18th International Conference on the Application of Computer
Science and Mathematics in Architecture and Civil Engineering
K. Gürlebeck and C. Könke (eds.)
Weimar, Germany, 07–09 July 2009

ESTIMATING UNCERTAINTIES FROM INACCURATE MEASUREMENT DATA USING MAXIMUM ENTROPY DISTRIBUTIONS

T. Most*

**Research Training Group 1462*
Bauhaus-Universität Weimar, Germany
E-mail: thomas.most@uni-weimar.de

Keywords: Maximum entropy, general distribution function, bootstrapping

Abstract. *Modern engineering design often considers uncertainties in geometrical and material parameters and in the loading conditions. Based on initial assumptions on the stochastic properties as mean values, standard deviations and the distribution functions of these uncertain parameters a probabilistic analysis is carried out. In many application fields probabilities of the exceedance of failure criteria are computed. The out-coming failure probability is strongly dependent on the initial assumptions on the random variable properties. Measurements are always more or less inaccurate data due to varying environmental conditions during the measurement procedure. Furthermore the estimation of stochastic properties from a limited number of realisation also causes uncertainties in these quantities. Thus the assumption of exactly known stochastic properties by neglecting these uncertainties may not lead to very useful probabilistic measures in a design process [1].*

In this paper we assume the stochastic properties of a random variable as uncertain quantities caused by so-called epistemic uncertainties. Instead of predefined distribution types we use the maximum entropy distribution [2] which enables the description of a wide range of distribution functions based on the first four stochastic moments. These moments are taken again as random variables to model the epistemic scatter in the stochastic assumptions. The main point of this paper is the discussion on the estimation of these uncertain stochastic properties based on inaccurate measurements. We investigate the bootstrap algorithm [3] for its applicability to quantify the uncertainties in the stochastic properties considering imprecise measurement data. Based on the obtained estimates we apply standard stochastic analysis on a simple example to demonstrate the difference and the necessity of the proposed approach.

1 INTRODUCTION

In practical applications of probabilistic models generally a very limited number of observations is available to choose a suitable distribution type and to estimate the corresponding stochastic parameters. Due to the small number of samples and varying measurement conditions these distributions can not be estimated exactly. Probabilistic measures as the failure probability are strongly dependent on the initial assumptions on these distributions. Thus the assumption of exactly known stochastic properties by neglecting these uncertainties may not lead to very useful results in a design process as shown by [1]. Thus in this paper we consider these uncertain distributions by using maximum entropy distributions based on uncertain stochastic moments. In many studies it has been shown that this formulation is very flexible to model different distribution types. We quantify the parameter uncertainties by the bootstrap method and use a random variable description of the stochastic parameters to estimate the resulting variation of the failure probability.

2 MAXIMUM ENTROPY DISTRIBUTION

2.1 Determination of density function

Based on the entropy principle proposed by [4] entropy distributions are defined to be those which maximize the information entropy measure

$$H = -\int_{D_x} f_X(x) \log(f_X(x)) dx \tag{1}$$

where f_X is the probability density function and D_x is the domain of the random variable X. Introducing $m+1$ moment constraints

$$\mu_i' = E[X^i] = \int_{D_x} x^i f_X(x) dx \quad i = 0, 1, \dots, m \tag{2}$$

with $\mu_0' = 1$, such an entropy distribution can be expressed as

$$f_X(x) = \exp\left(\lambda_0 + \sum_{i=1}^{m} \lambda_i x^i\right). \tag{3}$$

In many studies e.g. in [5], [6] and [2] it was shown that the first four moments are sufficient to describe a wide range of distribution types. The formulation using absolute moments in Equation 2 can be modified for the central moments

$$\mu_i = E[(X - \bar{X})^i] = \int_{D_x} (x - \bar{X})^i f_X(x) dx, \quad i = 1, \dots, m \tag{4}$$

where the mean value $\bar{X}$ of a random variable is its first absolute moment.

$$\bar{X} = E[X] = \int_{D_x} x f_X(x) dx. \tag{5}$$

The first central moment μ_1 is zero and the second central moment is the variance σ_X^2.

$$\sigma_X^2 = E[(X - \bar{X})^2] = \int_{D_x} (x - \bar{X})^2 f_X(x) dx. \tag{6}$$

According to [5] the entropy distribution based on central moment constraints reads

$$f_X(x) = \exp\left(\nu_0 + \sum_{i=1}^{m} \nu_i (x - \bar{X})^i\right) \tag{7}$$

where $\exp(\nu_0) = 1/c$ is a constant normalizing the area under the density function. If a standardized random variable is defined

$$Y = \frac{X - \bar{X}}{\sigma_X} \tag{8}$$

the maximum entropy distribution can be obtained from the standardized central moment constraints

$$k_i = \frac{\mu_i}{\sigma_X^i}, \quad i = 1, \ldots, m, \quad k_1 = 0, \quad k_2 = 1, \tag{9}$$

where the third and fourth standardized central moments are the skewness γ_1 and the the kurtosis γ_2

$$\gamma_1 = k_3 = \frac{\mu_3}{\sigma_X^3}, \quad \gamma_2 = k_4 = \frac{\mu_4}{\sigma_X^4}. \tag{10}$$

