A FUNCTIONAL BAYESIAN METHOD FOR THE SOLUTION OF INVERSE PROBLEMS WITH SPATIO-TEMPORAL PARAMETERS

AUTHORS:

Zenon Medina-Cetina
International Centre for Geohazards / Norwegian Geotechnical Institute
Roger Ghanem
Amy L. Rechenmacher
University of Southern California

CORRESPONDENCE:

Zenon Medina-Cetina
International Centre for Geohazards/ Norwegian Geotechnical Institute
Postboks 3930 Ullevål Stadion 0806 Oslo Norway
Phone (+47) 22021008
Fax (+47) 22 23 04 48
E-mail: zenon.medina-cetina@ngi.no

ABSTRACT

A Functional Bayesian (FB) methodology is introduced for the calibration of constitutive parameters spatially distributed within a model. The probabilistic solution to the inverse problem consists of assimilating the uncertainty captured from the actual material responses into the material parameters. A case study is introduced to illustrate the applicability of the method, where a soil model built in LS-DYNA is parameterized using surface displacement fields read from stereo digital images taken during a series of triaxial tests performed under similar conditions. The implementation of the FB method yields probability density functions of the parameters and its corresponding correlation structure. The parameters field is efficiently sampled using the Polynomial Chaos Decomposition method (PC) which allows for spatial non-stationary and non-Gaussian material representations. The posterior integration is performed via Markov Chain Monte Carlo techniques. Results show extended inferences about the material behaviour due to probabilistic description of the material variability.

Keywords:

Functional Bayesian, Probabilistic calibration, Stereo digital images, Inverse problems, Triaxial tests, Polynomial Chaos Decomposition, 3D Digital Image Correlation

INTRODUCTION

This work introduces a probabilistic methodology for the calibration of constitutive models based on a Functional Bayesian method (FB) (Medina-Cetina, 2006). The probabilistic methodology relies on two main components: first, in the collection of physical evidence containing uncertainty measures about the material behavior; and second, in the inclusion of these measurements into the computational solution of inverse problems solved by the Bayesian paradigm. In this work, the physical evidence consists in 3D surface full-field displacements captured by stereo digital images over the boundary of deforming specimens. Also, the material modelling implemented into he Bayesian solution introduces a functional spatial representation of the material properties by utilizing the Polynomial Chaos Decomposition method (PC) (Sakamoto and Ghanem; 2002a and 2002b). This approach allows for the sampling of non-Gaussian and non-stationary parameter fields, which is an adequate deductive assumption for starting the calibration.

A case study is presented describing the calibration of a soil model as a way to illustrate the applicability of the method. The soil model includes linear elastic constitutive parameters assigned to different 'material' regions distributed inside a 3D Finite Element Model (3D-FEM) built in LS-DYNA (2003). The soil model is introduced in Figure 1a containing five vertical layers and three concentric layers split vertically in two sections allowing for up to 25 'materials'. Figure 1b presents the spatial location of the materials as projected over the normalized vertical cross section indicated in Figure 1a. Notice that this projection is the same for ±90° around the indicated position.

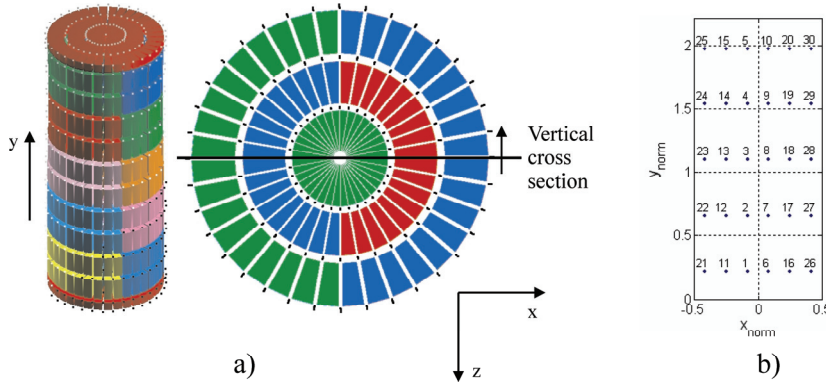

Figure 1. a) Soil model showing the spatial distribution of the 'materials', and b) location of the sampling materials as projected over the specimen normalized vertical cross section (locations 1-6, 2-7, 3-8, 4-9, 5-10 are assumed to belong to the same 'material' regions).

