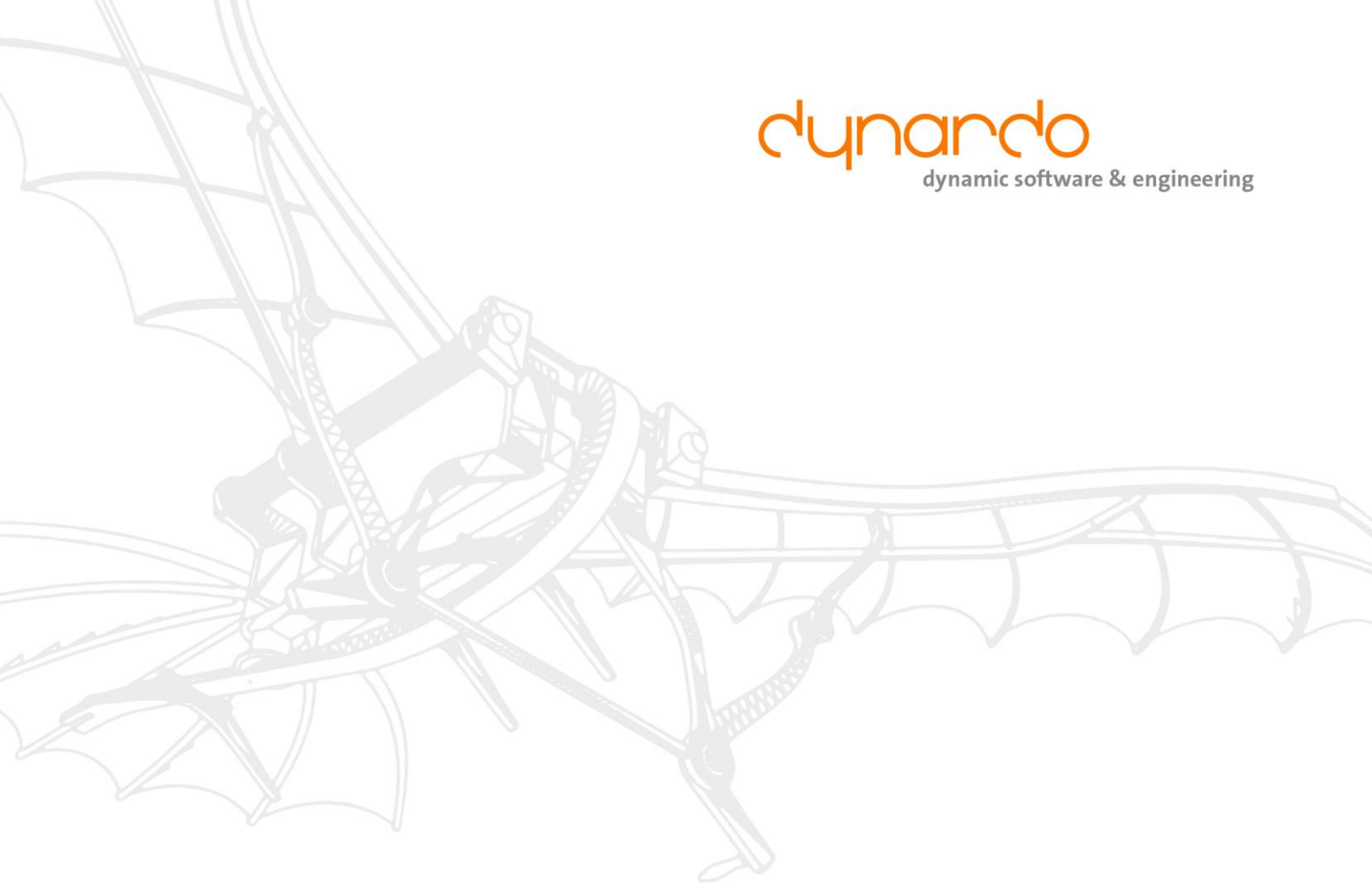
dynardo
dynamic software & engineering



Lectures

Identification of the parameters of complex constitutive models: Least squares minimization vs. Bayesian updating

Thomas Most

presented at the Reliability Conference in München 2010
Source: www.dynardo.de/en/library

Identification of the parameters of complex constitutive models: Least squares minimization vs. Bayesian updating

Thomas Most
Research Training Group 1462, Bauhaus-Universität Weimar, Germany

In this study the common least-squares minimization approach is compared to the Bayesian updating procedure. In the content of material parameter identification the posterior parameter density function is obtained from its prior and the likelihood function of the measurements. By using Markov Chain Monte Carlo methods, such as the Metropolis-Hastings algorithm (Hastings 1970), the global density function including local peaks can be computed. Thus this procedure enables an accurate evaluation of the global parameter quality. However, the computational effort is remarkable larger compared to the minimization approach. Thus several methodologies for an efficient approximation of the likelihood function are discussed in the present study.

1 INTRODUCTION

In all fields of computational mechanics numerical models are used to analyze engineering problems. In material modeling many types of constitutive formulations exist for a wide range of materials. Generally the parameters of these models are identified from experimental data. Beside manual identification by try and error or based on experiences fully automatic identification procedures using optimization strategies have become popular in the last years. Nevertheless, not only the parameter values itself are of interest for a later numerical analysis, but also their accuracy which is mainly influenced by the measurement errors and the type of experiments. For very complex constitutive models it is very difficult to judge in advance which experiments are required to enable a sufficient identifiability of all parameters. Therefor an estimation of the parameter uncertainties is very useful. Based on the determined optimal parameter set as the result of a least squares minimization, where the gap between experimental and numerical response data is minimized, the parameter covariances can be estimated based on the measurement errors and a linearization of the model (Ledesma et al. 1996). This approach was derived for a convex optimization problem. Applications for non-convex problems using complex material models in geomechanical applications can be found in (Hofmann et al. 2009), (Knabe et al. 2009).

Alternatively to the ordinary least squares formulation Kalman filter identification has been introduced (Cividini et al. 1983), (Bittani et al. 1984). This technique is a Bayesian estimator using prior information to identify the parameters. Similar to the covariance analysis in least squares minimization the model itself is linearized and the parameter distributions are taken as Gaussian which implies a convex optimization problem.

Another method uses Bayesian inference to estimate the parameters including their statistical properties (Beck and Arnold 1977). This approach can be applied for non-convex problems and arbitrary parameter and measurement error distributions and works without any linearization of the numerical model.

Recently inverse approximation schemes using neural networks have been proposed to identify the constitutive parameters (Novák and Lehký 2006),(Most et al. 2007). In (Unger and Könke 2010) this approach has been extended using Bayesian neural networks in order to estimate the parameter accuracy. In opposite to the other procedures there measurement errors are neglected and the parameter uncertainty arises only from the identification procedure. This results in a vanishing uncertainty if the number of training samples is increased dramatically. Therefor this procedure is not discussed further in this study.

