
BLK E1A #07-03, 1 Engineering Drive 2, Singapore 117576
email: cvebm@nus.edu.sg

Abstract. In this paper options for a realistic evaluation of engineering data characterized by inconsistency regarding uncertainty and imprecision are discussed. The proposed methods are linked to the generalized uncertainty model fuzzy randomness. This enables a quantification of uncertainty and imprecision simultaneously with a smooth transition between fuzziness and randomness. Statistical information is exploited with traditional statistical methods, whereas imprecision is dealt with using fuzzy methods. Statistical uncertainty and imprecision are considered within the same model but not mixed with one another. In this manner, both components are reflected separately in the computational results from a subsequent structural or safety analysis. Quantification techniques are elucidated for three typical engineering cases of inconsistent information; (i) small sample size and expert knowledge, (ii) imprecise sample elements, and (iii) inconsistent environmental conditions and expert knowledge. The usefulness of the proposed quantification methods for a subsequent structural analysis and safety assessment is demonstrated by way of engineering examples.

Keywords: Inconsistent data; Imprecise data; Fuzzy methods; Fuzzy probabilities; Uncertain structural analysis; Safety assessment.

1. Introduction

The usefulness of the results from an engineering analysis depends significantly on the realistic modeling of the input parameters. Shortcomings, in this regard, may lead to biased computational results, wrong decisions, and serious consequences [18]. This applies, in particular, if the data are characterized by uncertainty and imprecision. A variety of mathematical models have been formulated to take account of the available information as realistically as possible [3, 6, 7, 10, 13, 14, 15, 17, 23, 24, 28, 29]. The usefulness and capabilities of these models have already been demonstrated in the solution of practical problems, for example, in civil/mechanical engineering [1, 4, 5, 8, 9, 11, 12, 14, 16, 19, 21, 22, 25].

In engineering practice the available information frequently appears as partly stochastic and partly imprecise – in a mixed stochastic/non-stochastic form. In those cases the model fuzzy randomness [19] provides a proper basis to utilize traditional statistical methods together with quantification methods from fuzzy set theory. In this manner, a broad spectrum of typical engineering cases can be covered; and the introduction of unwarranted information is avoided. This is demonstrated in the sequel with proposals of quantification techniques for three typical engineering situations. First, the quantification of data from a small sample together with expert knowledge is considered. This is associated with the problem of weak statistical information from estimations and tests. A solution is obtained by utilizing the statistical imprecision in the specification of fuzzy parameters and fuzzy distribution types of a fuzzy random quantity.

Second, samples with imprecise elements are evaluated, which requires the application of statistics with fuzzy quantities. For this purpose, fuzzy arithmetic is implemented in statistical estimations and tests. Third, inconsistent environmental conditions are dealt with together with expert knowledge. This leads to critical conditions for statistical estimations and tests. For solution, a separation of fuzziness and randomness is applied in the quantification procedure by constructing groups of consistent data.

In all three cases fuzzy random quantities are obtained which reflect the stochastic uncertainty and the imprecision of the underlying information simultaneously and separately. The fuzzy probability distributions are described as a bunch of distributions that cover all possible stochastic models within the range of imprecision. Bunch parameters are fuzzy quantities $\tilde{p}_t$, which include distribution parameters as well as parameters for the specification of the distribution type. Then, each crisp point from the $\tilde{p}_t$ specifies one real-valued random quantity associated with a certain membership degree according to fuzzy set theory. For a detailed description see [19]. This enables the utilization and combination of sophisticated and numerically efficient methods from stochastic mechanics [26, 27] and from interval [22] and fuzzy structural analysis [20] in subsequent engineering computations. The respective algorithms of fuzzy stochastic structural analysis and safety assessment are discussed in [19].

2. Small Sample size and Expert Knowledge

Assume that a concrete sample of small size is available. The sample elements are random realizations. The available information on the sample is insufficient, however, to describe a real-valued random variable free of doubt. The type of the distribution function and the parameters cannot be determined uniquely; additional uncertainty exists. Expert knowledge and experience are available from similar cases in the past. This uncertainty is rather non-stochastic and may be accounted for with the aid of fuzzy set theory [2, 30]. Statistical methods may be used as a basis for quantification, which are supplemented by fuzzy methods to finalize the modeling. Depending on the available information it is possible to formulate an imprecise parametric or nonparametric estimation problem. On this basis, the type and the parameters of the sought distribution are determined in as imprecise quantities, namely, as fuzzy quantities. These fuzzy quantities are, subsequently, lumped together as fuzzy parameters $\tilde{p}_t(\tilde{X})$, in which $\tilde{X}$ represents a fuzzy random quantity – for convenience, limited to the one-dimensional case. The $\tilde{p}_t(\tilde{X})$ may be determined from imprecise empirical statistical information extracted from the sample together with expert knowledge.

If, for example, the type of distribution is known with sufficient certainty, this implies an imprecise, parametric estimation problem. The sample functions applied in statistical methods yield more or less acceptable estimation values for the parameters of a distribution. In order to take account of the imprecision of the estimator, confidence intervals may be determined for the estimator in question. The probabilistic propositions for confidence intervals applied in statistical methods may then serve as additional information for the specification of the membership functions $\mu(p_t(X))$ of the $\tilde{p}_t(\tilde{X})$ in the present case. Expert knowledge is brought in with regard to

- the specification of the distribution type,
- the choice of the estimator,
- the construction of confidence intervals (type and levels),

– the assignment of membership degrees to the selected confidence levels, and

– the subsequent modification of the initial draft of the membership functions $\mu(p_t(X))$.

Table I. Sample of the cylinder compressive strength f_c of a concrete

Number i of realization	Compressive strength $x_i = f_{ci}[N/mm^2]$	Number i of realization	Compressive strength $x_i = f_{ci}[N/mm^2]$
1	28.3	11	26.8
2	31.5	12	35.3
3	35.2	13	26.3
4	29.8	14	23.1
5	27.6	15	20.2
6	30.7	16	29.2
7	25.2	17	25.7
8	34.6	18	34.2
9	28.9	19	24.8
10	19.2	20	22.8

Table II. Statistical estimation and assignment of membership values for $\tilde{m}_x$ and $\tilde{\sigma}_x$

Estimation	Confidence level	m_X	σ_X	α-level
Point	- –	27.97	4.75	1.00
Interval	0.50	[27.24, 28.70]	[4.35, 5.43]	0.75
	0.75	[26.71, 29.23]	[4.05, 5.92]	0.50
	0.90	[26.13, 29.81]	[3.77, 6.52]	0.25
	0.99	[24.93, 31.01]	[3.34, 7.92]	0.00

Suppose that a sample of size 20 is available for the cylinder compressive strength f_c of a concrete according to Table 2. A normal distribution is assumed based on expert knowledge, and the parameters m_x and σ_x are determined as fuzzy values $\tilde{m}_x$ and $\tilde{\sigma}_x$. For this purpose interval estimations are applied. From the 20 measured values of the compressive strength the central confidence intervals for the confidence levels 0.50, 0.75, 0.90, and 0.99 are determined. Dependencies between the parameters are not taken into account. Additionally, common point estimations are used to specify crisp values for the expected value (as the mean value of the sample) and the standard deviation (based on the sample variance). The results (Table 2) are then taken as a basis for the specification of the parameters as fuzzy quantities. Membership values are assigned to the estimation results by subjective assessment. That is, the confidence intervals are interpreted as being α-level sets of the fuzzy values $\tilde{m}_x$ and $\tilde{\sigma}_x$; see Table 2. The mean values of the fuzzy numbers are taken from the point estimations. Eventually, the fuzzy quantities $\tilde{m}_x$ and $\tilde{\sigma}_x$ are obtained according to Fig. **??**. As dependencies between the parameters in the interval estimations are neglected, interaction between $\tilde{m}_x$ and $\tilde{\sigma}_x$ is not obtained.