Estimates of the statistics required for the Bayesian formulation are obtained from a control-based experimental database consisting of a series of triaxial tests on sand specimens where stereo digital images captured local kinematic effects over the samples' surfaces (Rechenmacher et al, 2005; Medina-Cetina and Rechenmacher, 2006). This approach allows for assimilating the uncertainty generated by the global and local material heterogeneous responses into the model calibration.

The solution to the inverse problem is obtained by integrating the posterior defined by the Bayesian paradigm. The posterior integration requires defining the a-priori marginal pdfs and correlation structure of the constitutive parameters, so that the prior can be formulated; it also requires defining the covariance matrices of the observations and of the predictions, so that the likelihood can be formulated. These components of information rely on inferences from observations over the same experiment repeated several times, and on predictions of the same model based on multiple simulations. Once the prior and likelihood are properly defined, the posterior is integrated via Markov-Chain Monte Carlo (MCMC) using the Metropolis-Hastings (MH) rule (Robert and Casella, 2004).

RATIONALE

The Bayes theorem naturally defines the solution to the inverse problem as

$$\pi(\boldsymbol{\theta} \mid \mathbf{d}_{obs}) = \frac{f(\mathbf{d}_{obs} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta} \mid \mathbf{u})}{\int f(\mathbf{d}_{obs} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta} \mid \mathbf{u})d\boldsymbol{\theta}} = \frac{f(\mathbf{d}_{obs} \mid \boldsymbol{\theta}, g(\boldsymbol{\theta}))\pi(\boldsymbol{\theta} \mid \mathbf{u})}{\int f(\mathbf{d}_{obs} \mid \boldsymbol{\theta}, g(\boldsymbol{\theta}))\pi(\boldsymbol{\theta} \mid \mathbf{u})d\boldsymbol{\theta}} \qquad \text{(Eq. 1)}$$

where the prior $\pi(\boldsymbol{\theta} \mid \mathbf{u})$ introduces the a-priori state of information associated to the a set of parameters $\boldsymbol{\theta}$, that eventually may be dependent on another set of parameters $\mathbf{u}$ known as hyper-parameters. The likelihood $f(\mathbf{d}_{obs} \mid \boldsymbol{\theta})$ represents the a-priori state of information associated to the potential of the parameters $\boldsymbol{\theta}$ to match the observations $\mathbf{d}_{obs}$ or to help to match them if they are embedded into a predictive model g (i.e. 3D-FEM). The posterior $\pi(\boldsymbol{\theta} \mid \mathbf{d}_{obs})$ is thus the joint probability function between the a-priori states of information associated to both the prior and the likelihood.

POLYNOMIAL CHAOS DECOMPOSITION

Since the vector of constitutive parameters is referenced to space, it can be defined as the random field $\boldsymbol{\theta} = \boldsymbol{\theta}(\mathbf{X})$, where $\mathbf{X}$ is a set of spatial parameters contained in 3D-FEM as showed in Figure 1b. In this case, an efficient way to sample realizations of spatially distributed 'material' properties is by using the PC method, which can be implemented from Sakamoto and Ghanem (2002a and 2002b). This approach enables the simulation

of non-stationary and non-Gaussian random fields through functional representations, which enhances the capability of the calibration method to accommodate multiple forms of the marginal pdfs $f(\theta_j) = f(\theta(X_j))$ during the MCMC integration.