2 LEAST SQUARES MINIMIZATION

2.1 Maximum Likelihood formulation

A deterministic models which relates a set of responses $\mathbf{x} = \mathbf{x}(\mathbf{p})$ with the model parameters $\mathbf{p}$ is assumed to be given. In the parameter identification procedure the measurements $\mathbf{x}^* = [x_1^* \dots x_m^*]^T$ are obtained by experiments.

The likelihood of the parameters is proportional to the conditional probability of measurements $\mathbf{x}^*$ from a given parameter set $\mathbf{p}$ (Ledesma et al. 1996):

$$L = k \cdot f(\mathbf{x}^*|\mathbf{p}). \tag{1}$$

If the chosen model is correct the gap between numerical responses and measurements $(\mathbf{x}^* - \mathbf{x})$ is caused only by the measurement errors. Thus $P(\mathbf{x}^*|\mathbf{p})$ is equivalent to the probability of reproducing the measurement errors. Assuming a multivariate normal distribution we obtain

$$P(\mathbf{x}^* - \mathbf{x}) = \frac{1}{\sqrt{(2\pi)^m |\mathbf{C_{xx}}|}} \exp\left[-\frac{1}{2}(\mathbf{x}^* - \mathbf{x})^T \mathbf{C_{xx}}^{-1} (\mathbf{x}^* - \mathbf{x})\right]. \tag{2}$$

Maximizing the likelihood L is equivalent to minimize $S = -2 \ln L$. This yields to the well known least squares objective function:

$$J = (\mathbf{x}^* - \mathbf{x})^T \mathbf{C_{xx}}^{-1} (\mathbf{x}^* - \mathbf{x}), \tag{3}$$

where $\mathbf{C_{xx}}$ is the covariance matrix of the measurements. In (Ledesma et al. 1996) it is shown, that the optimal weights used in the objective function are

$$\mathbf{W} = \mathbf{C_{xx}}^{-1}. \tag{4}$$

If the objective function is linearized, the following update scheme is obtained

$$\Delta\mathbf{p} = \left(\mathbf{A}^T \mathbf{C_{xx}}^{-1} \mathbf{A}\right)^{-1} \mathbf{A}^T \mathbf{C_{xx}}^{-1} \Delta\mathbf{x}, \tag{5}$$

where $\mathbf{A}$ is the sensitivity matrix of the parameters with respect to the model responses

$$\mathbf{A} = \frac{\partial \mathbf{x}}{\partial \mathbf{p}}. \tag{6}$$

The update scheme can be used to obtain the optimal parameter set with the minimum objective function value. This approach is a gradient based method which leads to the global minimum without fail only for a convex optimization problem. If this method is applied for a non-convex problem, it may stuck in a local minimum. This is the case for more complex constitutive models. For this reason often global optimization schemes are used to identify the optimal parameter set. One of these methods is presented in the next section.

If the optimal parameter set is finally identified, the covariances of the parameters $\mathbf{C_{pp}}$ can be estimated by using the linearized relation from Eq.(5) which leads to the so-called Markov estimator

$$\mathbf{C_{pp}} = \left(\mathbf{A}_{opt}^T \mathbf{C_{xx}}^{-1} \mathbf{A}_{opt}\right)^{-1}. \tag{7}$$

In (Hofmann et al. 2009) and (Knabe et al. 2009) this estimator is applied also for non-convex optimization problem, which means that around the global optimum a local convex problem is assumed and the model is linearized at this optimum. Both, parameter and measurement distributions are assumed to be Gaussian within this approach.

2.2 Particle swarm optimization

In our study a population based global optimization algorithm, the particle swarm optimization (PSO), is applied which enables detecting the global optimum even if several local minima exist. Each population consists of a given number of particles where each particle position is equivalent to a set of parameters $\mathbf{p}_i^k$ and is updated in each iteration step k by the simple scheme (Kennedy and Eberhart 1995)

$$\begin{aligned} \mathbf{p}_i^{k+1} &= \mathbf{p}_i^k + \mathbf{v}_i^{k+1} \\ \mathbf{v}_i^{k+1} &= \omega \mathbf{v}_i^k + c_1 \mathbf{r}_1 (\mathbf{P}_i^k - \mathbf{p}_i^k) + c_2 \mathbf{r}_2 (\mathbf{P}_g^k - \mathbf{p}_i^k) \end{aligned} \tag{8}$$

where c_1 and c_2 are constants and $\mathbf{r}_1$ and $\mathbf{r}_2$ are vectors of uniformly distributed random variables between zero and one. $\mathbf{P}_i^k$ is the best position of a single particle and $\mathbf{P}_g^k$ is the global best position. We extend the original concept with passive congregation (He et al. 2004)

$$\mathbf{v}_i^{k+1} = \omega \mathbf{v}_i^k + c_1 \mathbf{r}_1 (\mathbf{P}_i^k - \mathbf{p}_i^k) + c_2 \mathbf{r}_2 (\mathbf{P}_g^k - \mathbf{p}_i^k) + c_3 \mathbf{r}_3 (\mathbf{R}_i^k - \mathbf{p}_i^k) \tag{9}$$

where an additional term is added to the original scheme to decrease the risk of running into a local minimum. This term $\mathbf{R}_i^k$ is defined as the best position of a randomly chosen particle. The required constants c_1, c_2 and c_3 are given in (He et al. 2004). If new particle positions are outside the parameter boundaries they are corrected by a harmony search scheme according to (Li et al. 2007).

3 BAYESIAN ESTIMATORS

3.1 Kalman filter

Kalman filter technology was firstly applied for parameter identification in (Cividini et al. 1983), (Bittani et al. 1984). More recent applications can be found in (Bolzon et al. 2002), (Fedele et al. 2006), (Furukawa and Pan 2009). In this approach the measurement noise is assumed to be a time-dependent white Gaussian random process characterized by zero mean, zero cross-correlation, and a time-dependent covariance matrix

$$\begin{aligned} E(\boldsymbol{\epsilon}^{(t)}) &= \mathbf{0}, \\ E(\boldsymbol{\epsilon}^{(t)} \boldsymbol{\epsilon}^{(s)^T}) &= \mathbf{0}, s \neq t \\ E(\boldsymbol{\epsilon}^{(t)} \boldsymbol{\epsilon}^{(t)^T}) &= \mathbf{C}_{\mathbf{xx}}^{(t)}. \end{aligned} \tag{10}$$