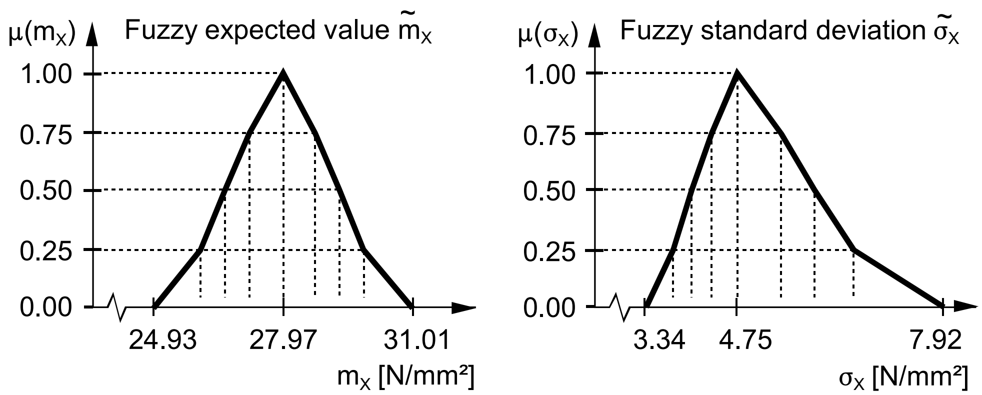

Figure 1. Fuzzy expected value $\tilde{m}_x$ and fuzzy standard deviation $\tilde{\sigma}_x$

3. Imprecise Sample Elements

Imprecision of sample elements may occur, for example, due to imprecise readings of (analog) measuring devices or as a reflection of imprecise individual care of personnel in tests. This imprecision can be expressed in form of fuzzy numbers for the measured values representing the sample elements. It is then possible to construct a fuzzy random quantity $\tilde{X}$ directly from the imprecise data material. The corresponding fuzzy parameters $\tilde{p}_t(\tilde{X})$ for the description of the fuzzy random quantity $\tilde{X}$ can be estimated based on statistical estimations and tests extended to deal with fuzzy arguments. This requires a proper application of fuzzy arithmetic in these algorithms. For a numerical evaluation, the fuzzy analysis based on α-level optimization according to [20] may be utilized. This framework enables an implementation of algorithms of mathematical statistics as the mapping model of a fuzzy analysis. Each fuzzy sample element is then treated as a fuzzy input quantity of the mapping model. The fuzzy result represents the sought parameter $\tilde{p}_t(\tilde{X})$.

As an example, the sample elements from Table 2 are assumed to possess an imprecision of $\pm 2\,\mathrm{N/mm^2}$ due to imprecise readings of the measuring device. This provides information for a modeling of the sample elements as fuzzy triangular numbers denoted by $\tilde{x}_i = < x_{i\,\mu=0l}, x_{i\,\mu=1}, x_{i\,\mu=0r} >$. The values from Table 2 are assessed with $\mu = 1$, from where the linear branches of the membership function decrease down to $\mu = 0$ at the points of the maximum deviation $\pm 2\,\mathrm{N/mm^2}$; see Table 3.

In order to compute the empirical parameters, common statistics (sample functions) are applied with the fuzzy values $\tilde{x}_i$ as arguments. The fuzzy sample mean is then obtained with

$$\tilde{\bar{x}} = \frac{1}{n}\sum_{i=1}^{n} \tilde{x}_i\,, \tag{1}$$

in which n is the sample size. The linearity of this mapping model leads to a fuzzy triangular number for the fuzzy sample mean, which is completely specified by the membership levels $\mu = 1$ and $\mu = 0$ as shown in Fig. 2, $\tilde{\bar{x}} = < 25.97, 27.97, 29.97 > \mathrm{N/mm^2}$. In contrast to this, the mapping model for computing the standard deviation of the sample is nonlinear and even non-monotonic,

$$\tilde{s}_x = \sqrt{\frac{1}{n-1}\left[\sum_{i=1}^{n}\tilde{x}_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n}\tilde{x}_i\right)^2\right]}. \quad (2)$$

This requires a more sophisticated evaluation technique. In the example, α-level optimization [20] is applied

Table III. Fuzzy sample elements of the cylinder compressive strength f_c of a concrete

Number i of fuzzy realization	Fuzzy compressive strength $\tilde{x}_i = \tilde{f}_{ci}[\mathrm{N/mm^2}]$	Number i of fuzzy realization	Fuzzy compressive strength $\tilde{x}_i = \tilde{f}_{ci}[\mathrm{N/mm^2}]$
1	$< 26.3, 28.3, 30.3 >$	11	$< 24.8, 26.8, 28.8 >$
2	$< 29.5, 31.5, 33.5 >$	12	$< 33.3, 35.3, 37.3 >$
3	$< 33.2, 35.2, 37.2 >$	13	$< 24.3, 26.3, 28.3 >$
4	$< 27.8, 29.8, 31.8 >$	14	$< 21.1, 23.1, 25.1 >$
5	$< 25.6, 27.6, 29.6 >$	15	$< 18.2, 20.2, 22.2 >$
6	$< 28.7, 30.7, 32.7 >$	16	$< 27.2, 29.2, 31.2 >$
7	$< 23.2, 25.2, 27.2 >$	17	$< 23.7, 25.7, 27.7 >$
8	$< 32.6, 34.6, 36.6 >$	18	$< 32.2, 34.2, 36.2 >$
9	$< 26.9, 28.9, 30.9 >$	19	$< 22.8, 24.8, 26.8 >$
10	$< 17.2, 19.2, 21.2 >$	20	$< 20.8, 22.8, 24.8 >$

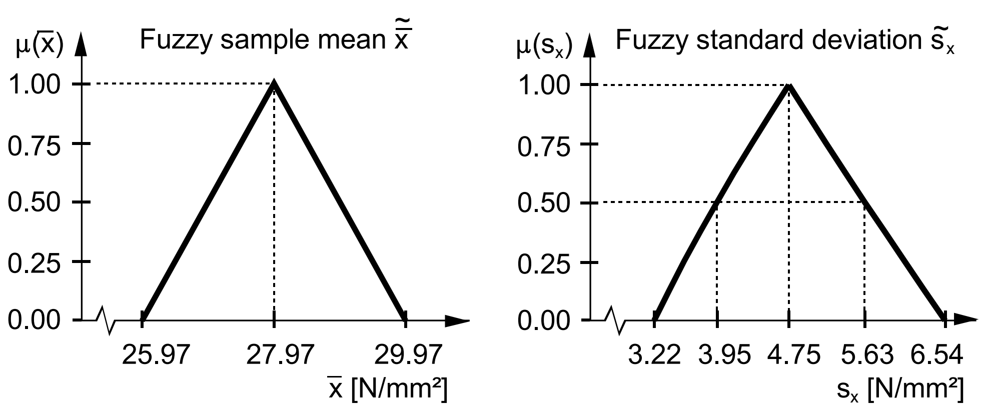

Figure 2. Fuzzy mean $\tilde{\bar{x}}$ and fuzzy standard deviation $\tilde{s}_x$ of the sample from Table 3

to evaluate Eq. (2). The membership function $\mu(s_x)$ is obtained with nonlinear branches; see Fig. 2.