PRIOR

Any probabilistic model for the elastic parameters must respect known physical constraints such as positivity, together with statistics obtained from observations. A typical empirical representation of the a-priori information about elastic random parameters is the multivariate log-normal *prior*, which referenced to the spatial domain $\mathbf{X}$ can be defined as

$$\pi(\boldsymbol{\theta} \mid \mathbf{u}) \propto \exp\left[-\frac{1}{2}\left(\ln(\boldsymbol{\theta}(\mathbf{X})) - \boldsymbol{\theta}^*{}_{prior}(\mathbf{X})\right)^T \left(\mathbf{C}_{\boldsymbol{\theta}^*{}_{prior}(\mathbf{X})}\right)^{-1} \left(\ln(\boldsymbol{\theta}(\mathbf{X})) - \boldsymbol{\theta}^*{}_{prior}(\mathbf{X})\right)\right] \qquad \text{(Eq. 2)}$$

where $\boldsymbol{\theta}^*{}_{prior}(\mathbf{X})$ and $\mathbf{C}_{\boldsymbol{\theta}^*_{prior}}(\mathbf{X})$ are the mean vector and covariance matrix respectively of the underlying log-normal parameters $\boldsymbol{\theta}(\mathbf{X})$. These parameters represent the values of a random field evaluated at the spatial locations associated with 'material' regions in the computational model.

LIKELIHOOD

The likelihood defines a joint probability function of a set of 'actual' observations $\mathbf{d} = d_1, d_2, d_3, \ldots, d_n$, given a set of parameters $\boldsymbol{\theta}(\mathbf{X})$. If the observations show a random behaviour, a simple and common approach to define the shape of the likelihood is to consider that variables $d_1, d_2, d_3, \ldots, d_n$ are independent of each other, and that there is present an additive error due to modelling and experimentation defined as $\Delta\mathbf{d} = \Delta\mathbf{d}_{pred} + \Delta\mathbf{d}_{obs} = \{\mathbf{d}_{pred} - \mathbf{d}\} + \{\mathbf{d} - \mathbf{d}_{obs}\} = \mathbf{d}_{pred} - \mathbf{d}_{obs} = g(\boldsymbol{\theta}(\mathbf{X})) - \mathbf{d}_{obs}$, with $g(\boldsymbol{\theta}(\mathbf{X}))$ being linear or nonlinear, and Gaussian $\Delta\mathbf{d}$ with zero mean and covariance $\mathbf{C}_{\Delta\mathbf{d}} = C_{\mathbf{d}_{pred}} + C_{\mathbf{d}_{obs}}$. Thus, the resulting likelihood is defined as

$$f(\Delta\mathbf{d} \mid \boldsymbol{\theta}(\mathbf{X})) \propto \exp\left[\sum_{k=1}^{K} -\frac{1}{2}\left(g(\boldsymbol{\theta}(\mathbf{X})) - \mathbf{d}_{obs}\right)_k^T \left(\mathbf{C}_{\Delta\mathbf{d}}\right)_k^{-1} \left(g(\boldsymbol{\theta}(\mathbf{X})) - \mathbf{d}_{obs}\right)_k\right] \qquad \text{(Eq. 3)}$$

with k representing the different deformation stages where observations and predictions are evaluated.

FUNCTIONAL BAYESIAN POSTERIOR

The posterior $\pi(\boldsymbol{\theta}(\mathbf{X}) \mid \mathbf{d}_{obs})$ represents the joint density function of a vector set of spatially varying 'material' parameters $\boldsymbol{\theta}(\mathbf{X})$, conditioned on a set of displacement observations $\mathbf{d}_{obs}$ which are obtained over the sample's surface at different stages of deformation k and which coincide with the predictions $g(\boldsymbol{\theta}(\mathbf{X}))$ in space and time. Therefore, the posterior for a Gaussian likelihood and log-normal prior can be formulated as