Using a linearization of the model the following update scheme can be derived

$$\begin{aligned} \hat{\mathbf{p}}^{(t)} &= \hat{\mathbf{p}}^{(t-1)} + \mathbf{K}^{(t)} (\mathbf{x}^* - \mathbf{x}(\hat{\mathbf{p}}^{(t-1)})) \\ \hat{\mathbf{C}}_{\mathbf{pp}}^{(t)} &= \hat{\mathbf{C}}_{\mathbf{pp}}^{(t-1)} - \mathbf{K}^{(t)} \mathbf{A}^{(t)} \hat{\mathbf{C}}_{\mathbf{pp}}^{(t-1)} \end{aligned} \tag{11}$$

where $\mathbf{K}$ is the Kalman gain matrix

$$\mathbf{K}^{(t)} = \hat{\mathbf{C}}_{\mathbf{pp}}^{(t-1)} \mathbf{A}^{(t)^T} \left[\mathbf{A}^{(t)} \hat{\mathbf{C}}_{\mathbf{pp}}^{(t-1)} \mathbf{A}^{(t)^T} + \mathbf{C}_{\mathbf{xx}}^{(t)} \right]^{-1} \tag{12}$$

and $\mathbf{A}$ is again the sensitivity matrix. As initial setting generally the prior information is taken

$$\hat{\mathbf{p}}^{(0)} = \mathbf{p}_0, \quad \hat{\mathbf{C}}_{\mathbf{pp}}^{(0)} = \mathbf{C}_{\mathbf{pp}_0}. \tag{13}$$

Due to the linearization of the model and the assumed Gaussian measurement noise, the parameter distributions are implicitly assumed to be Gaussian. In a static approach the time-step t is equivalent to the iteration step and the measurement covariance matrix remains constant.

(Cividini et al. 1983) has shown, that for vanishing prior information, where $(\mathbf{C}_{\mathbf{pp}_0})^{-1}$ tends to zero, the estimated parameter covariance is equivalent to the Markov estimator.

3.2 Bayesian updating

Bayesian updating is based on the well-known Bayes' theorem which shows the relation between one conditional probability and its inverse

$$P(A|B) = \frac{P(B|A)\, P(A)}{P(B)}. \tag{14}$$

Assuming a prior multivariate distribution of the parameters $\pi_{prior}(\mathbf{p})$ and a likelihood-function $\pi(\mathbf{x}^*|\mathbf{p})$ for the distribution of the measurements the conditional probability of the parameters can be obtained as (Beck and Arnold 1977)

$$\pi_{posterior}(\mathbf{p}) = \pi(\mathbf{p}|\mathbf{x}^*) = \frac{\pi_{prior}(\mathbf{p}) \cdot \pi(\mathbf{x}^*|\mathbf{p})}{\pi(\mathbf{x}^*)}. \tag{15}$$

Since the normalization constant $\pi(\mathbf{x}^*)$ is difficult to determine due to the only implicitly given distribution function, realizations of $\pi_{posterior}(\mathbf{p})$ are obtained by Markov-Chain Monte Carlo Simulation (Hastings 1970). For jointly Gaussian measurement errors the likelihood function is equivalent to Eq.(2).

3.3 Metropolis-Hastings algorithm

The Metropolis-Hastings algorithm (Hastings 1970) enables the sampling of an implicitly given not necessarily scaled density function. Therefor it is very important in interferential problems. This method is a Monte Carlo Simulation using a first order Markov chain, where the next sample depends only on the current state.

The algorithm generates a sequence of samples $\{\mathbf{p}^{(1)}, \mathbf{p}^{(2)}, \ldots, \mathbf{p}^{(n)}\}$ which statistically converge to the given distribution, which is in our case the Bayesian posterior multivariate parameter distribution. The algorithm works as follows:

1. Sample candidate $\mathbf{p}^*$ from jumping distribution $q(\mathbf{p}^*, \mathbf{p}^{(t-1)})$
2. Calculate
$$\alpha = \min\left[1, \frac{\pi(\mathbf{p}^*|\mathbf{x}^*)}{\pi(\mathbf{p}^{(t-1)}|\mathbf{x}^*)} \frac{q(\mathbf{p}^{(t-1)}, \mathbf{p}^*)}{q(\mathbf{p}^*, \mathbf{p}^{(t-1)})}\right] \tag{16}$$
3. Sample uniformly distributed $U \in (0.0; 1.0)$
4. If $U \leq \alpha$ accept $\mathbf{p}^{(t)} = \mathbf{p}^*$, otherwise $\mathbf{p}^{(t)} = \mathbf{p}^{(t-1)}$
5. Return to step 1

The jumping distribution is taken generally to be a Gaussian or uniform distribution. If it is symmetric $q(\mathbf{p}^{(t-1)}, \mathbf{p}^*)/q(\mathbf{p}^*, \mathbf{p}^{(t-1)})$ is equal to one. The initial parameter set is chosen randomly in the given parameter ranges. In this study the prior parameter distribution is taken as a uniform distribution in these parameter ranges. Thus the Metropolis-Hastings algorithm will only accept samples inside this boundaries. At the beginning the generated sequence is biased depending on the starting point and the jumping distribution. But after a certain burn-in phase it turns into an ergodic process. Thus the statistical evaluation of the parameter distributions is done be neglecting a given number of the first samples.

The variances of the jumping distribution strongly influences acceptance rate (AR) and convergence speed. If the distribution variances are taken too small, the convergence is very low and almost all samples will be accepted. An optimal statistical evaluation is enabled if the acceptance rate is between 10 and 30 % as shown in (Unger 2009).

For non-convex problems the variances of the jumping distribution have to be chosen in that way, that the peaks of the likelihood function are sufficiently covered. This is shown in Figure 1 and Table 1 depending on the measurement errors. If the measurement errors are small and the variances of the jumping distribution are to small, only one local optimum is covered by the sampling procedure. But if the variances are chosen large enough the acceptance rate is very small and a large number of samples is necessary to obtain statistically accurate results.

		Bayes' updating			
σ_ϵ	σ_q	AR	No. samples	$\tilde{\sigma}_{X_1}$	$\tilde{\sigma}_{X_2}$
0.05	1.0	0.02	100000	0.81	0.80
0.10	1.5	0.05	44000	0.77	0.76
0.20	1.5	0.15	13000	0.79	0.81
0.05	0.5	0.04	44000	0.13	0.12
0.10	0.5	0.20	10000	0.78	0.78
0.20	0.5	0.50	4000	0.79	0.79

Table 1: Estimated parameter variation for a non-convex problem with different measurement errors and variances of the jumping distribution

3.4 Scaling of measurement errors

In the previous section it was mentioned, that for small measurement errors a large number of samples may be required if the optimization problem is non-convex. For this reason two possibilities to reduce this numerical effort are discussed.