The fuzzy sample elements $\tilde{x}_i$ enter Eq. (1) and Eq.(2), simultaneously. Thus, a relationship exists between the fuzzy sample mean and the fuzzy standard deviation of the sample. This is referred to as interaction between the fuzzy quantities $\tilde{\bar{x}}$ and $\tilde{s}_x$. This interaction is shown in Fig. 3 for the membership level $\alpha = 0$. Certain combinations of crisp values from $\tilde{\bar{x}}$ and $\tilde{s}_x$ cannot appear. An analytical or numerical determination of this interaction is, however, virtually excluded due to the tremendous computational effort even for a small sample. In the example a numerical approximation solution was determined with the aid of systematic and random-oriented simulations. The effect of the number of fuzzy realizations on the interaction relationship becomes apparent when only the first seven sample elements from Table 3 are considered; see Fig. 4. Not only the position but also the shape of the fuzzy set $\{\tilde{\bar{x}}, \tilde{s}_x\}$ shows a deviation from the illustration in Fig. 3. As a consequence of the same support widths of the fuzzy realizations $\tilde{x}_i$ the minimum and maximum sample means are, in each case, coupled with the same standard deviation of the sample. This property is lost in the general case. As demonstrated for $\tilde{\bar{x}}$ and $\tilde{s}_x$, interaction generally exists between all empirical parameters including the distribution type. The fact that the fuzzy realizations themselves may also be interactive may even lead to non-connected sets for the empirical parameters. Due to the numerical complications in the determination of the interaction, an approximation may be pursued. Or, the interaction may even be neglected; see Fig. 3. Although this means that non-justified parameter combinations are included and thus enter subsequent computations, the "exact" solution is completely contained in this approximation. The negligence of interaction leads to an envelope curve of those parameter combinations, which can actually appear.

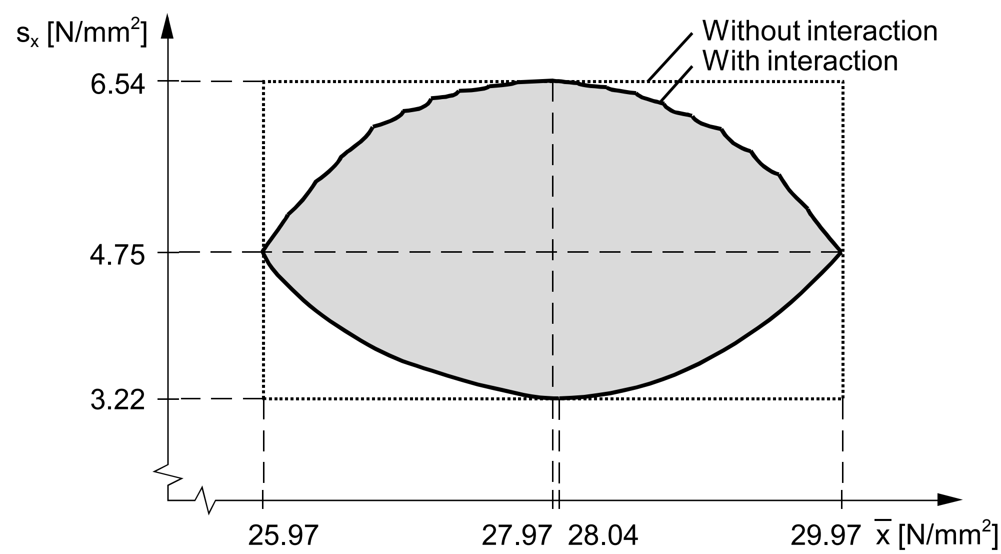

Figure 3. Numerical approximation of the interaction between the fuzzy sample mean $\tilde{\bar{x}}$ and the fuzzy standard deviation $\tilde{s}_x$ for the 20 fuzzy realizations from Table 3

The fuzzy parameters computed from the sample are the basis for the specification of the fuzzy probability distribution function needed for further processing of fuzzy random quantities in engineering computations. In the example, a normal distribution is assumed for the fuzzy random quantity. The functional parameters are then estimated by the fuzzy sample mean $\tilde{\bar{x}}$ as fuzzy expected value $\tilde{m}_x$, and by the fuzzy standard deviation $\tilde{s}_x$ of the sample as fuzzy standard deviation $\tilde{\sigma}_x$ of the fuzzy random quantity. The obtained fuzzy probability density function $\tilde{f}(x)$ and the fuzzy probability distribution function $\tilde{F}(x)$ are

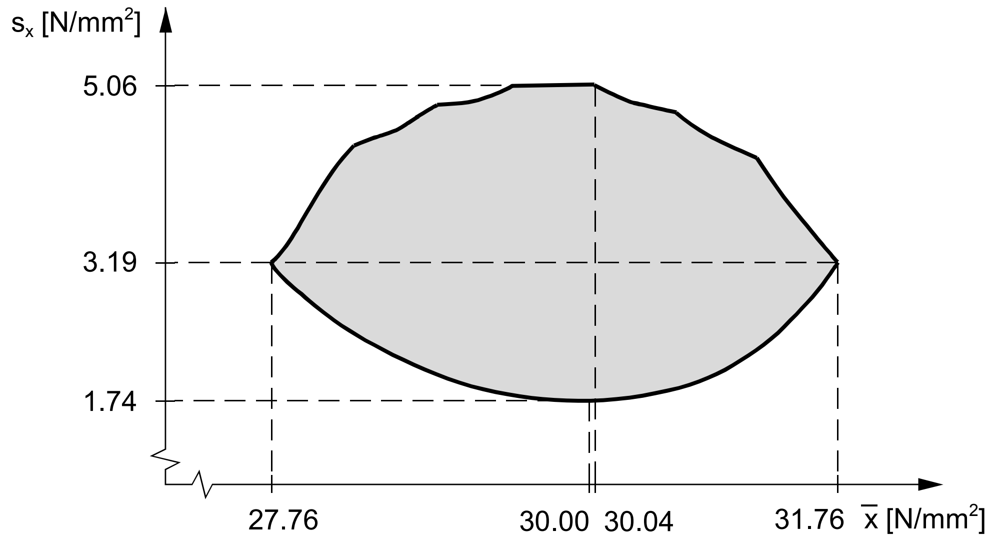

Figure 4. Numerical approximation of the interaction between $\tilde{\bar{x}}$ and $\tilde{s}_x$ for the first seven fuzzy realizations from Table 3

shown in Figs. 5 and 6, respectively. The illustrations show the functions with and without the consideration of the interaction between $\tilde{m}_x$ and $\tilde{\sigma}_x$. Negligence of the interaction between $\tilde{m}_x$ and $\tilde{\sigma}_x$ leads to envelope curves enclosing the exact fuzzy functions $\tilde{f}(x)$ and $\tilde{F}(x)$. The interaction between $\tilde{m}_x$ and $\tilde{\sigma}_x$ excludes the simultaneous occurrence of extrema of the expected value and standard deviation; see Fig. 3. This influences, in particular, the tails of the fuzzy functions $\tilde{f}(x)$ and $\tilde{F}(x)$. The probability mass in the tails is higher if the interaction is neglected. This leads to an overestimation of failure probabilities in a subsequent structural safety assessment. This overestimation is, however, not tremendous and leads to a slightly conservative safety assessment, which is rather welcome.

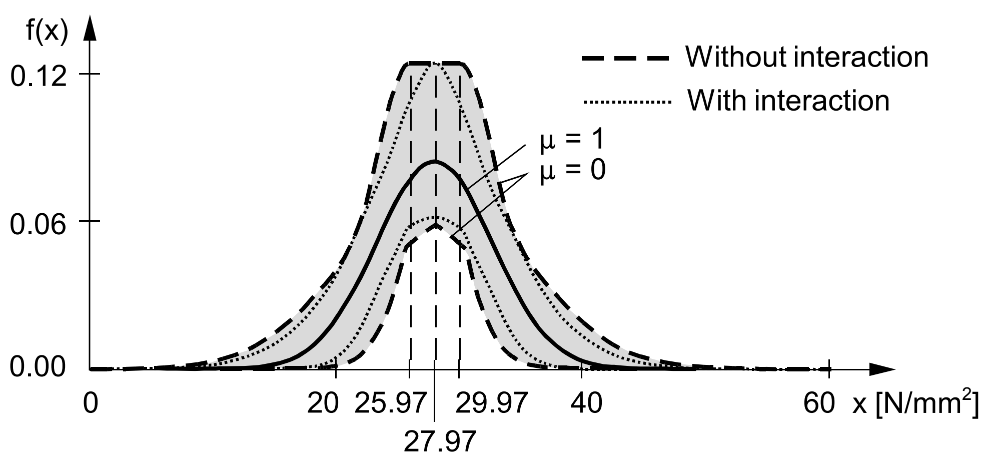

Figure 5. Fuzzy probability density function $\tilde{f}(x)$ with and without consideration of the interaction between $\tilde{m}_x$ and $\tilde{\sigma}_x$

