SOIL PARAMETER IDENTIFICATION INCLUDING SOIL HETEROGENEITY

ANNEKE HOMMELS and FRANS MOLENKAMP
Department of Civil Engineering and Applied Geosciences,
Delft University of Technology, The Netherlands

ABSTRACT

Geomechanical models are indispensable for reliable design of engineering structure, simulating processes and hazard and risk evaluation. These models are however far from perfect. Errors are introduced by fluctuations in the input, by poorly known parameters in the model and due to geological uncertainties. Especially geological uncertainties are still not adequately addressed. To overcome these problems, the Ensemble Kalman filter has been implemented, in order to incorporate measurements into the deterministic model to improve the poorly known parameters and consequently the model results. This allows for observations of on-going processes to be used for enhancing the quality of subsequent model predictions based on improved knowledge of the soil parameters. The Ensemble Kalman filter analyses the state of the system each time data becomes available and improves the model state and the uncertain soil parameters using these data.

Almost all natural soils have highly variable properties and they are rarely homogeneous. Soil heterogeneity can be classified into lithological heterogeneity and spatial soil variability, which are modelled using Sequential Indicator Simulation and the Random Finite Element Method respectively.

In a case study, based on the construction of a road embankment on soft clay the Ensemble Kalman filter is not only used for a straight forward identification of several uncertain soil parameters of the soil layers below the embankment, but also for assessing the lithological variability using Sequential Gaussian Simulation and the spatial soil variability by means of the Random Finite Element Method..

INTRODUCTION

In geotechnical engineering context inverse modelling or back analysis consists in finding the values of the mechanical parameters, or of other quantities characterizing a soil or rock mass, that when introduced in the stress analysis of the problem under examination lead to results (e.g. displacements, stresses) as close as possible to the corresponding in situ measurements. About 25 years ago the use of data from in situ measurements as input for inverse modelling was proposed by Cividini et al. (1983) and Gioda (1985). By then inverse modelling was mostly referred to as back analysis or parameter estimation. The optimal state of the system is obtained by minimizing the discrepancy between the observed values in the system and the forecasted state of the system within a certain time interval. These type of methods are called variational methods. In other fields of science another type of inverse modelling has been developed: sequential modelling or filtering. In a filter, the state of the system is analysed each time data becomes available. The total computational effort compared to the variational methods is the same, but a filter is easier to implement and the implementation of the adjoint model is not required.
Recent developments have shown a powerful sequential technique indicated as the Ensemble Kalman filter (EnKF). The general formulations of this filter will be discussed in the following sections.

Soil heterogeneity can be classified into two main categories. The first is lithological heterogeneity, which can be manifested in the form of thin soft or stiff layers embedded in a stiffer or softer media (layered cake model) or the inclusion of pockets of different lithology within a more or less uniform soil mass. The second source of heterogeneity can be attributed to inherent spatial soil variability, which is the variation of soil properties from one point to another in space due to different deposition conditions and different loading histories (Elkateb et al., 2003). Since in geostatistics the variable is considered to be a realization of a random function, it is possible to simulate an infinite number of realizations to represent the variability and spatial characteristics in the experimental data (Bastante et al., 2008).
Indicator kriging can be used to characterize the spatial variabilty of the categorical variables (e.g. the presence of sand, clay, silt etc at a certain location) to reflect the geometric patterns (Goovaerts, 1997). However when kriging is used a smoothing effect of the reality is usually produced, which typically underestimates the real data dispersion (Bastante et al., 2008). The Random Finite Element Method (RFEM), combines the Finite Element Method (FEM) with the random field theory. A random field is generated using the Local Average Subdivision technique (LAS) (Fenton and Vanmarcke, 1991), which describes the spatial variability of geotechnical parameters througout a soil layer.
The general formulation of Sequential Indicator Simulation (SISIM) as wel as the Random Finite Element Method (RFEM) will be explained in the next section.
The effects of both types of soil heterogeneity on the soil parameter identification process using the Ensemble Kalman filter will be presented.

ENSEMBLE KALMAN FILTER

Evensen introduced the Ensemble Kalman Filter (EnKF) in 1994 and the theoretical formulations as well as an overview of several applications are described in Evensen (1994 and 2003). The EnKF was designed to resolve two major problems related to the use of the Extended Kalman filter (EKF). The first problem relates to the use of an approximate closure scheme in the Extended Kalman Filter, and the other one to the huge computational requirements associated with the storage and forward integration of the error covariance matrix *P*. For further details the reader is referred to the references.