From this standardized constraints the distribution parameters can be obtained very efficiently as shown in [2] and [7]. The final maximum entropy distribution is than obtained for the standardized random variable Y

$$f_Y(y) = \frac{1}{c'} \cdot \exp\left(\sum_{i=1}^{m} \nu'_i y^i\right), \tag{11}$$

and finally for the original random variable X as

$$\begin{aligned} f_X(x) &= \frac{1}{c' \cdot \sigma_X} \cdot \exp\left(\sum_{i=1}^{m} \nu'_i \left(\frac{x - \bar{X}}{\sigma_X}\right)^i\right) \\ &= \frac{1}{c} \cdot \exp\left(\sum_{i=1}^{m} \nu_i (x - \bar{X})^i\right), \\ c &= c' \cdot \sigma_X, \quad \nu_i = \frac{\nu'_i}{\sigma_X^i}. \end{aligned} \tag{12}$$

Special types of the maximum entropy distribution are the uniform distribution ($\nu'_1 = \nu'_2 = \nu'_3 = \nu'_4 = 0$, $c' = \sqrt{12}$), the exponential distribution ($\nu'_1 = -1$, $\nu'_2 = \nu'_3 = \nu'_4 = 0$, $c' = e$) and the normal distribution ($\nu'_2 = -0.5$, $\nu'_1 = \nu'_3 = \nu'_4 = 0$, $c' = \sqrt{2\pi}$).

In the Figures 1, 2 and 3 the density functions of maximum entropy distributions based on the first four moments compared to these of other common distribution types are shown. The figures indicate a very good agreement for the log-normal distribution with a coefficient of variation of 20% and a sufficient agreement for the log-normal and Gumbel distribution with higher variation. For the Rayleigh distribution the agreement is sufficient for positive values above 0.5, but for smaller values a remarkable deviation is observed. Apart from this case the maximum entropy distribution allows a very flexible representation of random variable distributions.

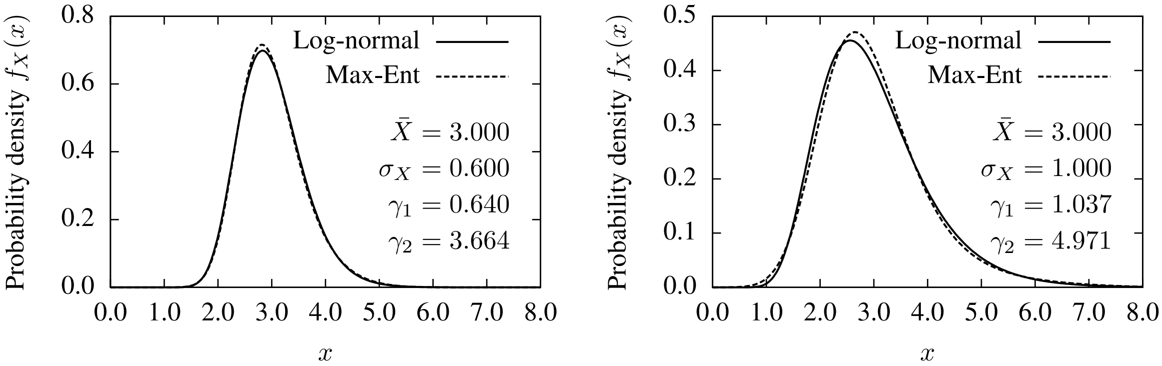


Figure 1: Log-normal and corresponding maximum entropy distributions for different standard deviations

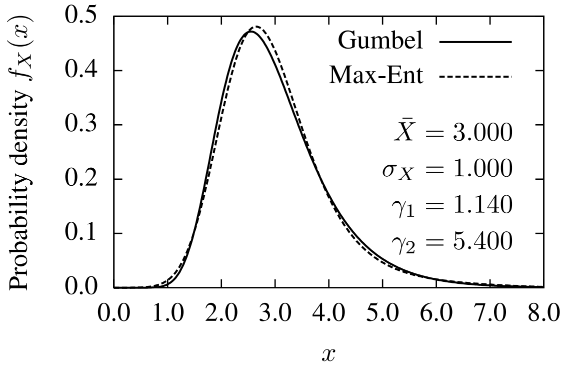


Figure 2: Gumbel and Max-Ent distributions

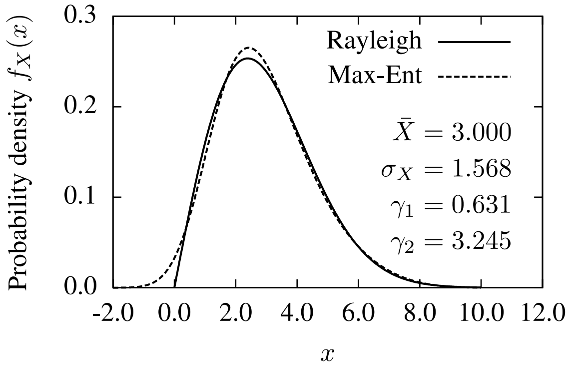


Figure 3: Rayleigh and Max-Ent distributions

2.2 Sampling of realizations

If we want to apply the obtained maximum entropy density function for a probabilistic analysis the generation of discrete samples out of the given distribution is necessary. For an arbitrary type of distributions we can perform this procedure for example with the Metropolis-Hastings algorithm [8] by using only the density function itself. If we want to apply more sophisticated sampling procedures, e.g. Latin Hypercube Sampling, the easiest way to generate samples is to transform samples from a uniform or normal distribution by the inverse cumulative distribution function.