$$\pi(\boldsymbol{\theta}(\mathbf{X}) \mid \mathbf{d}_{obs})_{Log} \propto \exp\left[\begin{array}{l} \sum_{k=1}^{K} -\frac{1}{2}\left(g(\boldsymbol{\theta}(\mathbf{X})) - \mathbf{d}_{obs}\right)_k^T \left(\mathbf{C}_{\Delta \mathbf{d}}\right)_k^{-1} \left(g(\boldsymbol{\theta}(\mathbf{X})) - \mathbf{d}_{obs}\right)_k \\ -\frac{1}{2}\left(\ln(\boldsymbol{\theta}(\mathbf{X})) - \boldsymbol{\theta}^*{}_{prior}(\mathbf{X})\right)^T \left(\mathbf{C}_{\boldsymbol{\theta}^*{}_{prior}(\mathbf{X})}\right)^{-1} \left(\ln(\boldsymbol{\theta}(\mathbf{X})) - \boldsymbol{\theta}^*{}_{prior}(\mathbf{X})\right) \end{array}\right] \quad \text{(Eq. 4)}$$

POSTERIOR INTEGRATION

The integration of the posterior is performed via MCMC method, as an efficient way to sample $\pi(\boldsymbol{\theta}(\mathbf{X}) \mid \mathbf{d}_{obs})$ of multivariate complex shapes where no analytical solutions are available. Also, a nice property about MCMC is that it converges to the target joint density as the sample integration grows.

For this work, the decision rule that determines which samples are 'accepted' or 'rejected' is the Metropolis-Hastings (MH) criteria (Robert and Casella, 2005). Under these premises, the posterior integration at the MH 'state' of the chain *s + 1* iteration is obtained by sampling a candidate point **Y** from a proposal distribution $q(\cdot \mid \hat{\boldsymbol{\theta}}_s(\mathbf{X}))$, where the candidate point **Y** is accepted or rejected as the next step of the chain with probability given by:

$$\alpha(\hat{\boldsymbol{\theta}}_s(\mathbf{X}), \mathbf{d}_{obs}) = \min\left\{1, \frac{\boldsymbol{\pi}(\mathbf{Y} \mid \mathbf{d}_{obs}) q(\hat{\boldsymbol{\theta}}_s(\mathbf{X}) \mid \mathbf{Y})}{\boldsymbol{\pi}(\hat{\boldsymbol{\theta}}_s(\mathbf{X}) \mid \mathbf{d}_{obs}) q(\mathbf{Y} \mid \hat{\boldsymbol{\theta}}_s(\mathbf{X}))}\right\} \quad \text{(Eq. 5)}$$

For the MCMC sampling the distribution of interest $f(\cdot \mid \mathbf{d}_{obs})$ appears as a ratio, so that the constant or proportionality cancels out. Additionally, the evaluation of the posterior requires discarding the first iterations called the burn-in points, before it reaches the stationary condition from which the statistical inferences are generated.

CASE STUDY

In order to illustrate the applicability of FB method, this section introduces results of the probabilistic calibration of one triaxial test, conditioned on data captured during the deformation of the specimen (0 – 0.2% of axial strain), assumed to be linear elastic. The proposed constitutive model is the linear elastic, with the Young's Modulus E as the only spatially varying parameter within the model. The initial estimate of $E = 92.33$ MPa, calculated as the secant from the triaxial stress-strain curve. Values of the Poisson ratio and density are taken as constants for all the 'material' regions ($\nu = 0.3$ and $\rho = 1717.13$ kg/m^3).

The components for the formulation of the posterior defined by equation 4 include a Gaussian-type likelihood (equation 3), a log-normal-type prior (equation 2), log-normal marginal pdfs $f(\theta(X_j))$ with the test mean and standard deviation taken from the series of experiments ($E = 92.33$ MPa and $\sigma_E = 45.98$ MPa respectively), an isotropic correlation structure $\rho_{\boldsymbol{\theta}(\mathbf{X})}$ with random correlation length parameter δ modelled as a lognormal variable ($\bar{\delta} = 0.5$, and $\sigma_\delta = 2.0$), and the PC definition of the random field $\theta(X)$ embedded into the prior (third-order, fourth-dimensional). Medina-Cetina (2006) presents the details on the estimates on each of the elements required for the Bayesian formulation.