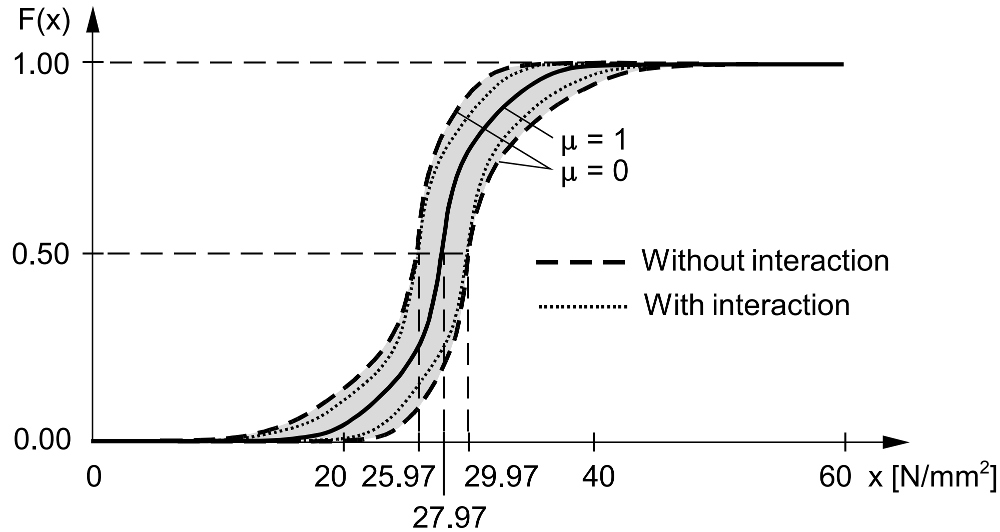

Figure 6. Fuzzy probability distribution function $\tilde{F}(x)$ with and without consideration of the interaction between $\tilde{m}_x$ and $\tilde{\sigma}_x$

4. Inconsistent Environmental Conditions and Expert Knowledge

This situation appears if the sample has been generated under varying environmental conditions. It then defies a traditional statistical evaluation and needs special treatment. The varying environmental conditions may include, for example, involvement of different manufacturers, changes in the type of aggregates / additives from different suppliers, varying hardening conditions (temperature, humidity), and variations in the motivation of the personnel. In those cases, expert knowledge is usually available to separate fuzziness and randomness present in the statistical data material. This separation can be realized by characterizing the environmental conditions with attributes such as a specific supplier for aggregates or a certain team of employees in the production process. Observed realizations with the same attributes are lumped together in a single *group*. These groups are subsets of the population. Each group of realizations with the same attributes is treated as a separate sample. These samples can then be evaluated using statistical methods as they comply with the preconditions in form of constant environmental conditions. The statistical evaluation yields empirical parameter values including a distribution type for each group. For all groups the set $\underline{S}$ of statistical propositions is obtained. Each element of $\underline{S}$ is assigned to a subset of the population. Hence, the set $\underline{S}$ describes the set of real random quantities contained in the observed realizations. The differences between the elements of the set $\underline{S}$ represent imprecision, which may be modeled as fuzziness of the population. The elements contained in $\underline{S}$ and, thus, the associated real random quantities may be assessed with membership values. This results in the fuzzy set $\tilde{\underline{S}}$. The real random quantities together with their membership values form a fuzzy random quantity, which is described by $\tilde{\underline{S}}$.

The fuzzy set $\tilde{\underline{S}}$ can be constructed in parametric or in a non-parametric manner. The parametric construction of $\tilde{\underline{S}}$ involves a distribution assumption from expert knowledge. Then, the membership functions of the empirical distribution parameters may be constructed using histograms. In the non-parametric construction of $\tilde{\underline{S}}$ empirical distribution functions are used, and a direct fuzzification of the probability distribution function curve is pursued.

Parametric Quantification It is presumed that the groups of sample elements with same attributes and their corresponding empirical parameters are known. The parameter values constitute a sample for which a histogram is constructed. The parameter value is plotted along the abscissa, which is subdivided into subsets. In the normal manner the number of sample elements, which is the number of empirical parameter values, per subset is plotted on the ordinate. Then, the histogram can be used as a basis for constructing the membership function of the respective fuzzy parameter.

As an example, let specimens of a concrete be available from different concrete plants. Tests are carried out to measure the cylinder compressive strength f_c. The specimens are labeled, and the concrete plant and work team are registered. Specimens with the same identification (same attributes) are each lumped together in a group. In the example, twelve groups with a different number of specimens (sample size) are identified. By this means, randomness and fuzziness are separated. The statistical evaluation of the measured cylinder compressive strength f_c yields empirical parameters for each group. The sample mean $\overline{x}$ and the standard deviation s_x of the samples are computed; see Table 4.

Table IV. Sample mean $\overline{x}$ and standard deviation s_x of the cylinder compressive strength f_c of the concrete for twelve groups of specimens (twelve samples)

Label of group	Sample size	Sample mean $\overline{x}[N/mm^2]$	Standard deviation $s_x[N/mm^2]$
1	54	27.3	5.3
2	48	26.6	4.9
3	42	29.2	4.2
4	38	31.4	3.8
5	44	28.3	5.6
6	48	29.4	3.2
7	55	26.4	5.0
8	47	30.1	4.6
9	64	28.3	5.9
10	53	27.9	3.8
11	75	29.6	6.3
12	52	27.8	4.7

The values listed in Table 4 are used to construct histograms for the sample mean $\overline{x}$ and the standard deviation s_x of the samples; see Fig. 7. The chosen subset widths are $1.0\,N/mm^2$ for $\overline{x}$ and $0.75\,N/mm^2$ for s_x. Each of the empirical parameters is modeled using fuzzy triangular numbers. The method of least squares is applied to determine the linear membership functions. The derived fuzzification suggestions are shown in Fig. 7.

Due to the fact that the values $\overline{x}$ and s_x for each group originate from the same sample, interaction exists between the fuzzy quantities $\tilde{\overline{x}}$ and $\tilde{s}_x$. Analog to the analysis of stochastic dependencies between random variables, the interaction relationship may be determined by evaluating the value pairs $(\overline{x}, s_x)$ obtained. These pairs are plotted in a coordinate system, and the interaction relationship is estimated for different membership levels. This procedure is illustrated in Fig. 8 for the membership level $\alpha = 0$. Assuming a normal distribution, the empirical fuzzy parameters $\tilde{\overline{x}}$ and $\tilde{s}_x$ are adopted as the fuzzy distribution parameters

$\tilde{m}_x$ and $\tilde{\sigma}_x$, respectively, of the fuzzy probability distribution. If the assumed distribution type is different for the individual groups, this may be accounted for with a compound distribution and fuzzy parameter for the mixing ratio.