For the presented maximum entropy distribution the cumulative distribution function is given as

$$F_X(x) = \int_{-\infty}^{x} f_X(\tau) d\tau = \int_{-\infty}^{x} \frac{1}{c} \cdot \exp\left(\sum_{i=1}^{m} \nu_i (\tau - \bar{X})^i \right) d\tau, \quad i = 1, 2, 3, 4. \tag{13}$$

For the general case the integral in Equation 13 can not be solved analytically. Thus we apply a numerical integration to obtain F_X. Based on a standard normally distributed variable Z discrete samples x_i of the maximum entropy variable X can be obtained as

$$x_i = F_X^{-1}[\Phi(z_i)]. \tag{14}$$

If we consider a pice-wise linear cumulative distribution function obtained by numerical integration, the values of the inverse distribution function F_X^{-1} can be directly calculated.

2.3 Extension for multivariate distributions

An arbitrary number of random variables can be arranged in a random vector

$$\mathbf{X} = [X_1, X_2, \dots, X_n]^T \tag{15}$$

with the mean value vector

$$\bar{\mathbf{X}} = \left[\bar{X}_1, \bar{X}_2, \dots, \bar{X}_n\right]^T. \tag{16}$$

The corresponding covariance matrix, containing the pair-wise values of the covariance function, is defined as

$$\mathbf{C_{XX}} = E\left[\left(\mathbf{X} - \bar{\mathbf{X}}\right)\left(\mathbf{X} - \bar{\mathbf{X}}\right)^T\right]. \tag{17}$$

This leads to the coefficient of correlation between two random variables X_1 and X_2 as follows

$$\rho_{12} = \frac{E\left[\left(X_1 - \bar{X}_1\right)\left(X_2 - \bar{X}_2\right)\right]}{\sigma_{X_1}\sigma_{X_2}}. \tag{18}$$

For a single random variable the maximum entropy distribution is obtained by considering only the moment constraints. For the multivariate distribution the correlations between each pair of random variables has to be taken into account as well. This would lead to an optimization problem with $4n$ optimization parameters with $4n$ constraints from the marginal moment conditions and $n(n+1)/2$ constraints from the correlation conditions, where n is number of random variables. This concept was recently applied in [9] to determine the joint density function. For a larger number of random variables the solution of this optimization problem is numerically very demanding.

In our study we avoid this high dimensional optimization problem by applying the Nataf model [10], [11] to construct multivariate distributions. In this model a vector of standard normally distributed random variables

$$\mathbf{Z} = [Z_1, Z_2, \dots, Z_n]^T \tag{19}$$

is obtained by the marginal transformation of the original random vector $\mathbf{X}$ as

$$Z_i = \Phi^{-1}[F_{X_i}(X_i)], \quad i = 1, \dots, n. \tag{20}$$

By assuming that $\mathbf{Z}$ is jointly normal distributed, the joint probability density function of $\mathbf{X}$ reads

$$f_{\mathbf{X}}(\mathbf{x}) = f_{X_1}(x_1)\dots f_{X_n}(x_n)\frac{\phi_n(\mathbf{z}, \mathbf{C_{ZZ}})}{\phi(z_1)\phi(z_2)\dots\phi(z_n)}, \tag{21}$$

where $z_i = \Phi^{-1}[F_{X_i}(x_i)]$, $\phi(.)$ is the standard normal probability density function and $\phi_n(\mathbf{z}, \mathbf{C_{ZZ}})$ is the n-dimensional standard normal density depending on the covariance matrix of $\mathbf{Z}$. The elements of this covariance matrix are the correlation coefficients of $\mathbf{Z}$

$$C_{ZZ}(Z_i, Z_j) = \tilde{\rho}_{ij}, \tag{22}$$

which are defined in terms of the correlation coefficients ρ_{ij} of the original random vector $\mathbf{X}$ as

$$\begin{aligned} \rho_{ij} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\frac{x_i - \bar{X}_i}{\sigma_{X_i}}\right)\left(\frac{x_j - \bar{X}_j}{\sigma_{X_j}}\right) f_{X_i}(x_i) f_{X_j}(x_j)\frac{\phi_2(z_i, z_j, \tilde{\rho}_{ij})}{\phi(z_i)\phi(z_j)}\, dx_i\, dx_j, \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\frac{x_i - \bar{X}_i}{\sigma_{X_i}}\right)\left(\frac{x_j - \bar{X}_j}{\sigma_{X_j}}\right)\phi_2(z_i, z_j, \tilde{\rho}_{ij})\, dz_i\, dz_j. \end{aligned} \tag{23}$$

By applying the presented Nataf model the multivariate distribution function is obtained by solving the optimization problem with four parameters for each random variable independently. The successful application of the model requires a positive definite covariance matrix $\mathbf{C_{ZZ}}$ and continuous and strictly increasing distribution functions $F_{X_i}(x_i)$. In our study Equation 23 is solved iteratively to obtain $\tilde{\rho}_{ij}$ for each pair of marginal distributions from the known correlation coefficient ρ_{ij}.