The MCMC integration starts with the condition $\hat{\boldsymbol{\theta}}_{s=0}(\mathbf{X}) = \boldsymbol{\theta}_{prior}(\mathbf{X})$, meaning that each one of the constitutive parameters lying inside the model assumes the global estimate of the Young's modulus. The 'accepted' estimates during the MCMC-MH are divided in the burn in and the stationary phases. Convergence to the stationary condition is continuously checked by evaluating the trend of the mean of the samples and the mixing through the distribution of the decision parameter α. The stationarity condition is assessed by computing the running average for each vector of estimates $\hat{\boldsymbol{\theta}}_s(\mathbf{X})$ at step s, and by computing thereafter the first derivative function which indicates the inflection point between the burn in and the stationary phases. In this case, the stationary domain starts approximately at the 4,000th sample, delimiting the domain from which statistical inferences can be made. The sequence of sample estimates $\hat{\boldsymbol{\theta}}_s(\mathbf{X})$ corresponding to the burn in and the stationary phases are presented in Figure 2. The stationary domain includes 3,500 samples and five full periods of data, which are considered adequate for generating statistical inferences. Estimates from the burn in phase show the sampling effort of the calibration method to move away from the 'homogeneous' condition known a-priori, while estimates included in the stationary phase are more likely to achieve the target performance of the actual test.

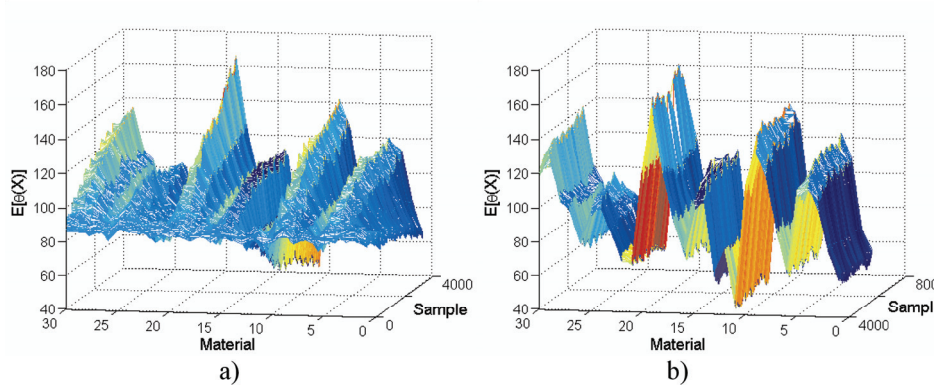

Figure 2. MCMC-MH sampling of Young's moduli during the a) burn in and b) stationary phases.

Regarding the 'material' parameters, Figures 3 shows the mean and standard deviation of the Young's modulus estimates. From Figure 3a, it is observed that the mean distribution exhibits the highest values of Young's moduli on the upper segment of the specimen, slightly deviated to the right; while the lowest values are concentrated almost uniformly at the bottom. Also, relative lower values are observed along the boundaries of the specimen. Figure 3b shows the standard deviation surface following a similar trend than the mean surface, with the highest standard deviation values in the upper segment, also slightly deviated to the right, and with the lowest values at the bottom.

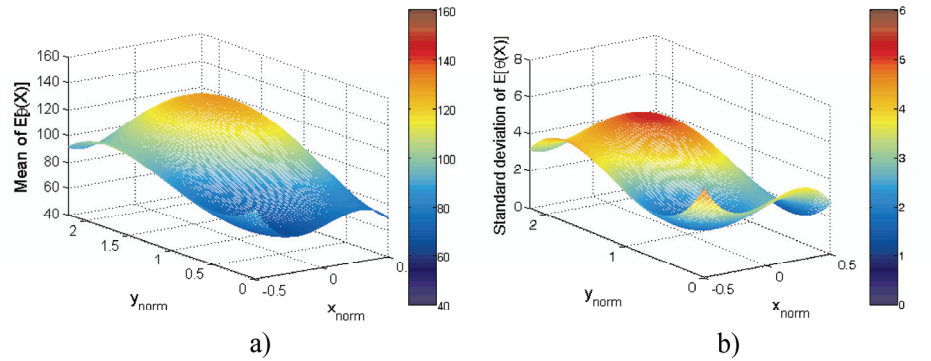

Figure 3. Posterior of $E[\theta(X)]$: a) mean and b) standard deviation.