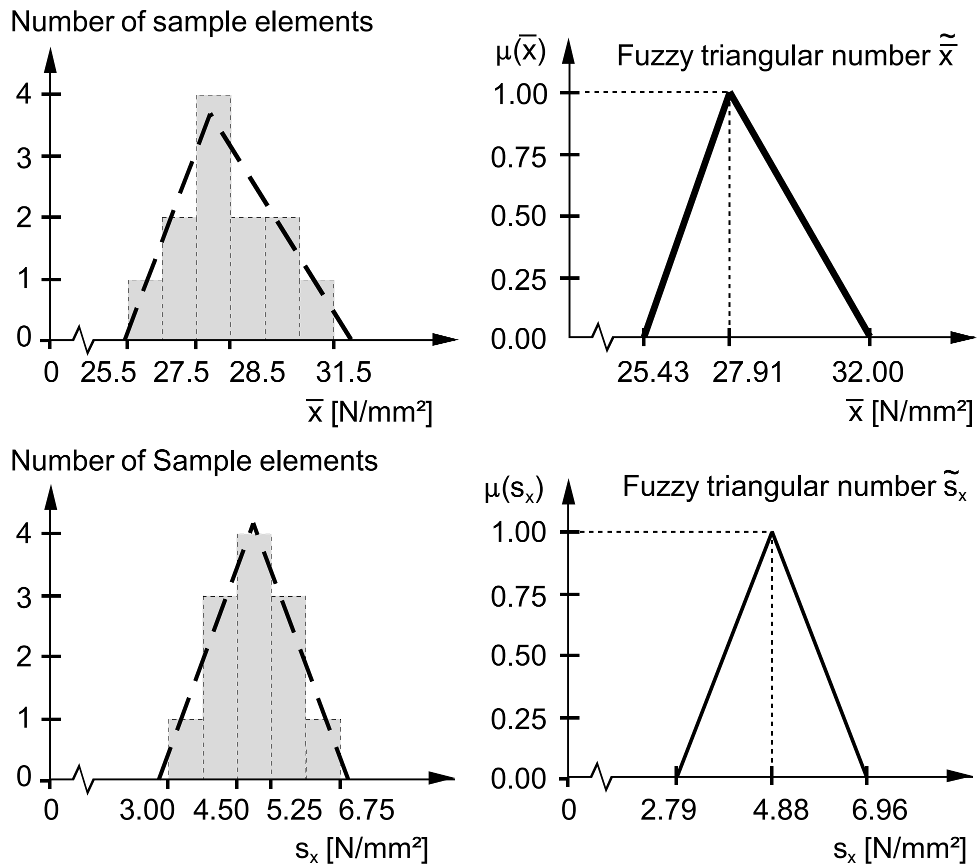

Figure 7. Histograms and fuzzification of the sample mean $\overline{x}$ and the standard deviation s_x assigned to the groups (samples) of the cylinder compressive strength f_c

Non-parametric Quantification The starting point is again the separation of randomness and fuzziness by constructing groups of observed realizations. Then, empirical distribution functions are constructed for the individual groups. The set of empirical distribution functions for all groups is then taken as the basis to determine fuzzy quantities for the functional values of an overall empirical distribution function.

The example from the parametric quantification is reused for demonstration. For each group, a histogram is constructed from the realizations to determine an empirical distribution function. The subset widths and the subset positioning on the abscissa must be the same for all histograms for all groups. The subsets are defined as half-closed intervals $[x_l, x_r)$ on the real number line. The number of observed realizations in the subsets is generally different for the individual groups. The histograms for the first two groups from Table 4 are shown in Fig. 9.

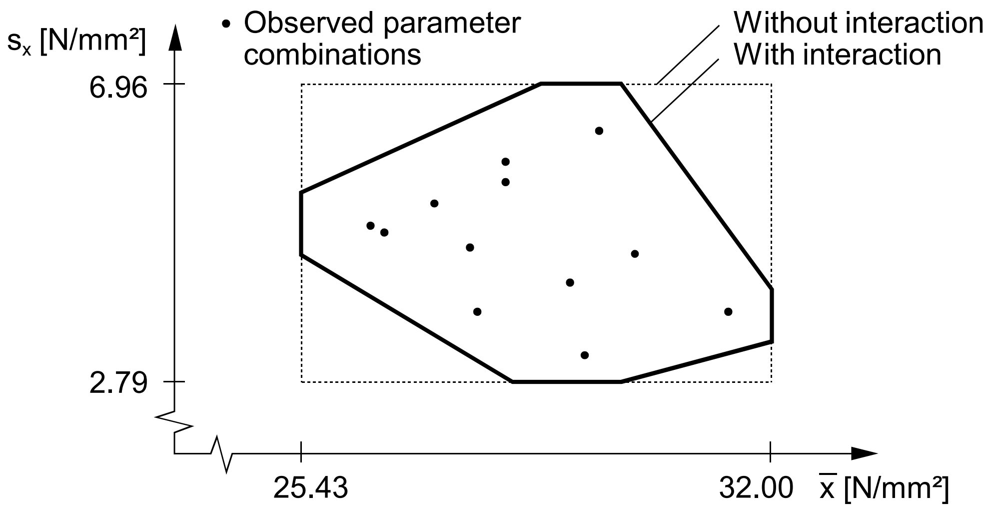

Figure 8. Estimation of the interaction between $\tilde{\bar{x}}$ and $\tilde{s}_x$

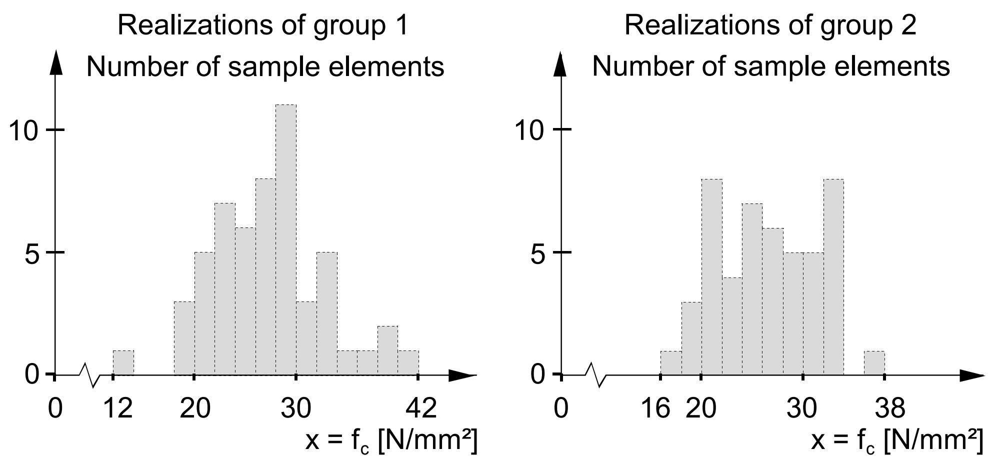

Figure 9. Histograms for the realizations of groups 1 and 2 from Table 4

For each group the empirical probability distribution function

$$F_i^e(x) = \frac{n_{i,k}(x)}{n_i} \tag{3}$$

is developed from the corresponding histogram. In the above, i denotes the group number, n_i is the number of all elements (realizations) in group i, and $n_{i,k}(x)$ is the number of those elements k (in group i), whose values x_k are smaller than x. The values x of the observed realizations are determined by the left-hand subset boundaries (that is, by the x_l of the half-closed intervals $[x_l, x_r))$ in the histograms; these mark discrete positions on the abscissa. The evaluation of all groups yields a bunch of discrete empirical distribution functions. The functional values $F_i^e(x = f_c)$ are listed in Table 4.