3 ESTIMATING PARAMETER UNCERTAINTIES BY THE BOOTSTRAP METHOD

Based on a given number of observations the uncertainties of random parameters shall be estimated. An infinite number of observations x_i of a random variable X is necessary to obtain the statistical properties as mean value and variance exactly. For a finite number of observations the estimators of these parameters give only approximate solutions. In our study we want to estimate these approximation errors of the statistical properties. For this purpose we utilize the bootstrapping approach which was introduced by [3]. In the bootstrapping method the estimating properties of an estimator are obtained by sampling from an approximate distribution which is generally the empirical distribution of the observed data. This method assumes independent and identically distributed observations and constructs a number of re-samples by random sampling with replacement from the observation dataset. Based on this resampling sets the properties of the estimated parameters including their distribution functions, confidence intervals etc. can be obtained. Generally the number of samples in a single resampling set is taken equal to the number of observations and the number of resampling sets has to be chosen very large.

The standard approach uses a fixed set of samples to extract the bootstrap sample sets. This analysis serves the statistical errors only cause by the small sample observations. In our study we extend this method for imprecise observations. Each observation is assumed to be an independent single random variable with known mean and standard deviation and a given distribution type which is taken here as normally distributed. For each bootstrap sample set first the observations are sampled from the given random variables and than the standard bootstrap extraction is applied.

4 NUMERICAL EXAMPLE: RELIABILITY ANALYSIS OF A SHALLOW FOUNDATION

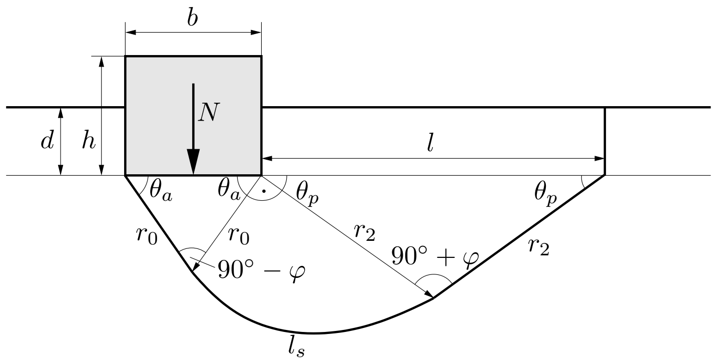
We apply the presented maximum entropy concept within this example for the modeling of the soil parameters and evaluate the influence of the parameter uncertainties on the failure probability. The limit state function of the bearing failure of a simple strip foundation with pure vertical loading can be derived based on [12] and [13] as

$$g = b \cdot (\gamma_s \cdot d \cdot N_{d0} + \gamma_s \cdot b \cdot N_{b0} + c \cdot N_{c0}) - N - bh\gamma_c \tag{24}$$

where b, d and h are geometrical properties as indicated in Figure 4 and γ_s and γ_c are the specific weights of the soil and the concrete, respectively. The bearing capacity factor are given as

$$N_{d0} = \tan^2\left(45^\circ + \frac{\varphi}{2}\right) \cdot e^{\pi \cdot \tan\varphi}, \quad N_{b0} = (N_{d0} - 1)\tan\varphi, \quad N_{c0} = \frac{N_{d0} - 1}{\tan\varphi}. \tag{25}$$

In our investigation we consider the friction angle φ and the cohesion c of the soil as random variables. All other quantities are taken as deterministic values. Based on the measurements



$$\vartheta_a = 45^\circ + \frac{\varphi}{2}, \qquad \vartheta_p = 45^\circ - \frac{\varphi}{2}, \quad r_0 = \frac{b}{2\cos\vartheta_a}$$

$$r_2 = r_0 \exp\left(\frac{\pi}{2}\tan\varphi\right), \qquad l = 2r_2\cos\vartheta_p, \quad l_s = \frac{r_2 - r_0}{\sin\varphi}$$