In the EnKF, an ensemble of possible state vectors, which are randomly generated using a Monte Carlo approach, represents the statistical properties of the state vector. The algorithm does not require a tangent linear model, which is required for the EKF, and is very easy to implement.

At initialisation, an ensemble of *N* initial states $(\xi_N)_0$ are generated to represent the uncertainty at time step $k=0$. In the measurement update step, the matrix $E^f(k+1)$ defines an approximation of the assimilated covariance matrix $P^a(k+1)$.

The time update equations for the EnKF for each estimate of the state are:

$$\xi_i^f(k+1) = f(\xi_i^a(k)) + G(k)w(k) \tag{1.1}$$

$$x^f(k+1) = \frac{1}{N}\sum_{1}^{N}\xi_i^f(k+1) \tag{1.2}$$

$$E^f(k+1) = [\xi_1^f(k+1) - x^f(k+1),\ldots,\xi_N^f(k+1) - x^f(k+1)] \tag{1.3}$$

In equations (1.1) and (1.2), $G(k)$ and $w(k)$ are the noise input matrix and process noise respectively and x^f is the 'forecasted' statevector.

In the measurement update step, the assimilated error covariance matrix $P^a(k+1)$ can be calculated as follows:

$$P^a(k+1) = \frac{1}{N-1}E^f(k+1)(E^f(k+1))^T \tag{1.4}$$

Equation (1.4) is never really calculated due to huge computational requirements, as mentioned before, but is split up:

$$K(k+1) = \frac{E^f((E^f)^T H)}{N-1}\left[\frac{HE^f}{N-1}((E^f)^T H^T) + R\right]^{-1} \tag{1.5}$$

$$\xi_i^a(k+1) = \xi_i^f(k+1) + K\left\{y_o(k+1) - H\xi_i^f(k+1) + \varepsilon\right\} \tag{1.6}$$

In equation (1.5), *K* is the Kalman gain, *H* is the observational operator and *R* is the measurement noise covariance matrix with zero mean. In equation (1.6): y_o is the vector, which contains the observations at time step $k+1$ and ε is an additional noise (Evensen, 2003)

GEOLOGICAL UNCERTAINTY

Usually in the field of geomechanics the Finite Element Method (FEM) is used for numerical simulations. In this method, one particular value for a soil parameter is assigned to a soil mass or a soil layer, which remains constant throughout the mass or the layer. However, in nature, this is not really the case and soil is a heterogeneous material. Soil heterogeneity can be subdivided into two main categories. The first is lithological heterogeneity, which can be manifested in the form of thin stiff or soft layers embedded in a softer or stiffer media or the inclusion of pockets of different lithology within a more or less uniform soil mass. The second source of heterogeneity can be attributed to inherent spatial soil variability, which is the variation of soil properties from one point to another in space due to different deposition conditions and different loading histories. For this paper the authors will simulate lithological heterogeneity using Sequential Indicator Simulation (SISIM) and the spatial variability will be simulated using the Random Finite Element Method (RFEM). These techniques will be explained in the following subsections.

Sequential Indicator Simulation

In genaral: consider the problem of modelling the categorical variable (e.g. soiltype) at location **u.** The categorical variable can be expressed as a series of K indicator variables:

$$i(\mathbf{u};k) = 1 \quad \text{if category } k \text{ prevails at location } \mathbf{u}$$

$$i(\mathbf{u};k) = 0 \quad \text{otherwise} \qquad (1.7)$$

$$\text{for } k = 1, \ldots, K$$

An indicator variable is often interpreted as the probability for a category to prevail at a certain location: the probability is 1 if it does prevail and 0 if it does not prevail. Hard local data are coded into 0 and 1, soft or imprecise data can be coded with values between 0 and 1.

The local hard and soft data are used to estimate the distribution of the categorical variable at locations where no information is available. Simple or ordinary kriging (depending on the amount of original data) is used for simulating the prevailing indicator variable, and therefore the name Indicator Kriging (IK). SISIM applies IK in a sequential fashion where a precise category is drawn using Monte Carlo simulation at each location. All locations are visited sequentially with an increasing level of conditioning. A random path is followed to avoid artifacts (Deutsch and Journel, 1998 and Deutsch, 2006).