Table V. Functional values of the empirical distribution functions $F_i^e(x = f_c)$ for all groups i of specimens

$x = f_c$ [N/mm^2]	Group i: 1	2	3	4	5	6	7	8	9	10	11	12
12	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
14	0.019	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.016	0.000	0.000	0.000
16	0.019	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.016	0.000	0.000	0.000
18	0.019	0.021	0.000	0.000	0.000	0.000	0.036	0.000	0.016	0.000	0.013	0.000
20	0.074	0.083	0.000	0.000	0.023	0.000	0.091	0.000	0.063	0.000	0.067	0.000
22	0.167	0.250	0.000	0.000	0.114	0.000	0.273	0.043	0.172	0.094	0.120	0.096
24	0.296	0.333	0.071	0.026	0.295	0.042	0.327	0.064	0.266	0.170	0.227	0.231
26	0.407	0.479	0.262	0.105	0.409	0.167	0.436	0.191	0.328	0.283	0.320	0.346
28	0.556	0.604	0.476	0.184	0.523	0.354	0.655	0.340	0.422	0.509	0.400	0.577
30	0.759	0.708	0.595	0.395	0.705	0.563	0.764	0.532	0.625	0.717	0.507	0.731
32	0.815	0.813	0.786	0.632	0.750	0.833	0.855	0.702	0.766	0.887	0.653	0.788
34	0.907	0.979	0.810	0.711	0.795	0.917	0.927	0.766	0.844	0.962	0.733	0.904
36	0.926	1.000	0.905	0.816	0.909	0.979	0.964	0.830	0.922	0.981	0.827	0.923
38	0.944	1.000	1.000	1.000	0.932	1.000	1.000	0.979	0.969	0.981	0.933	0.962
40	0.981	1.000	1.000	1.000	0.977	1.000	1.000	0.979	0.969	1.000	0.947	1.000
42	1.000	1.000	1.000	1.000	0.977	1.000	1.000	1.000	0.969	1.000	0.973	1.000
44	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.984	1.000	0.987	1.000
46	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.987	1.000
48	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

For each discrete value x from Table 4 the functional values $F^e(x)$ are taken as a basis to model fuzzy functional values $\tilde{F}^e(x)$ to cover all groups at once. At each (discrete) position x a histogram is constructed using the functional values of the empirical distribution functions. The abscissa is subdivided into suitable subsets in the interval $[0, 1)$; and the number of functional values assigned to each subset is plotted on the ordinate. Than, fuzzy numbers are generated from the histograms by simple approximation schemes such as least squares algorithm. In this generation process the properties of the probability measure must be observed. In the present case fuzzy triangular numbers and fuzzy numbers with a polygonal membership function are chosen. The fuzzification process is shown in Fig. 10 for three selected values $x = f_c$. The fuzzification results for all $x = f_c$ are listed in Table 4. The interval bounds of the support as well as the mean value are indicated for each fuzzy probability $\tilde{F}^e(x)$. The obtained fuzzy probabilities $\tilde{F}^e(x)$ for discrete $x = f_c$ are functional values of the sought fuzzy probability distribution function $\tilde{F}(x)$.

This non-parametric representation can finally be replaced by a parametric fuzzy probabilistic model in the form of an envelope. For this purpose, different membership levels α are considered for the determination of fuzzy parameters of the fuzzy probability distribution and for the description of the distribution type. The aim is to determine bounding distribution functions of the fuzzy random variable for each membership level. The entirety of all included probabilistic models then reflects the sought fuzzy probability distribution.

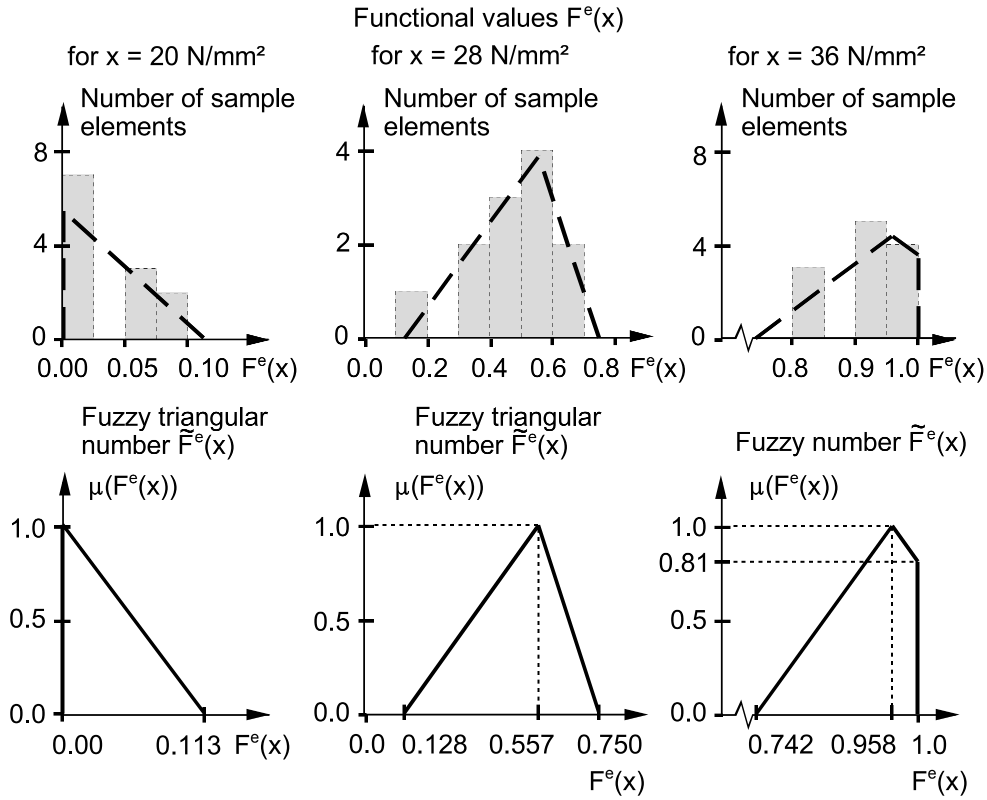

Figure 10. Histograms and membership functions of the functional values of the empirical distribution function $F^e(x)$ for $x = f_c = 20\,N/mm^2$, $x = f_c = 28\,N/mm^2$, and $x = f_c = 36\,N/mm^2$

In this example a compound distribution comprised of a normal distribution (ND) and a logarithmic normal distribution (LND) with a constant ratio of components is adopted. It is assumed that the expected value and standard deviation are the same for both distributions; the minimum value of the component logarithmic normal distribution is specified to be $x_0 = 5\,N/mm^2$. The expected value, standard deviation, and ratio of components are chosen to be free fuzzy parameters of the compound distribution

$$\tilde{F}(x) = \tilde{a} \cdot \tilde{F}^{NV}(x) + (1 - \tilde{a}) \cdot \tilde{F}^{LNV}(x)\,. \tag{4}$$

The subsequent evaluation is restricted to the membership levels $\alpha = 0$ and $\alpha = 1$. The free parameters required for approximating the distribution functions of the originals are determined by the method of least squares. The distribution function $F_1(x)$ for the membership level $\alpha = 1$ is obtained from the values of $F^e_1(x)$. The boundaries of the membership level $\alpha = 0$ are obtained in each case from all values of $F^e_{0l}(x)$ and $F^e_{0r}(x)$, respectively. The following constraints are taken into account:

– All $F^e_{0l}(x) > 0$ lie above the approximation function $F_{0l}(x)$

– All $F^e_{0r}(x) < 1$ lie below the approximation function $F_{0r}(x)$

Table VI. Support bounds $F^e_{0l}(x)$ and $F^e_{0r}(x)$, and mean values $F^e_1(x)$ of the fuzzy probability $\tilde{F}^e(x)$ for all $x = f_c[N/mm^2]$ from Table 4

x	$F^e_{0l}(x)$	$F^e_1(x)$	$F^e_{0r}(x)$	x	$F^e_{0l}(x)$	$F^e_1(x)$	$F^e_{0r}(x)$
12	0.000	0.000	0.000	32	0.603	0.799	0.975
14	0.000	0.000	0.018	34	0.652	0.925	1.000
16	0.000	0.000	0.018	36	0.742	0.958	1.000
18	0.000	0.000	0.035	38	0.913	1.000	1.000
20	0.000	0.000	0.113	40	0.949	1.000	1.000
22	0.000	0.000	0.369	42	0.966	1.000	1.000
24	0.000	0.283	0.417	44	0.983	1.000	1.000
26	0.025	0.358	0.492	46	0.984	1.000	1.000
28	0.128	0.557	0.750	48	1.000	1.000	1.000
30	0.331	0.763	0.825	–	–	–	–