Figure 4: Bearing failure surface of a shallow foundation with pure vertical loading

published in [14] where several specimen of Frankfurter clay have been analyzed, the variation of the stochastic parameters are calculated. In the numerical analysis the spatial variability of the soil parameters has to be locally averaged in order to obtain single random variables for the soil parameters. According to [15] this averaging can be performed by an integration of the correlation function. If we assume an isotropic and exponential correlation function

$$\rho(l_H, \Delta x) = \exp\left(-2\frac{|\Delta x|}{l_H}\right) \tag{26}$$

with the correlation length l_H the following variance reduction function is obtained

$$\Gamma^2(l_H, l_A) = \frac{1}{2}\left(\frac{l_H}{l_A}\right)^2 \left[\frac{2l_A}{l_H} - 1 + \exp\left(-\frac{2l_A}{l_H}\right)\right] \tag{27}$$

where l_A specifies the averaging length which is in our case the length of failure surface given in Figure 4 as

$$l_F = r_0 + l_s + r_2. \tag{28}$$

The value of the reduction function is obtained using the mean value of the friction angle and the foundation geometry. The mean value of the friction angle from the samples in [14] is obtained as 21.2°. We assume the foundation geometry parameters as $b = 3$ m and $d = h = 1$ m which leads to a failure length of $l_F = 13.7$ m. Together with a correlation length of the soil parameters of $l_H = 3.0$ m according to [16] the value of the reduction function is obtained as $\Gamma = 0.44$.

Before applying the bootstrap method the original samples from [14] are modified by the reduction function as follows

$$\varphi_i' = \bar{\varphi} + \Gamma(\varphi_i - \bar{\varphi}), \quad c_i' = \bar{c} + \Gamma(c_i - \bar{c}). \tag{29}$$

In Figure 5 the histograms of the modified samples are shown together with the corresponding log-normal and maximum entropy density functions. For the friction angle both density functions are almost identical whereas for the cohesion we remark a significant difference. Additionally to the sample distributions the histograms from 10000 bootstrap samples from the

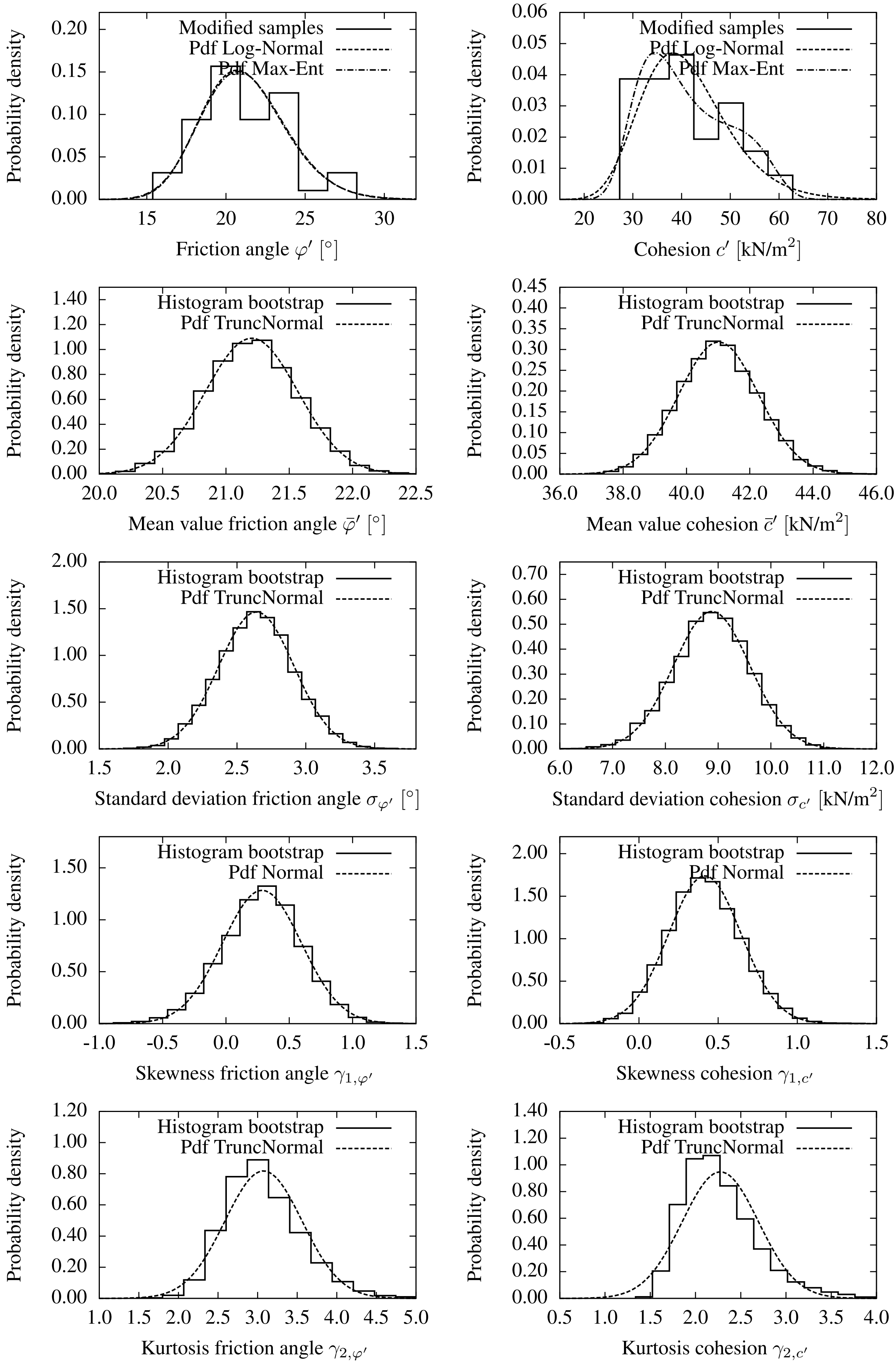


Figure 5: Maximum entropy distributions of the modified soil parameter samples and obtained distributions of the uncertain stochastic parameters from the bootstrap approach assuming exact measurements