The first idea is to scale the covariances of the measurement errors by a constant factor a

$$\mathbf{C}_{\mathbf{xx}}^{scaled} = a \cdot \mathbf{C}_{\mathbf{xx}}^{orig}. \tag{17}$$

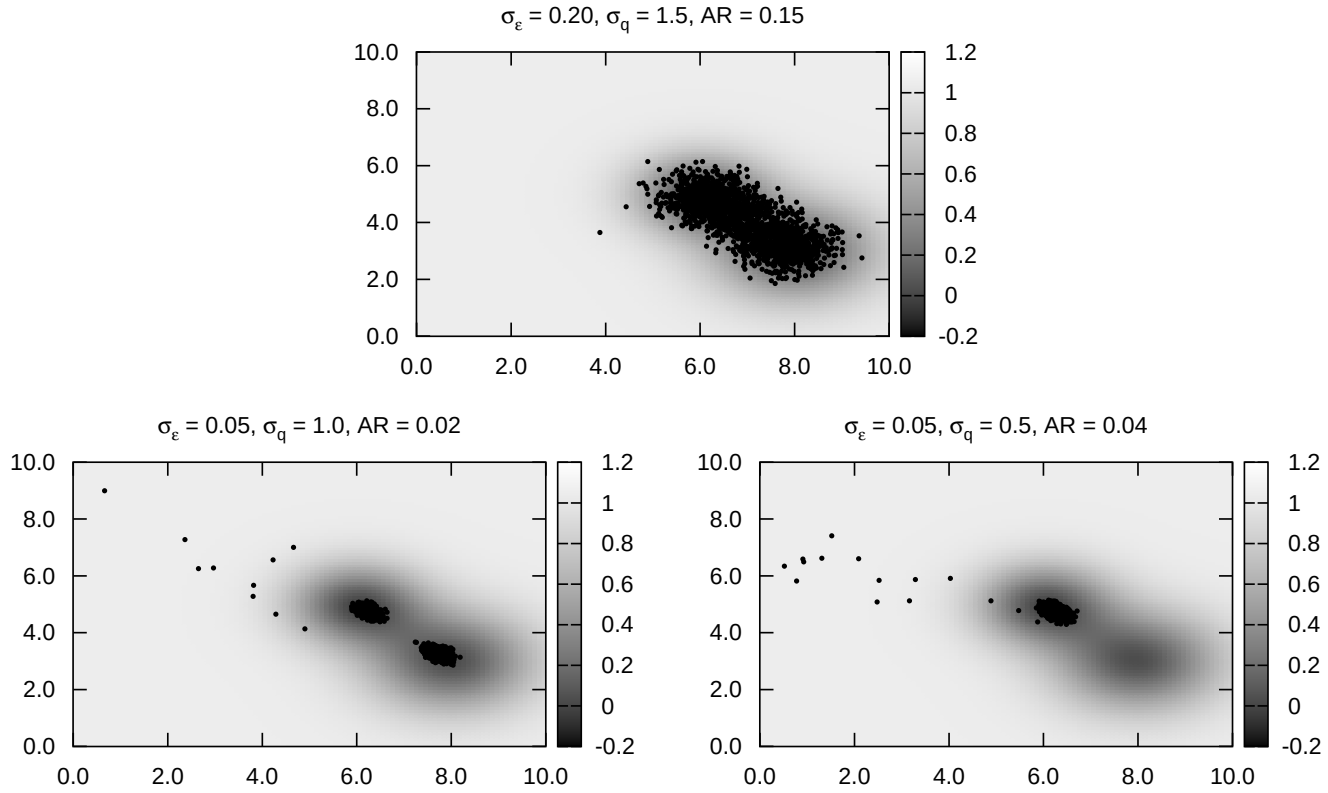


Figure 1: Samples from Metropolis-Hastings algorithm for a non-convex problem with different measurement errors

Based on the scaled likelihood function a sequence of samples is computed with the Metropolis-Hastings algorithm. The estimates of the parameter mean values and covariances of the original distribution can be computed as follows

$$\bar{p}_k = \frac{1}{n}\sum_{i=1}^{n} p_k^{(i)} \frac{f_{orig}(\mathbf{p}^{(i)})}{f_{scaled}(\mathbf{p}^{(i)})}$$
$$C_{kl} = \frac{1}{n-1}\sum_{i=1}^{n} (p_k^{(i)} - \bar{p}_k)(p_l^{(i)} - \bar{p}_l) \frac{f_{orig}(\mathbf{p}^{(i)})}{f_{scaled}(\mathbf{p}^{(i)})} \tag{18}$$

where the joint densities functions f_{orig} and f_{scaled} have to be known. Both are equivalent to the corresponding likelihood function scaled by the normalization constant. Since this normalization constant is not known it has to be estimated from the samples. In our opinion this is only possible by assuming a certain distribution type. This would suspend the advances of the Bayesian updating approach and would limit the method again to convex optimization problems. Therefor this scaling procedure is not investigated further in this study.

3.5 Approximation of the likelihood function

Another approach to reduce the numerical effort is to approximate the likelihood function by a suitable meta-model and using this approximation function instead of real model solutions. In (Orlande et al. 2008) this was shown by using radial basis function approximation. Dependent on the variance of the measurement noise the likelihood function contains sharp peaks around the local optima as shown in Figure 2. This requires an increased number of support points in the meta-model approach. For this reason in this study only the exponential term is approximated, which is equivalent to the objective function in Eq.(3). Then the measurement noise variance is only a scaling factor with no influence on the approximation quality. As meta-model approach interpolating Moving Least Squares (Most and Bucher 2008) is applied, since it can handle clustered support points and high gradients in the approximation function.

In Figure 3 and Table 2 the results with original and approximated likelihood function are compared for the non-convex problem. The results indicate a very good agreement in the parameter estimates. The numerical

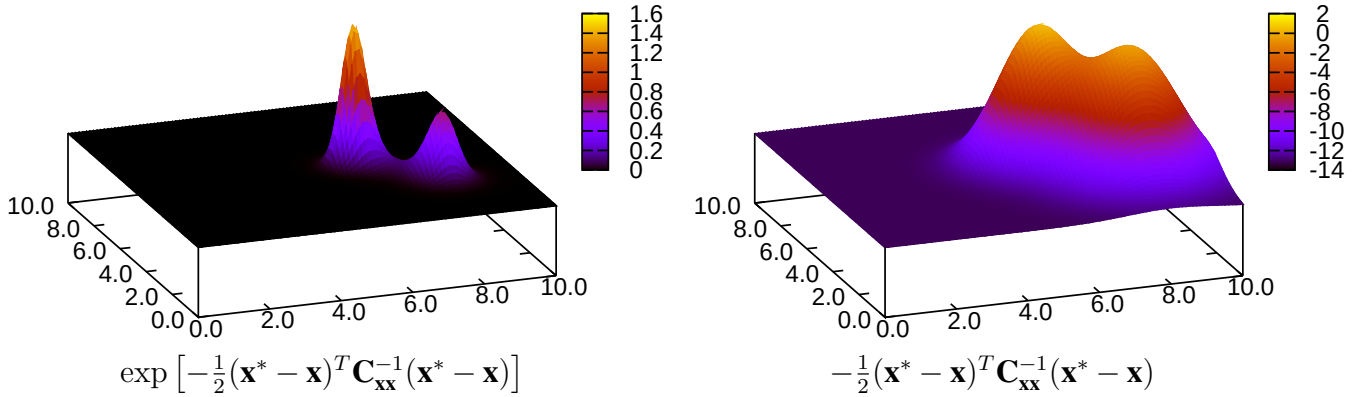


Figure 2: Likelihood function with sharp peaks and corresponding exponential term for small measurement errors