Furthermore, it is also possible to depict the marginal pdfs of the constitutive parameter over each 'material' region, and their corresponding correlation structure. Three typical relative density distributions of $\hat{\boldsymbol{\theta}}_s(\mathbf{X})$ evaluated at control points $X_{j=1}$(-0.25, 0.22), $X_{j=3}$(-0.25, 1.10) and $X_{j=5}$ (-0.25, 1.99) are introduced in Figure 4. Figure 4a shows an increment on the uncertainty from the bottom to the top of the specimen as discussed on Figure 3b, and the correspondence with the order of magnitude of the mean values as discussed on Figure 3a. On the other hand, Figure 4b presents the corresponding graphical representation of the empirical correlation matrix of the $\hat{\boldsymbol{\theta}}_s(\mathbf{X})$ estimates.

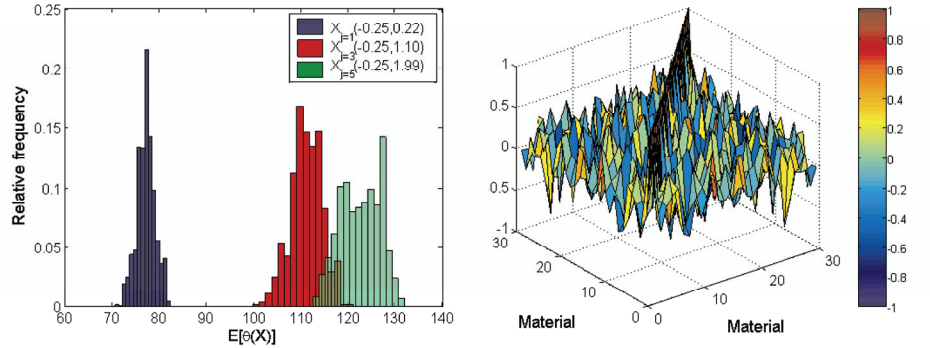

Figure 4. Statistics from Young's modulus posterior: a) relative frequency histograms at three control points over the model normalized vertical cross section (Figure 1) and b) graphical representation of the parameters correlation matrix.

Finally, a comparison between the predictions of the 3D-FEM model and the actual observations is presented in Figure 5 (at 0.2% of axial strain). For this particular case, predictions are assessed solving the forward problem based on the mean estimates of $\hat{\boldsymbol{\theta}}(\mathbf{X})$, while observations represent displacement fields ***u***, ***v*** and ***w*** as captured by the stereo digital images during the triaxial test. From this figure it is observed that there is a very good agreement between vertical displacements, and certainly a more limited model response on the horizontal displacements. The relative differences between predictions and observations are closer in the vertical displacements since the order of magnitude is different with respect to the horizontal displacements. Also, it is worth noticing that these differences are still within the order of accuracy estimated by the 3D-Digital Image Correlation measurements 3D-DIC (approximately ± 0.02 mm); and also that some of them may be associated to the limitations of the model itself.