fixed observation sets are shown for the mean value, the standard deviation, the skewness and the kurtosis. The figure indicates a very good agreement of these histograms with a normal distribution except for the kurtosis where an un-symmetric distribution seems to be more suitable. Since negative values are not possible for the mean value, the standard deviation and the kurtosis of the soil parameters, the normal distribution is truncated below zero. In Table 1 the calculated numerical values from the standard bootstrap samples are given. Additionally the results assuming imprecise measurements with the observation errors $\sigma_{\varphi'_i} = 2.0°$ and $\sigma_{c'_i} = 5.0$kN/m^2 are given. The table indicates that some of the stochastic parameters obtained from the classical bootstrap method are strongly correlated. Due to the observation noise these correlation values are reduced and the variation of the mean value and of the standard deviation are increased. The values of the skewness and kurtosis are slightly changed in the direction of a normal distribution while their variation remains almost unchanged.

	Friction angle φ' [°]		Cohesion c' [kN/m^2]	
Samples				
Mean value $\bar{X}$	21.19		41.04	
Standard deviation σ_X	2.68		9.01	
Skewness γ_1	0.32		0.43	
Kurtosis γ_2	3.15		2.26	
Bootstrap results assuming exact measurements				
Bootstrap estimates	Mean	Std.	Mean	Std.
$\bar{X}$	21.19	0.37	41.03	1.26
σ_X	2.64	0.27	8.88	0.73
γ_1	0.29	0.31	0.42	0.23
γ_2	3.09	0.51	2.28	0.43
Bootstrap correlations				
$\rho_{\bar{X}\sigma_X}$	0.21		0.37	
$\rho_{\bar{X}\gamma_1}$	0.01		-0.59	
$\rho_{\bar{X}\gamma_2}$	-0.13		-0.43	
$\rho_{\sigma_X\gamma_1}$	0.14		-0.09	
$\rho_{\sigma_X\gamma_2}$	-0.28		-0.36	
$\rho_{\gamma_1\gamma_2}$	0.48		0.78	
Bootstrap results assuming imprecise measurements				
Bootstrap estimates	Mean	Std.	Mean	Std.
$\bar{X}$	21.19	0.46	41.04	1.44
σ_X	3.30	0.33	10.18	0.91
γ_1	0.14	0.33	0.27	0.27
γ_2	2.96	0.61	2.55	0.47
Bootstrap correlations				
$\rho_{\bar{X}\sigma_X}$	0.09		0.23	
$\rho_{\bar{X}\gamma_1}$	0.01		-0.27	
$\rho_{\bar{X}\gamma_2}$	-0.2		-0.21	
$\rho_{\sigma_X\gamma_1}$	0.09		0.02	
$\rho_{\sigma_X\gamma_2}$	-0.04		-0.18	
$\rho_{\gamma_1\gamma_2}$	0.25		0.49	

Table 1: Soil parameter estimation: classical and bootstrap results from exact and imprecise measurements

Parameter distributions	N [kN]	β_{FORM}	P_F
Log-normal	1000	3.84	$6.1 \cdot 10^{-5}$
	1100	3.42	$3.1 \cdot 10^{-4}$
	1200	3.05	$1.1 \cdot 10^{-3}$
	1300	2.70	$3.4 \cdot 10^{-3}$
Maximum entropy	1000	4.04	$2.7 \cdot 10^{-5}$
	1100	3.55	$1.9 \cdot 10^{-4}$
	1200	3.12	$8.9 \cdot 10^{-4}$
	1300	2.74	$3.1 \cdot 10^{-3}$

Table 2: Results of classical reliability analysis with increasing loading

Based on the mean values of the bootstrap results from the exact measurements we perform a classical reliability analysis using the FORM approach. We apply log-normal and maximum entropy distribution types for the soil parameters φ' and c' and assume no correlation between both. Using the geometry values $b = 3$ m and $d = h = 1$ m and the specific weights $\gamma_s = 17.7$ kN/m^3 and $\gamma_c = 25.0$ kN/m^3 the reliability index and the corresponding failure probability are calculated for an increasing load N. The obtained results are given in Table 2. The results indicate that with decreasing load the deviation of the failure probability between both distribution types increases since the tails of the distributions become more significant. Thus for small values of the failure probability the choice of the distribution type has an enormous influence on the results of the reliability analysis.