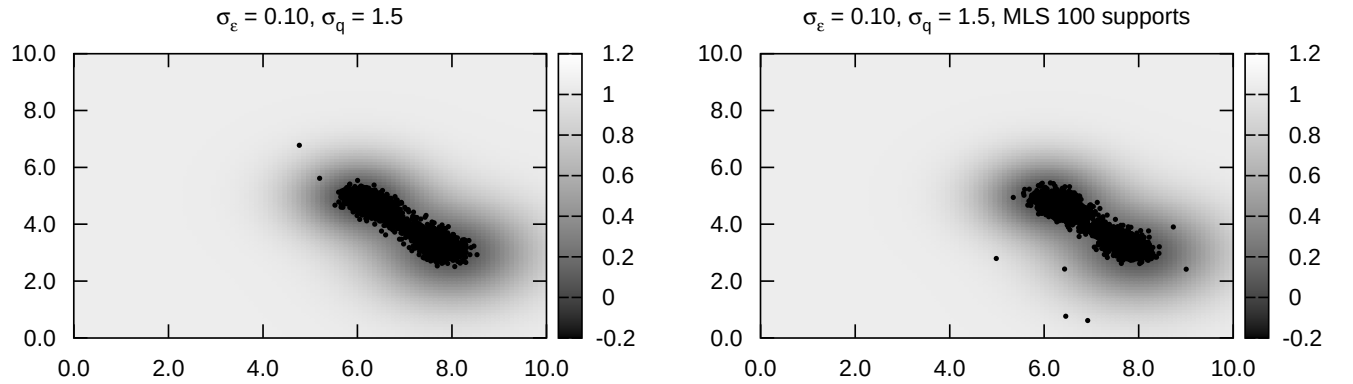


Figure 3: Computed parameter samples by using the original and approximated likelihood function

effort is reduced dramatically, since only 100 evaluations of the real model are required instead of 44000 by using the original likelihood function.

MLS approximation $\sigma_\epsilon = 0.10$, $\sigma_q = 1.5$				
	$\tilde{\mu}_{X_1}$	$\tilde{\mu}_{X_2}$	$\tilde{\sigma}_{X_1}$	$\tilde{\sigma}_{X_2}$
Original likelihood	7.22	3.83	0.77	0.76
MLS 100 supports	7.23	3.98	0.71	0.70
MLS 200 supports	6.91	4.08	0.78	0.73
MLS 500 supports	7.05	3.95	0.73	0.73

Table 2: Estimated parameters by using the original and approximated likelihood function

4 NUMERICAL EXAMPLES

4.1 1D elasto-plastic model without hardening

In this analytical example an one-dimensional elasto-plastic model with two parameters, the Young's modulus E and the yield stress σ_Y is investigated. As reference values $E_{ref} = 2.1 \cdot 10^{11}\text{N/m}^2$ and $\sigma_{Y,ref} = 5.0 \cdot 10^{8}\text{N/m}^2$ are chosen. By applying constant Gaussian measurement noise σ_ϵ for all observations 1000 samples are generated and for each sample the optimal parameter set is obtained by a gradient-based optimization method. The resulting variations of the identified parameters are shown in Table 3.

Additionally the Markov estimator is used at the parameter reference values and the Kalman filter approach is applied by using a large prior covariance $\mathbf{C}_{\mathbf{pp}}^0 = diag[100, 100] \cdot 10^8\text{N/m}^2$ in order to reduce the effect of the prior information. The results of both methods coincide excellent with the sample values due to the convexity of the problem. In Figure 5 the resulting parameter samples are shown.

Finally the Bayesian updating procedure is used whereby the prior distributions are chosen to be uniform between the optimization bounds $E \in (1.0; 3.0)$, $\sigma_Y \in (2.0; 10.0)$. The resulting estimates are given in Table 3. In Figure 5 the samples from the Metropolis-Hastings algorithm are shown, where the burn-in phase can be seen.