The following values are obtained for the free distribution parameters and the functional parameter *a* of the implemented distribution function:

- Approximation of $F^e_1(x)$: $m_x = 27.66 N/mm^2$, $\sigma_x = 4.34 N/mm^2$, $a = 0.00$
- Approximation of $F^e_{0l}(x)$: $m_x = 34.29 N/mm^2$, $\sigma_x = 4.81 N/mm^2$, $a = 0.00$
- Approximation of $F^e_{0r}(x)$: $m_x = 23.30 N/mm^2$, $\sigma_x = 4.44 N/mm^2$, $a = 1.00$

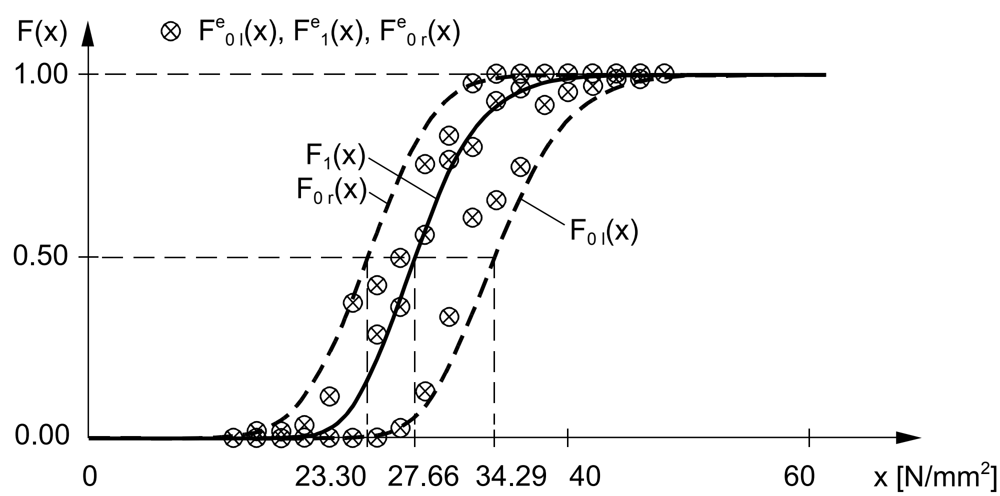

Figure 11. Functional values of the empirical probability distribution functions $F^e_1(x)$, $F^e_{0l}(x)$, and $F^e_{0r}(x)$, as well as the approximation functions $F_1(x)$, $F_{0l}(x)$, and $F_{0r}(x)$

The computed distribution functions $F_1(x)$, $F_{01}(x)$, and $F_{0r}(x)$ are shown in Fig. 11 together with the adopted functional values of the empirical distribution function from Table 4.

The fuzzy distribution parameters and the fuzzy functional parameter ã of the sought fuzzy probability distribution according to Eq. (4) may be expressed as fuzzy triangular numbers (confined to $\alpha = 0$ and $\alpha = 1$):

– $\tilde{m}_x =< 23.30, 27.66, 34.29 >$ N/mm^2,
– $\tilde{\sigma}_x =< 4.34, 4.34, 4.81 >$ N/mm^2, and
– $\tilde{a} =< 0.00, 0.00, 1.00 >$.

The interaction relationship between $\tilde{m}_x$, $\tilde{\sigma}_x$ and ã may be determined numerically (Sect. 3), or may be approximately estimated on the basis of the available information. A possible estimation of the interaction is shown in Fig. 12.

In the example, the interaction between $\tilde{m}_x$, $\tilde{\sigma}_x$ and ã has only a very slight effect, and may be neglected without a significant effect. The fuzzy probability density functions and the fuzzy probability distribution functions are compared in Figs. 13 and 14, with and without consideration of interaction. The approximation functions $F_{01}(x)$ and $F_{0r}(x)$ as well as the corresponding probability density functions $f_{01}(x)$ and $f_{0r}(x)$ are also shown in the figures.

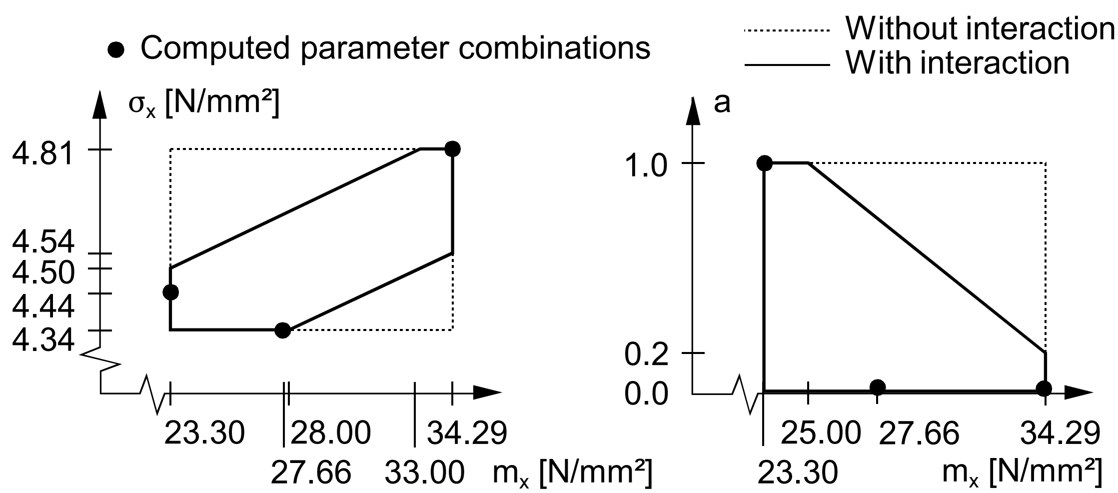

Figure 12. Estimation of the interaction between $\tilde{m}_x$, $\tilde{\sigma}_x$ and ã

5. Conclusions

Inconsistent data represent a common case of available information in civil engineering practice. These data must be properly evaluated and described numerically to obtain realistic results in a subsequent structural analysis, safety assessment or structural design. The evaluation of inconsistent data is, however, problematic.

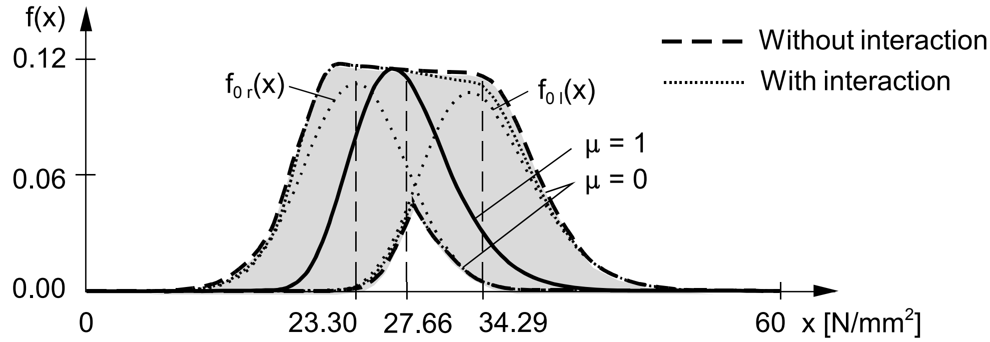

Figure 13. Fuzzy probability density function $\tilde{f}(x)$ with and without consideration of interaction between $\tilde{m}_x$, $\tilde{\sigma}_x$ and $\tilde{a}$; probability density functions $f_{0l}(x)$ and $f_{0r}(x)$ belonging to $F_{0l}(x)$ and $F_{0r}(x)$