In our final analysis we consider the variation in the stochastic parameters. Based on a first order Taylor series approximation the variation of the failure probability can be computed directly as

$$P_F(\mathfrak{p}) \approx P_F(\mathfrak{p}_0) + (\mathfrak{p} - \mathfrak{p}_0)^T \left. \frac{\partial P_F(\mathfrak{p})}{\partial \mathfrak{p}} \right|_{\mathfrak{p}_0}, \tag{30}$$

where $\mathfrak{p}$ is the vector of the uncertain stochastic parameters

$$\mathfrak{p} = [\bar{\varphi}', \sigma_{\varphi'}, \gamma_{1,\varphi'}, \gamma_{2,\varphi'}, \bar{c}', \sigma_{c'}, \gamma_{1,c'}, \gamma_{2,c'}] \,. \tag{31}$$

The required derivatives can be obtained for the FORM approach very efficiently as reported in [17]. Based on this Taylor series approximation the failure probability can be calculated quite accurate for each specific sample of the stochastic parameters close to the mean value vector $\mathfrak{p}_0$. This calculation is performed here for all 10000 bootstrap samples. The resulting histogram of the reliability index is shown in Figure 6 using log-normally distributed φ' and c' and in Figure 7 based on the maximum entropy distributions whereby for both cases exact and imprecise measurements are investigated.

Based on the assumption of almost normally distributed stochastic parameters $\mathfrak{p}$ the variance of the failure probability can be directly estimated from the covariance matrix of these parameters

$$\sigma^2_{P_F} \approx \left. \frac{\partial P_F(\mathfrak{p})}{\partial \mathfrak{p}} \right|^T_{\mathfrak{p}_0} \mathbf{C}_{\mathfrak{pp}} \left. \frac{\partial P_F(\mathfrak{p})}{\partial \mathfrak{p}} \right|_{\mathfrak{p}_0} . \tag{32}$$

In the Figures 6 and 7 the resulting normal density function of the reliability index is shown additionally. The figure indicates a very good agreement of the sample analysis and the Gaussian approximation. In Table 3 the resulting standard deviations of the reliability index for both methods including the 90% confidence intervals are given. The values clearly indicate, that

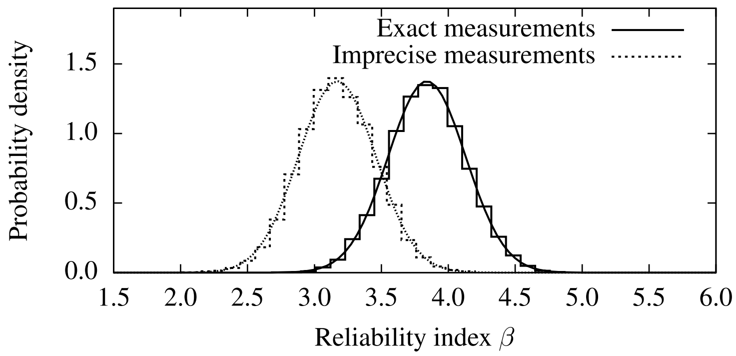


Figure 6: Variation of the reliability index using log-normally distributed soil parameters

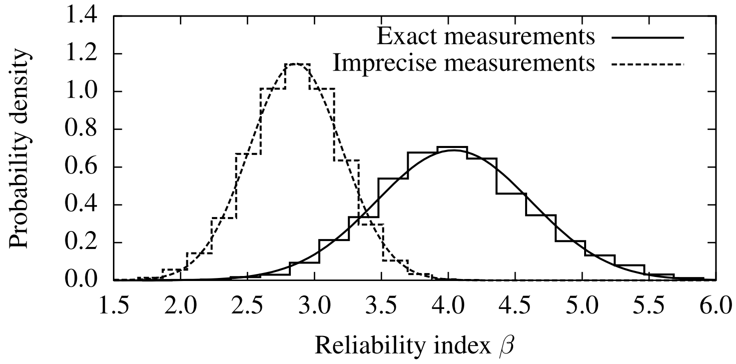


Figure 7: Variation of the reliability index based on maximum entropy distributions

the calculated variation based on log-normally distributed soil parameters is much less as with the maximum entropy distribution. This is caused by the fact that the obtained variation of the skewness and kurtosis from the bootstrap analysis has no influence on the log-normal distribution. Thus the uncertainty in the distribution function itself can be much better represented with the presented concept if the maximum entropy distribution is used. Additionally to the values from the coupled analysis with correlated stochastic parameters the variation of the reliability index is estimated assuming these parameters as uncorrelated. The obtained results, shown additionally in Table 3, indicate that especially for the maximum entropy distribution type this assumption leads to an over-estimation of the variation.

The consideration of imprecise measurements decreases the mean reliability index for both distributions types whereas for the log-normal distribution its variation remains almost unchanged but for the maximum entropy distributions the variation of the reliability index decreases.