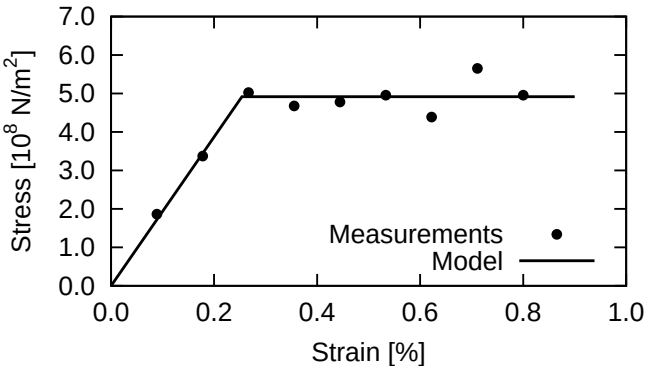


Figure 4: Stress-strain dependence of the 1D elasto-plastic model

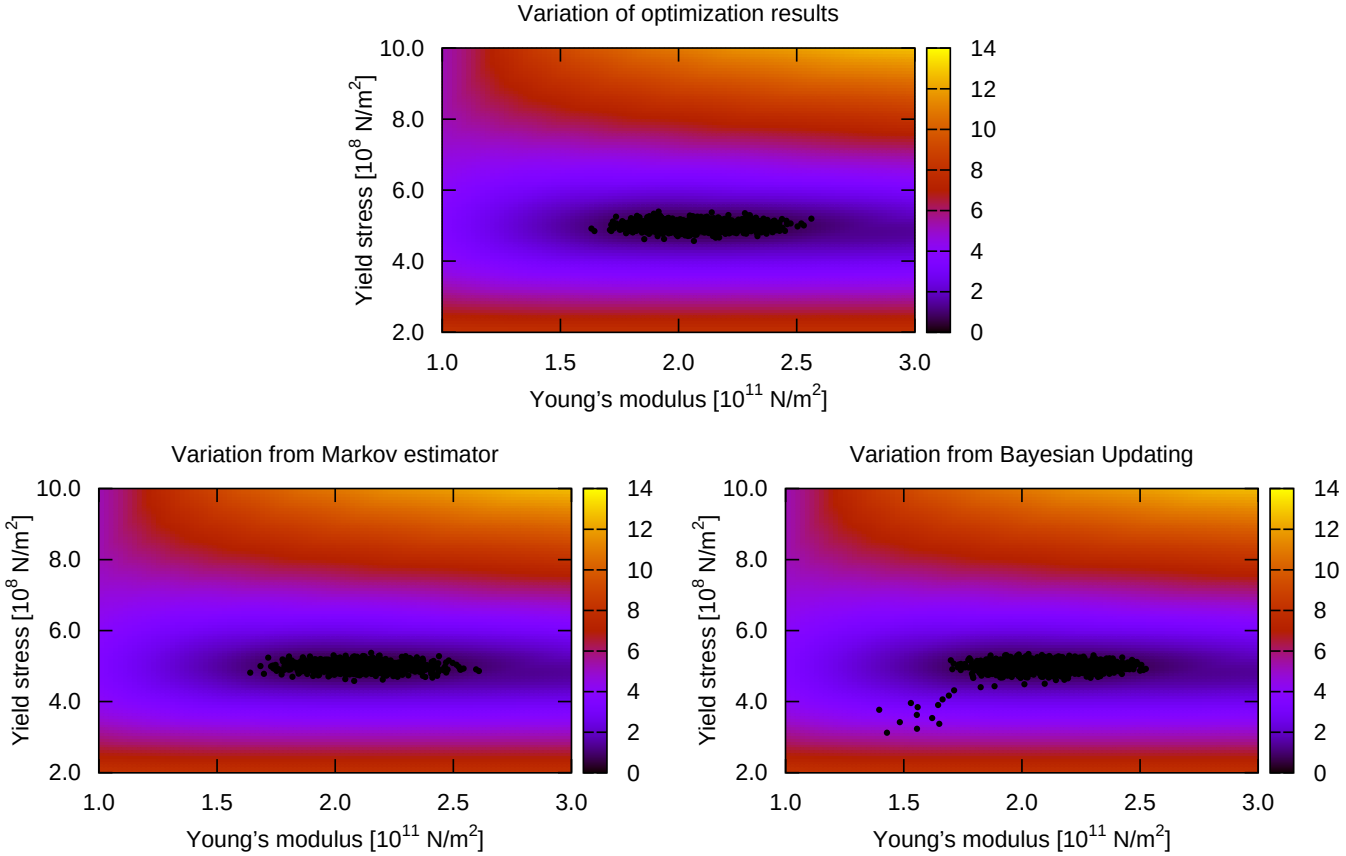


Figure 5: Deterministic objective function of the 1D elasto-plastic model including the parameter samples from the different approaches for $\sigma_\epsilon = 0.2 \cdot 10^8 \mathrm{N/m}^2$

			$\sigma_\epsilon[10^8\text{N/m}^2]$		
			0.05	0.10	0.20
Sample variation	σ_E	$[10^{11}\text{N/m}^2]$	0.025	0.052	0.157
	σ_{σ_Y}	$[10^8\text{N/m}^2]$	0.019	0.036	0.117
Markov estimator	$\tilde{\sigma}_E$	$[10^{11}\text{N/m}^2]$	0.025	0.050	0.151
	$\tilde{\sigma}_{\sigma_Y}$	$[10^8\text{N/m}^2]$	0.019	0.038	0.113
Kalman filter	$\tilde{\sigma}_E$	$[10^{11}\text{N/m}^2]$	0.025	0.050	0.151
	$\tilde{\sigma}_{\sigma_Y}$	$[10^8\text{N/m}^2]$	0.019	0.038	0.113
Bayes' updating	$\tilde{\sigma}_E$	$[10^{11}\text{N/m}^2]$	0.025	0.050	0.145
	$\tilde{\sigma}_{\sigma_Y}$	$[10^8\text{N/m}^2]$	0.019	0.038	0.115

Table 3: Estimated parameter variations for the 1D elasto-plastic model

The first 1000 samples are not considered in the statistical evaluation in order to get unbiased estimates. The results agree very well with the other approaches. Only for a larger measurement noise minor deviation can been observed which might be caused by the linearization in the Markov estimator and Kalman filter approaches.

4.2 Bilinear interface model for concrete cracking

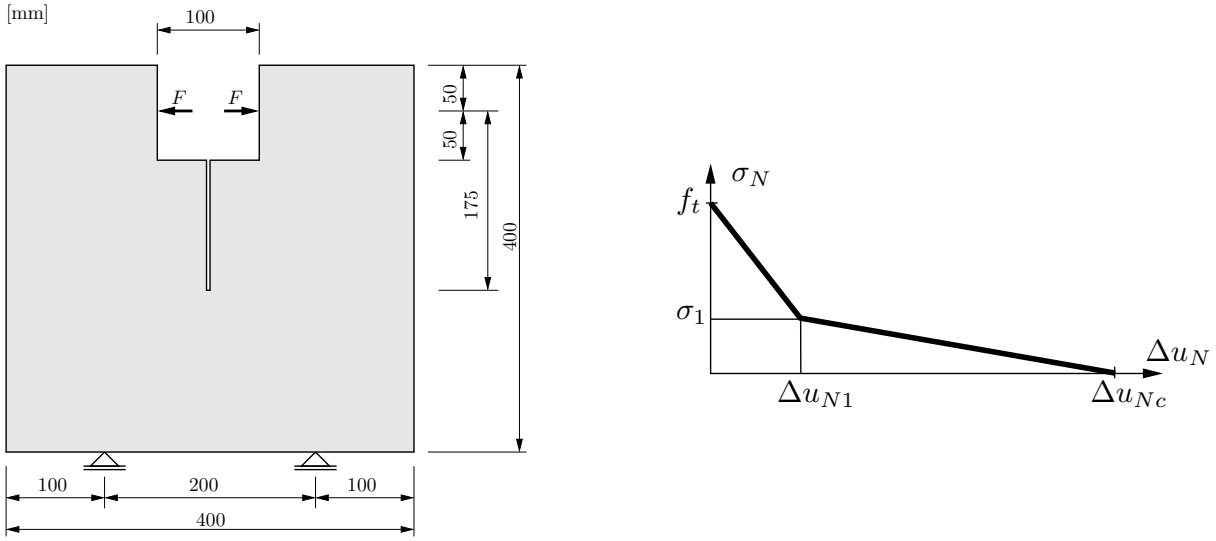


Figure 6: Wedge splitting test with system geometry and bilinear interface model

In this example the parameters of a cohesive crack model are identified. The investigations are based on the experimental program of (Trunk 1999). In Figure 6 the concrete specimen is shown with the corresponding dimensions. The numerical analysis is carried out using a finite element model with a predefined crack discretized by finite interface elements. Further details about the numerical suimulation can be found in (Most 2005). For the interface elements a bilinear softening law is used shown in principle in Figure 6. The remaining base material is assumed to be linear elastic. The following parameters have been indentified within the experimental program

- Young's modulus $E = 2.83 \cdot 10^{10}$ N/m^2
- Tensile strength $f_t = 2.27 \cdot 10^6$ N/m^2
- Specific fracture energy $G_f = 285$ Nm/m^2
- Softening shape parameters $\alpha_\sigma = \sigma_1/f_t = 0.163$, $\alpha_u = \Delta u_{N1}/\Delta u_{Nc} = 0.242$

The corresponding measurements and the numerical load displacement curve is shown in Figure 7