Random Finite Element Method

The RFEM combines the Finite Element Method (FEM) with the random field theory generated using the Local Average Subdivision (LAS) method (Fenton and Vanmarcke, 1991). In the FEM the uncertainty of a material is defined by its mean μ and its standard deviation σ. For the introduction of more spatial variability, the introduction of an additional statistical parameter, the spatial correlation length θ, is required. The spatial correlation length defines the distance beyond which there is minimal correlation and can be determined from for example CPT-data. A large value of θ indicates a strongly correlated mate-

rial, while a small value indicates a weakly correlated material. Now the random field can be generated and in the RFEM, this is done using LAS, based on a standard normal distribution (mean μ is zero and the standard deviation σ equals one) and a spatial correlation function $\rho(\tau)$. LAS generates a square random field by uniformly subdividing a square domain into smaller square cells, where each cell has a unique local average, which is correlated with surrounding cells. For an isotropic field, where θ is equal in all directions, the spatial correlation function $\rho(\tau)$ is given by

$$\rho(\tau) = \exp\left(-\frac{2}{\theta}|\tau|\right) \tag{1.8}$$

in which τ is the lag vector. The random field for a certain parameter E for cell i, based on a standard normal distribution, can then be transformed into for example a normal distribution:

$$E_i = \mu_E + \sigma_E Z_i \tag{1.9}$$

in which μ_E is the mean and σ_E is the standard deviation of parameter E; Z_i is the local average value for cell i.

Based on the given set of statistics of μ, σ and θ, there are an infinite number of possibilities for the random field. Again Monte Carlo simulations are used to express the spatial distribution (Hicks and Samy, 2002).

RESULTS

For this study, a four meter high embankment constructed on soft clay has been considered. A dedicated model for the determination of the settlement has been written using Smith and Griffiths (2004) . In the left of Figure 1 the three-layered foundation below the embankment is shown. The properties of the foundation, including the uncertain Young's modulii E are given in Table 1. The nodes at which the observations of the vertical displacement took place are indicated with a triangle. For the material model the purely elasto-plastic Mohr-Coulomb model is used. The groundwater table coincides with the original ground surface.

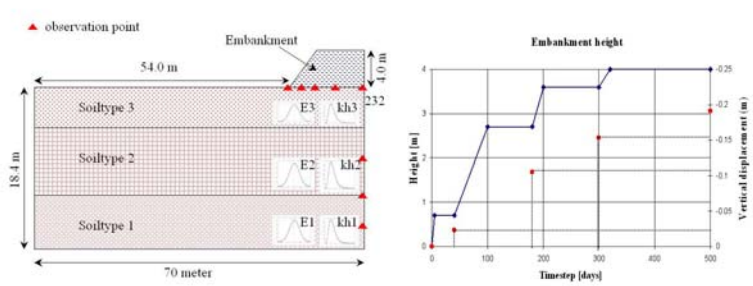

Figure 1: (left) three-layered foundation; the observation points are indicated with a red triangle; (right) embankment height (left axis) and vertical displacement (right) during the construction phases

The embankment is built in four lifts and at the end of each lift a consolidation period is considered. The graph in the right of figure 1 shows at the left axis the different stages during the construction versus the total time are shown, while at

the right axis the measured displacement versus the total time at the end of each lift at a certain observation point, is shown. The measurements were generated based on an assumed true state.

Soil layer	μ_E (kPa)	σ_E	ν	c	φ	ψ	γ(kN/m^3)	k_h(m/day)	k_v(m/day)
1	*3400*	*340*	*0.33*	*16*	*12.5*	*0*	*16.5*	*2.41E-05*	*2.94E-05*
2	*1400*	*140*	*0.33*	*22*	*13.5*	*0*	*15.6*	*8.34E-05*	*8.99E-05*
3	*4600*	*460*	*0.30*	*8*	*6.5*	*0*	*17.3*	*3.56E-04*	*1.03E-04*