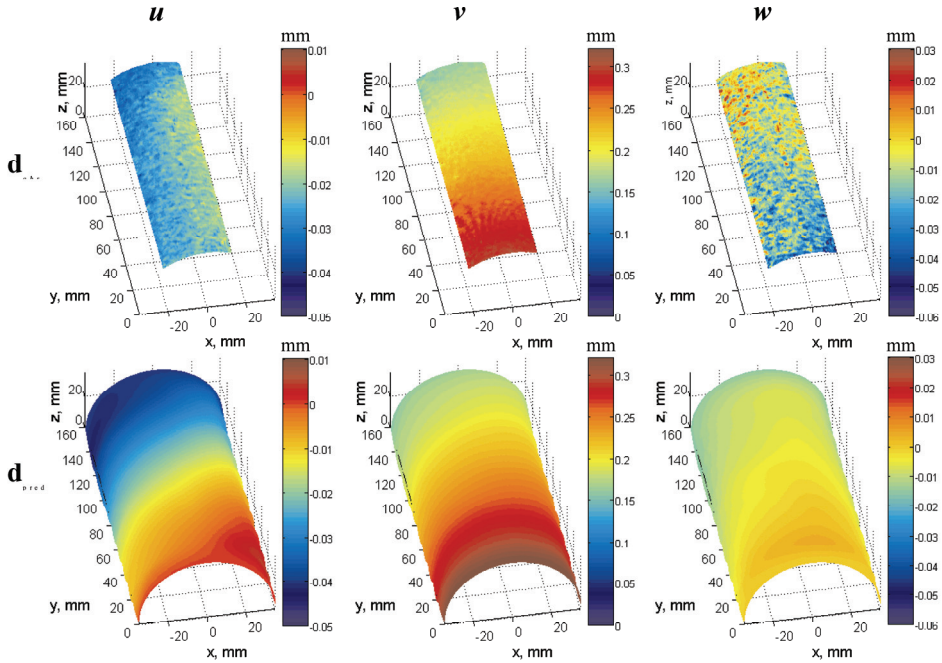

Figure 5. Comparison between observations (upper row) and predictions (lower row) calculated for the mean of the posterior estimates at 0.2% of axial strain.

CONCLUSIONS

This work introduced the Functional Bayesian FB method as an innovative approach for the calibration of a soil model. An appealing characteristic of the FB method is its ability to sample multidimensional non-stationary non-Gaussian constitutive parameters. The Polynomial Chaos PC method proved to be an efficient formulation for sampling this particular 'material' field.

An innovative technique consisting in the coupling of triaxial testing and 3D digital imaging allowed for the population of the experimental database used for the validation of the FB method. A unique characteristic of these methods was the capturing of global and local deformation effects on soil deforming specimens.

A case study was introduced to illustrate the applicability of the probabilistic calibration approach. The most relevant of the method application is the description of the constitutive parameters as pdfs.

The goal of developing a probabilistic methodology for the calibration of constitutive models was reached and validated when applied for the elastic model. A robust computational algorithm was developed and applied for the case study facilitating the understanding of the required components for the probabilistic calibration.

REFERENCES

1. Livermore Software Technology Corporation, http://www.lstc.com, 2003.

2. Medina-Cetina, Z. and Rechenmacher, A.L., "Image-based sensing of 3-D displacements for enhanced soil model calibration," Proceedings of the GeoCongress 2006: Geotechnical Engineering in the Information Technlology Age. Atlanta USA, Feb 26 - March 1, 2006.

3. Medina-Cetina Z. "Probabilistic Calibration of a Soil Model". PhD Dissertation, The Johns Hopkins University, Baltimore MD USA, 2006 https://dspace.library.jhu.edu/handle/1774.2/3069).

4. Rechenmacher, A.L., Medina-Cetina, Z. and Ghanem R. "Calibration of Heterogeneous, Probabilistic Soil Models," Proceedings of the 16th International Conference on Soil Mechanics and Geotechnical Engineering, Osaka Japan, Sept 12-16, 2005.

5. Rechenmacher A.L. and Medina-Cetina Z. "Calibration of constitutive models with spatially varying parameters". Journal of Geotechnical and Geoenvironmental Engineering, submitted.

6. Robert C. P. and Casella G. Monte Carlo Statistical Methods. Springer, New York, 2005.

7. Sakamoto S. and Ghanem R. "Simulation of multi-dimensional non-Gaussian non-stationary random fields". Probabilistic Engineering Mechanics, 17 (2), 167-176, 2002a.

8. Sakamoto S. and Ghanem R. "Polynomial chaos decomposition for the simulation of non-Gaussian non-stationary stochastic processes". Journal of Engineering Mechanics, 128 (2), 190-201, 2002b.