In a first step the measurements are taken as deterministic and 100 optimization runs with random start points using a gradient based and the particle swarm approach are carried out. The results are given in Table 4. Due to the existence of several local minima the gradient based methods does not converge to the global optimum in almost the half of the runs, whereby the PSO method is successfull in almost all cases . In Table 4

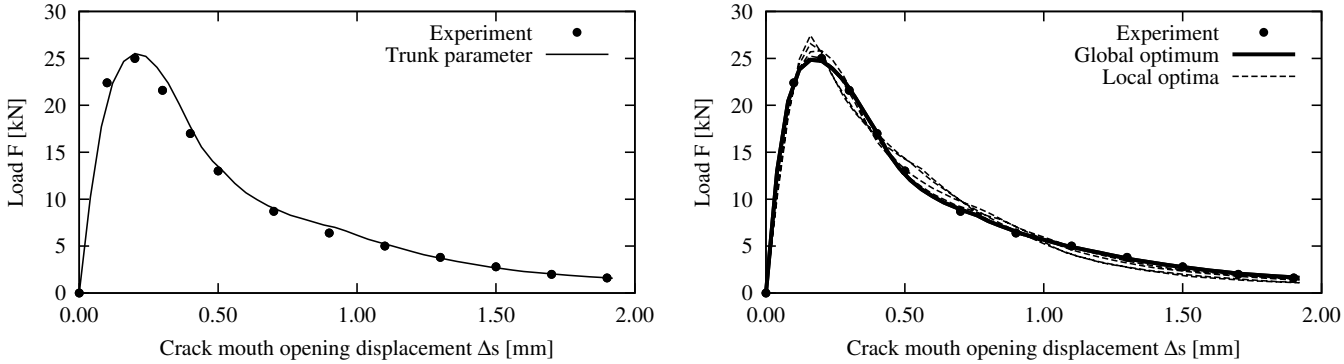


Figure 7: Load displacement curves for the wedge splitting test using the identified parameters

				Global		Total	
			Best	Mean	Std.	Mean	Std.
Gradient (success rate 51%)	E	$[10^{10}$ N/m$^2]$	3.85	3.89	0.11	3.72	0.39
	f_t	$[10^{6}$ N/m$^2]$	2.04	2.03	0.02	2.18	0.38
	G_f	$[10^{2}$ Nm/m$^2]$	2.84	2.84	0.01	2.81	0.07
	α_σ	$[10^{-1}]$	1.57	1.58	0.05	1.86	0.51
	α_u	$[10^{-1}]$	2.54	2.55	0.07	2.46	0.41
PSO (success rate 89%)	E	$[10^{10}$ N/m$^2]$	3.85	3.82	0.05	3.74	0.28
	f_t	$[10^{6}$ N/m$^2]$	2.04	2.05	0.01	2.13	0.31
	G_f	$[10^{2}$ Nm/m$^2]$	2.84	2.84	0.00	2.82	0.05
	α_σ	$[10^{-1}]$	1.57	1.57	0.02	1.70	0.42
	α_u	$[10^{-1}]$	2.54	2.53	0.03	2.44	0.32

Table 4: Optimization results for the wedge splitting test

the statistics of the runs ending close to the global optimum and of all runs are compared. The results indicate a very good accuracy of the PSO method which is even better than the gradient based approach.

In the next step the parameter uncertainties are estimated with the different approaches by assuming a constant measurement noise of $\sigma_\epsilon = 0.5 \cdot 10^6$N/m^2. In Table 5 and Figure 8 the results of these analyses are shown. They indicate are very good agreement of the Bayes' estimates with the variation from the sample analysis. The Markov estimator gives a good approximation of the variation, but some parameter dependencies are estimated not correct.

	Samples		Markov		Bayes	
	Mean	Std.	Mean	Std.	Mean	Std.
E	3.78	0.30	3.85	0.24	3.79	0.24
f_t	2.07	0.07	2.04	0.06	2.07	0.05
G_f	2.84	0.08	2.84	0.07	2.83	0.08
α_σ	1.62	0.32	1.57	0.18	1.62	0.26
α_u	2.51	0.26	2.54	0.17	2.52	0.24

$$Corr_{\mathbf{pp}}^{Samples} = \begin{bmatrix} 1.00 & -0.65 & 0.29 & -0.26 & 0.16 \\ & 1.00 & -0.13 & 0.40 & -0.34 \\ & & 1.00 & -0.31 & -0.65 \\ & & & 1.00 & 0.41 \\ & & & & 1.00 \end{bmatrix}$$

$$Corr_{\mathbf{pp}}^{Markov} = \begin{bmatrix} 1.00 & -0.59 & 0.29 & -0.83 & -0.40 \\ & 1.00 & -0.05 & 0.57 & -0.09 \\ & & 1.00 & 0.02 & -0.79 \\ & & & 1.00 & 0.00 \\ & & & & 1.00 \end{bmatrix}$$

$$Corr_{\mathbf{pp}}^{Bayes} = \begin{bmatrix} 1.00 & -0.56 & 0.17 & -0.09 & 0.12 \\ & 1.00 & -0.02 & 0.21 & -0.24 \\ & & 1.00 & -0.21 & -0.69 \\ & & & 1.00 & 0.49 \\ & & & & 1.00 \end{bmatrix}$$

Table 5: Parameter estimates and correlations for the wedge splitting test

5 CONCLUSIONS

The approximation of parameter uncertainties and dependencies by Bayes' updating is more accurate then using the Markov estimator. However, the Markov estimator is suitable as a rough estimate even for non-convex