Table 1: properties of the foundation below the embankment

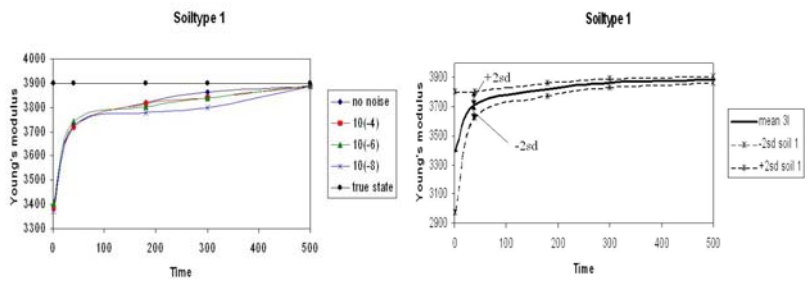

Figure 2: (left) influence of different measurement noises on the update process of the Young's modulus of soiltype 1; (right) decreasing uncertainty of the Young's modulus E of soiltype 1 in time

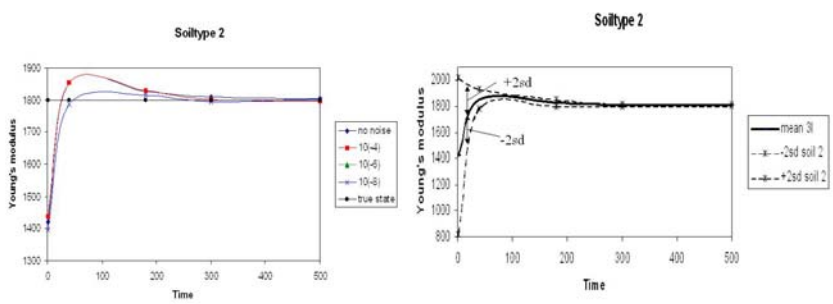

Figure 3: (left) influence of different measurement noises on the update process of the Young's modulus of soiltype 2; (right) decreasing uncertainty of the Young's modulus E of soiltype 2 in time

Several soil parameters were considered to be uncertain. Here only the results for the Young's moduli *E* are shown. For the three-layered foundation (left part of Figure 1) the update processes through time for the Young's modulus *E* of soil layer 1 is shown in the left part of Figure 2. In Figure 3 and Figure 4 the update processes through time is shown for soil layer 2 and 3 respectively. In the left parts of Figure 2, Figure 3 and Figure 4 the effect of different amounts of measurement noises on the update process are shown.

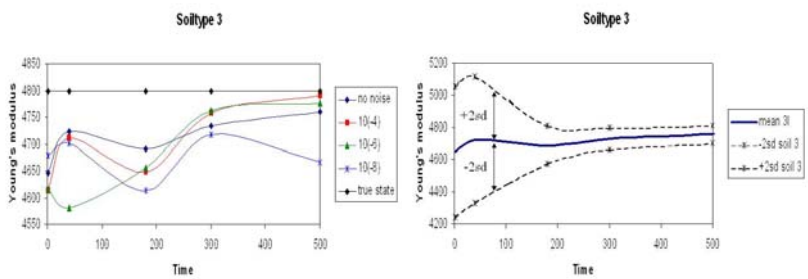

Figure 4: (left) influence of different measurement noises on the update process of the Young's modulus of soiltype 3; (right) decreasing uncertainty of the Young's modulus E of soiltype 3 in time

The properties of the normal distributions of the analysed Young's moduli *E* (no masurement noise) at the end of each lift for soil layer 1 are shown in the right part of Figure 2. In Figure 3 and Figure 4 these properties are shown for soil layers 2 and 3.

For soil layer 1 and 2 the true state is reached fast for different measurment noises and if the measurement noise decreases, the true state is reached more quickly (Figure 2 and Figure 3). However, in soil layer 3, the EnKF has some difficulties to reached the true state (Figure 4), except when noise measurement noise is included. This effect cannot be fully explained yet. The uncertainty of the Young's modulus E decreases tremendously in time for each layer.

Spatial variability in one-dimension

Consider now a vertical column below the crest of the embankment. The Young's moduli *E* are now spatially varying with depth using LAS. The soil column is loaded vertically (using an elastic model). In the left part of Figure 5, the uncertainty of the Young's modulus *E* in several subsequent elements at time step k=0 is shown; in the right part of Figure 5 the uncertainty of the Young's modulus *E* at time step k=100 is shown.