5 CONCLUSIONS

In this paper we have presented an approach to model uncertainties in the distributions of random variables by maximum entropy formulations based on the first four stochastic moments. Since these moments can not be estimated exactly for small-sample observations we model them as uncertain parameters. Based on a given set of observations we estimated the uncertainties utilizing the bootstrap method. We observed an almost normal distribution of the mean value,

Analysis method	Mean value	Stand. dev.	90% confidence interval
Exact measurements			
Log-normal distributions			
Classical FORM	3.840	-	-
Bootstrap samples	3.841	0.288	3.362 - 4.319
Gaussian estimate correlated	3.840	0.290	3.363 - 4.317
Gaussian estimate uncorrelated	3.840	0.336	3.287 - 4.393
Maximum entropy distributions			
Classical FORM	4.040	-	-
Bootstrap samples	4.053	0.583	3.142 - 5.075
Gaussian estimate correlated	4.040	0.579	3.088 - 4.992
Gaussian estimate uncorrelated	4.040	0.966	2.451 - 5.629
Imprecise measurements			
Log-normal distributions			
Classical FORM	3.172	-	-
Bootstrap samples	3.172	0.285	2.713 - 3.642
Gaussian estimate correlated	3.172	0.286	
Gaussian estimate uncorrelated	3.172	0.310	
Maximum entropy distributions			
Classical FORM	2.855	-	-
Bootstrap samples	2.855	0.347	2.28 - 3.415
Gaussian estimate correlated	2.855	0.347	
Gaussian estimate uncorrelated	2.855	0.430	

Table 3: Variation of the reliability index β using the bootstrap samples and Gaussian estimation by assuming exact and imprecise measurements

the standard deviation and the skewness but a skewed distribution of the kurtosis.

We have estimated the variation of the computed failure probability for a bearing failure problem of a shallow foundation based on the obtained variation of the stochastic parameters. As an outcome we can summarize that the consideration of the uncertainties in the skewness and kurtosis, which can be represented with the maximum entropy formulation, leads to remarkable larger variations of the failure probability as obtained with a standard distribution type, where only the mean value and the standard deviation are uncertain.

An approximation of the parameter distributions by a Gaussian density enables an efficient estimation of the failure probability variation which agreed very well with the results of the sample analysis in the investigated example.

Finally we can conclude that the presented extension of the bootstrap method is a very suitable and simple method to consider imprecise measurements even if varying measurement errors and distributions are assumed.

6 ACKNOWLEDGMENT

This research has been supported by the German Research Council (DFG) through Research Training Group 1462, which is gratefully acknowledged by the author.

REFERENCES

[1] A. Der Kiureghian. Analysis of structural reliability under parameter uncertainties. *Probabilistic Engineering Mechanics*, 23:351–358, 2008.

[2] M. Rockinger and E. Jondeau. Entropy densities with an application to autoregressive conditional skewness and kurtosis. *Journal of Econometrics*, 106:119–142, 2002.

[3] B. Efron. Bootstrap methods: another look at the Jackknife. *The Annals of Statistics*, 7(1):1–26, 1979.

[4] C.E. Shannon. The mathematical theory of communication. *Bell Systems Technical Journal*, 27:379–423, 1948.

[5] P.C. Basu and A.B. Templeman. An efficient algorithm to generate maximum entropy distributions. *International Journal for Numerical Methods in Engineering*, 20:1039–1055, 1984.

[6] A. Zellner and R.A. Highfield. Calculation of maximum entropy distributions and approximation of marginal posterior distributions. *Journal of Econometrics*, 37:195–2009, 1988.

[7] N. van Erp and P. van Gelder. Introducing entropy distributions. In C.-A. Graubner et al., editors, *Proc. 6th Inter. Probabilistic Workshop, November 26-27, 2008, Darmstadt, Germany*. 2008.

[8] W.K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57:97–109, 1970.

[9] C. Soize. Construction of probability distributions in high dimension using the maximum entropy principle: Application to stochastic processes, random fields and random matrices. *International Journal for Numerical Methods in Engineering*, 76:1583–1611, 2008.

[10] A. Nataf. Détermination des distributions de probabilités dont les marges sont données. *Comptes Rendus de l'Academie des Sciences*, 225:42–43, 1962.

[11] P.-L. Liu and A. Der Kiureghian. Multivariate distribution models with prescribed marginals and covariances. *Probabilistic Engineering Mechanics*, 1:105–112, 1986.

[12] K. Terzaghi. *Theoretical Soil Mechanics*. Whiley, New York, 1943.

[13] G.G. Meyerhof. Some recent foundation research and its application to design. *The Structural Engineer*, 31(6):151–167, 1953.

[14] C. Moormann. *Trag- und Verformungsverhalten tiefer Baugruben in bindigen Böden unter besonderer Berücksichtigung der Baugrund-Tragwerk- und der Baugrund-Grundwasser-Interaktion*. PhD thesis, Technische Universität Darmstadt, Germany, 2002.

[15] E.H. VanMarcke. *Random Fields - Analysis and Synthesis*. MIT-Press, Cambridge, 1983.

[16] C. Cherubini. Data and considerations on the variability of geotechnical properties of soils. In *Proceedings of the International Conference on Safety and Reliability, ESREL 97*, 1997.

[17] T. Most and T. Knabe. Reliability analysis of the bearing failure problem considering uncertain stochastic parameters. *Computers and Geotechnics, submitted*, 2009.