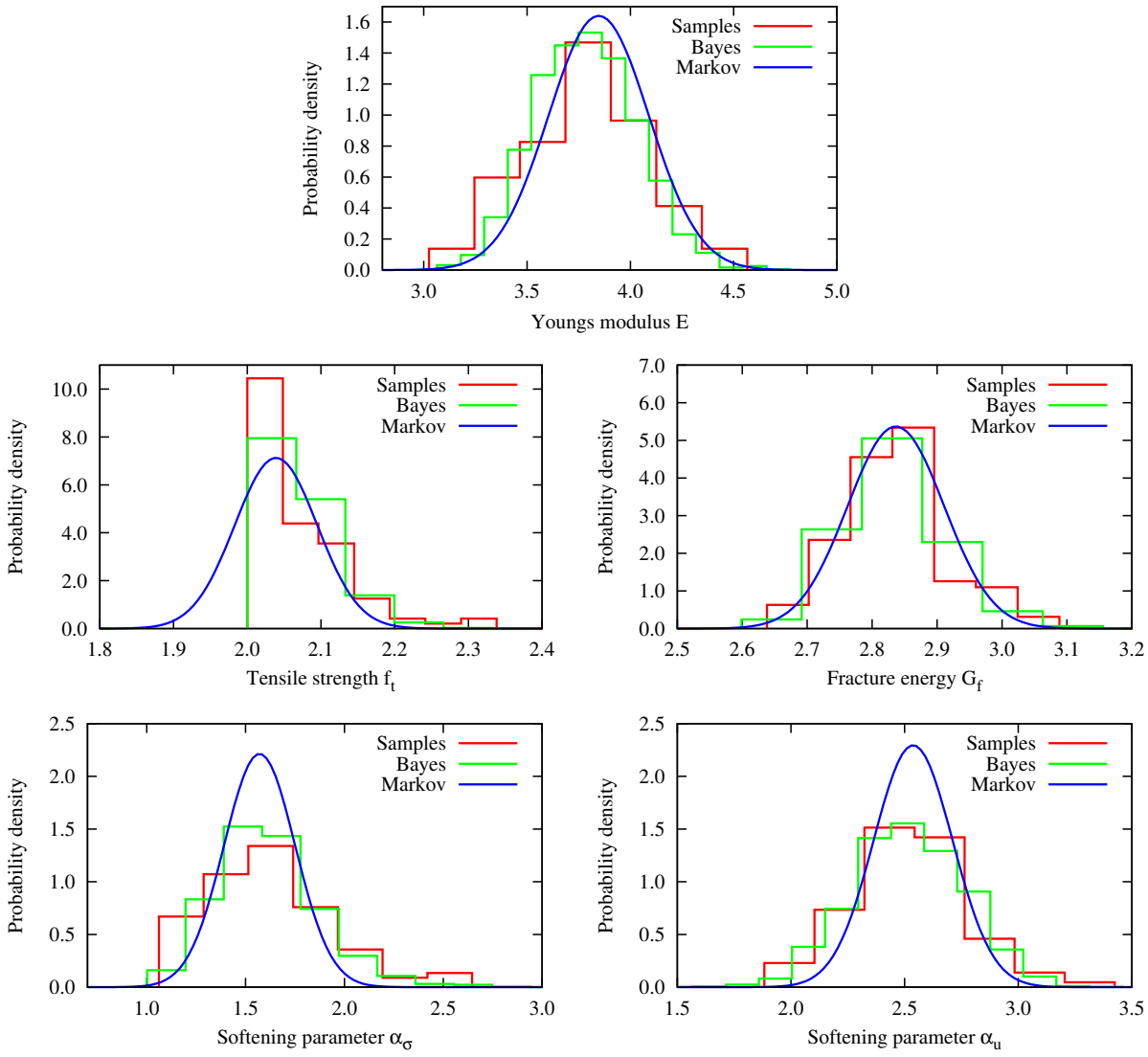


Figure 8: Estimated parameter distributions for the wedge splitting test

problems. A remarkable reduction of the numerical effort is possible by using meta-model approach to represent the likelihood function.

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

Beck, J. V. and K. Arnold (1977). *Parameter estimation in engineering and science*. New York: Wiley Interscience.

Bittani, S., G. Maier, and A. Nappi (1984). Inverse problems in structural elasto-plasticity: a Kalman filter approach. In A. Sawczuk and G. Bianchi (Eds.), *Plasticity Today*. London: Elsevier Applied Science.

Bolzon, G., R. Fedele, and G. Maier (2002). Parameter identification of a cohesive crack model by Kalman filter. *Computer Methods in Applied Mechanics and Engineering 191*, 2847–2871.

Cividini, A., G. Maier, and A. Nappi (1983). Parameter estimation of a static geotechnical model using a Bayes' approach. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts 20*, 215–226.

Fedele, R., G. Maier, and M. Whelan (2006). Stochastic calibration of local constitutive models through measurements at the macroscale in heterogeneous media. *Computer Methods in Applied Mechanics and Engineering 195*, 4971–4990.

Furukawa, T. and J. W. Pan (2009). Stochastic identification of elastic constants for anisotropic materials. *International Journal for Numerical Methods in Engineering 81*, 429–452.

Hastings, W. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika 57*, 97–109.

He, S., Q. Wu, J. Wen, J. Saunders, and R. Paton (2004). A particle swarm optimizer with passive congregation. *BioSystems 78*, 135–147.

Hofmann, M., T. Most, and G. Hofstetter (2009). Parameter identification for partially saturated soil models. In *Proc. 2nd Intern. Conf. Computational Methods in Tunnelling, Ruhr University Bochum, Germany, September 9-11, 2009*.

Kennedy, J. and R. Eberhart (1995). Particle swarm optimization. In *Proceedings of the IEEE International Conference on Neural Networks*, Volume 4.

Knabe, T., M. Zimmerer, T. Most, and T. Schanz (2009). Identification of constitutive parameters for geomaterials modeling using an optimization strategy. In *Proc. 2nd Intern. Conf. Computational Methods in Tunnelling, Ruhr University Bochum, Germany, September 9-11, 2009*.

Ledesma, A., A. Gens, and E. E. Alonso (1996). Estimation of parameters in geotechnical backanalysis – I. Maximum likelihood approach. *Computers and Geotechnics 18*, 1–27.

Li, L., Z. Huang, F. Liu, and Q. Wu (2007). A heuristic particle swarm optimizer for optimization of pin connected structures. *Computers and Structures 85*, 340–349.

Most, T. (2005). *Stochastic crack growth simulation in reinforced concrete structures by means of coupled finite element and meshless methods*. Ph. D. thesis, Bauhaus-Universität Weimar, Germany.

Most, T. and C. Bucher (2008). New concepts for moving least squares: An interpolating non-singular weighting function and weighted nodal least squares. *Engineering Analysis with Boundary Elements 32*, 461–470.

Most, T., G. Hofstetter, M. Hofmann, Novàk, and D. Lehky (2007). Approximation of constitutive parameters of material models by artificial neural networks. In B. H. V. Topping (Ed.), *Proc. 9th Intern. Conf. on the Application of Artificial Intelligence to Civil, Structural and Environmental Engineering, St. Julians, Malta, September 18-21, 2007*. Civil-Comp Press.

Novák, D. and D. Lehký (2006). ANN inverse analysis based on stochastic small-sample training set simulation. *Engineering Applications of Artificial Intelligence 19*, 731–740.

Orlande, H. R. B., M. J. Colaco, and G. S. Dulikravich (2008). Approximation of the likelihood function in the Bayesian technique for the solution of inverse problems. *Inverse Problems in Science and Engineering 16*, 677–692.

Trunk, B. (1999). *Einfluss der Bauteilgrösse auf die Bruchenergie von Beton*. Ph. D. thesis, Eidgenössische Technische Hochschule, Zürich, Switzerland.

Unger, J. (2009). *Neural networks in a multiscale approach for concrete*. Ph. D. thesis, Bauhaus-Universität Weimar.

Unger, J. F. and C. Könke (2010). An inverse parameter identification procedure assessing the quality of the estimates using Bayesian neural networks. *Applied Soft Computing Journal, submitted*.