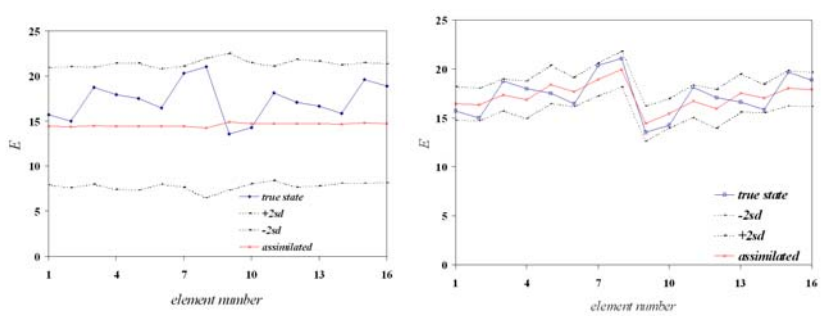

Figure 5: (left) uncertainty througout the column at k=0; (right) uncertainty througout the column at k=100

At the start of the update or assimilation process, the uncertainty of the Young's modulus *E* is very large. Figure 5 shows that there has been a large reduction of the parameter uncertainty. The true state can be found, for both loading steps,

between the range of the ensemble members and therefore the filter is performing well if spatial variability is included.

CONCLUSIONS

The case study of the construction of a road embankment, with three soil layers in the foundation, has shown that the Ensemble Kalman filter can be a very powerfull technique in the field of geomechanics. For soil layer 1 and 2 the true state is reached fast for different measurement noises. However, in soil layer 3, the EnKF has some difficulties to reached the true state, which cannot be fully explained yet. The uncertainty of the Young's modulus E decreases tremendously in time for each layer.

If spatial variability of the Young's modulus is taken into account, in the case of the vertically loaded soil column, the EnKF is also able to reduce the uncertainty of the Young's modulus.
The results of the incorporation of the lithological heterogeneity and the spatial variability on the inverse modelling process of the road embankment will be shown during the presentation.

ACRONYMS

EKF	Extended Kalman filter
EnKF	Ensemble Kalman Filter
FEM	Finite Element Method
IK	Indicator Kriging
LAS	Local Average Subdivision
RFEM	Random Finite Element Method
SISIM	Sequential Indicator Simulation

REFERENCES

Bastante, F.G., Ordóňez, C. Taboada, J. and Matías, J.M. (2008) *Comparison of indicator kriging, conditional indicator simulation and multiple-point statistics used to model slate deposits*, in Engineering Geology, 98, 50-59.

Cividini, A., Jurina, L. and Gioda, G. (1981) *Some aspects of 'characterization' problems in geomechanics*, in Int. Journal of Rock Mech. and Min. Science and Geom., 18, 487-503.

Deutsch, C.V. (2006) *A sequential indicator simulation program for categorical variables with point and block data: BlockSIS,* in Computers and Geosciences, 32, 1669-1681.

Deutsch, C.V. and Journel, A.G. (1998) *GSLIB: Geostatistical Software Library and User's Guide*, 2nd Edition, Oxford University Press.

Elkateb, T., Chalaturnyk, R. and Robertson, P.K. (2003) *An overview of soil heterogeneity: quantification and implications on geotechnical field problems*, in Canadian Geotechnical Journal, 40, 1-15.

Evensen, G. (1994) *Sequential data assimilation with non-linear quasi-geostrophic model using Monte Carlo methods to forecast error statistics*, in Journal of Geophysical Research, 99,143-162.

Evensen, G. (2003) *The Ensemble Kalman filter: theoretical formulation and practical implementation*, in Ocean Dynamics, 53, 343-367.

Fenton, G.A., and Vanmarcke, E.H. (1991) *Spatial variation in liquefaction risk assessment*, in Proceedings of the ASCE Geotechnical Engineering Congress, Boulder, Colorado, 594-607.

Gioda, G. (1985) *Some remarks on back analysis and characterization problems in geomechanics*, in Proc. 5th Int. Conf. On Numerical Methods in Geomechanics, Nagoya, 47-61.

Goovaerts, P. (1997) *Geostatistics for Natural Resources Evaluation*, 1st Edition, Oxford University Press.

Hicks, M.A. and Samy, K. (2002) *Influence of heterogeneity on undrained clay slope stability*, in The Quarterly Journal of Engineering Geology and Hydrogeology, 35, 41-49.

Smith, I.M. and Griffiths, D.V. (2004) *Programming the finite elememt method*, 4th Edition Chichester, Wiley and Sons.