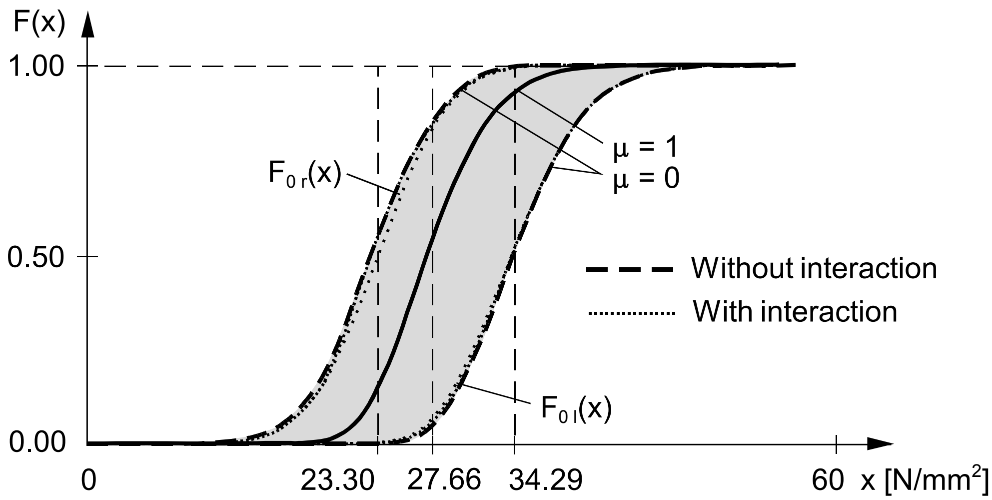

Figure 14. Fuzzy probability distribution function $\tilde{F}(x)$ with and without consideration of interaction between $\tilde{m}_x$, $\tilde{\sigma}_x$ and $\tilde{a}$; probability distribution functions $F_{0l}(x)$ and $F_{0r}(x)$

Stochastic uncertainty and imprecision appear simultaneously and in various configurations. For a proper treatment of this type of information, the model fuzzy randomness is proposed. This enables a separate and simultaneous treatment of statistical uncertainty and imprecision. Due to the variety of possible forms of available information, a general quantification algorithm cannot be formulated. The quantification has to be realized according to the conditions in each particular case. In the paper quantification guidelines for three selected typical cases of inconsistent data in civil engineering were presented by way of examples. Algorithms from traditional statistics have been utilized and combined with fuzzy methods for the inclusion of expert knowledge. The quantification results reflect the stochastic uncertainty and the imprecision of the available information in form of a fuzzy probability. This represents an envelope of all real-valued probabilistic models which meet the available information.

Further developments are focused on the development of a hybrid quantification algorithm for inconsistent data, which includes, simultaneously, more components beyond traditional statistics and fuzzy methods to extend the spectrum of cases covered and to further improve the quality of the quantification results. This leads, eventually, to a minimization of risks due to modeling errors and associated misinterpretations of structural behavior and safety.

Acknowledgements

The author gratefully acknowledges the financial support by National University of Singapore through the Ministry of Education Academic Research Fund.

References

1. Ayyub, B.: 1998, *Uncertainty Modeling and Analysis in Civil Engineering*. Boston London New York: CRC Press.
2. Bandemer, H. and S. Gottwald: 1995, *Fuzzy Sets, Fuzzy Logic Fuzzy Methods With Applications*. Wiley.
3. Bandemer, H. and W. Näther: 1992, *Fuzzy Data Analysis*. Dordrecht: Kluwer Academic Publishers.
4. Beer, M. and M. Liebscher: 2007, 'Designing Robust Structures – A Nonlinear Simulation Based Approach'. *Special Issue of Computers & Structures*. in press.
5. Ben-Haim, Y.: 2004, 'Uncertainty, probability and information-gaps'. *Reliability Engineering & System Safety* **85**(1-3), 249–266.
6. Ben-Haim, Y. and I. Elishakoff: 1990, *Convex Models of Uncertainty in Applied Mechanics*. Amsterdam: Elsevier.
7. Elishakoff, I.: 1999, *Whys and Hows in Uncertainty Modelling Probability, Fuzziness and Anti-Optimization*. Wien, New York: Springer.
8. England, J., J. Agarwal, and D. Blockley, 'The vulnerability of structures to unforeseen events'. *Special Issue of Computers & Structures*. , submitted.
9. Farkas, L., D. Moens, D. Vandepitte, and W. Desmet, 'Application of fuzzy numerical techniques for product performance analysis in the conceptual and preliminary design stage'. *Special Issue of Computers & Structures*. , submitted.
10. Fellin, W., H. Lessmann, M. Oberguggenberger, and R. Vieider (eds.): 2005, *Analyzing Uncertainty in Civil Engineering*. Berlin Heidelberg New York: Springer.
11. Hall, J. W., E. Rubio, and M. Anderson: 2004, 'Random sets of probability measures in slope hydrology and stability analysis'. *Special Issue of ZAMM - Zeitschrift für Angewandte Mathematik und Mechanik* **84**(10–11), 710–720.
12. Helton, J., J. Johnson, W. Oberkampf, and C. Sallaberry: 2006, 'Sensitivity analysis in conjunction with evidence theory representations of epistemic uncertainty'. *Reliability Engineering & System Safety* **91**, 1414–1434.
13. Helton, J. C. and D. E. Burmaster (eds.): 1996, *Special Issue on the Treatment of Aleatory and Epistemic Uncertainty*, Vol. 54 of *Reliability Engineering & System Safety*.
14. Helton, J. C. and W. L. Oberkampf (eds.): 2004, *Special Issue on Alternative Representations of Epistemic Uncertainty*, Vol. 85 of *Reliability Engineering & System Safety*.
15. Klir, G. J.: 2006, *Uncertainty and information : foundations of generalized information theory*. Hoboken: Wiley-Interscience.
16. Kreinovich, V. and S. A. Ferson: 2004, 'A new Cauchy-based black-box technique for uncertainty in risk analysis'. *Reliability Engineering & System Safety* **85**(1-3), 267–279.
17. Kruse, R. and K. Meyer: 1987, *Statistics with Vague Data*. Dordrecht: Reidel.
18. Möller, B. and M. Beer, 'Engineering Computation Under Uncertainty – Capabilities of Non-Traditional Models'. *Special issue of Computers & Structures*. in press.
19. Möller, B. and M. Beer: 2004, *Fuzzy Randomness – Uncertainty in Civil Engineering and Computational Mechanics*. Berlin: Springer.
20. Möller, B., W. Graf, and M. Beer: 2000, 'Fuzzy structural analysis using alpha-level optimization'. *Computational Mechanics* **26**, 547–565.

21. Muhanna, R. and R. Mullen: 1999, 'Formulation of Fuzzy Finite Element Methods for Solid Mechanics Problems'. *Computer-Aided Civil and Infrastructure Engineering* **14**, 107–117.
22. Muhanna, R. L., R. L. Mullen, and H. Zhang: 2007, 'Interval Finite Element as a Basis for Generalized Models of Uncertainty in Engineering Mechanics'. *Journal of Reliable Computing* **13**(2), 173–194.
23. Neumaier, A.: 2004, 'Clouds, Fuzzy Sets, and Probability Intervals'. *Reliable Computing* **10**(4), 249–272.
24. Pedrycz, W., A. Skowron, and V. Kreinovich (eds.): 2008, *Handbook of Granular Computing*. New York: Wiley.
25. Ross, T. J.: 2004, *Fuzzy Logic with Engineering Applications*. Wiley, 2nd edition.
26. Schenk, C. A. and G. I. Schuëller: 2005, *Uncertainty Assessment of Large Finite Element Systems*. Berlin Heidelberg: Springer.
27. Spanos, P. D. and G. Deodatis (eds.): 2007, *Computational Stochastic Mechanics*. Rotterdam: Millpress.
28. Viertl, R.: 1996, *Statistical Methods for Non-Precise Data*. Boca Raton New York London Tokyo: CRC Press.
29. Viertl, R. and W. Trutschnig: 2006, *Fuzzy histograms and fuzzy probability distributions*.
30. Zimmermann, H.-J.: 1992, *Fuzzy set theory and its applications*. Boston London: Kluwer Academic Publishers.