

ICFEP-UQ

Short guidance

Version 1.0.0

June 3, 2024

Contents

1	Welcome to ICSEP-UQ guidance!	2
2	Theory basis	3
3	Sensitivity analysis	4
4	Uncertainty propagation	4
5	Model assessment	5
6	Model calibration	6
7	Example to get started	8

List of Code Listings

1 Welcome to ICFEP-UQ guidance!

ICFEP-UQ is a general purpose and Matlab-based toolbox for modelling uncertainty in physical and mathematical systems. The application of the code can be generalized to ABAQUS/PLAXIS/etc and all kinds of engineering problems. The code is organized as a set of modules centered around core capabilities in Uncertainty Quantification (UQ).

Product owner:	Truong Le
Lead developers:	Ningxin Yang
Contributors:	Master students

All primary functionalities in ICFEP-UQ are categorized as:

- Sequential Bayesian inference
- Enriching experimental design
- Sensitivity Analysis
- Dimensionality reduction based surrogate models
- User-defined loglikelihood for combined data sources (ongoing)

Software/system preparation

Dependencies required:

MATLAB® R2019a and later; windows;
Uqlab package (version 2.1);
Statistics and Machine learning toolbox,
...

Installation:

For MATLAB, Imperial holds license. For UQlab opensource package, you can download it from [here](#) and following the basic installation within.

Contact:

n.yang23@imperial.ac.uk

2 Theory basis

Probabilistic theory

Sum rule and product rule are the basis of Bayes theorem.

Relationship between integration and expectation.

Basis knowledge on first and second moments.

...

See [more details](#)

Bayesian inference

Frequentist vs Bayesian

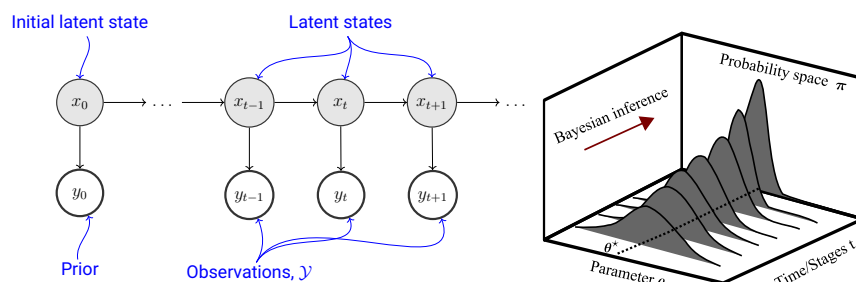
Bayes theorem

Inference - prior, likelihood and evidence

Sampling methods to posterior

Sequential Bayesian inference

In engineering problems, observations are acquired stage by stage based on construction/loading. Therefore, the Bayesian inference/model calibration is conducted step by step, which is also called *sequential Bayesian inference*. Based on a Markovian process, some problems are *filtering, smoothing and prediction*.



Components in UQ

Four components in uncertainty quantification

- Model assessment
- Model calibration

- Uncertainty propagation
- Sensitivity analysis

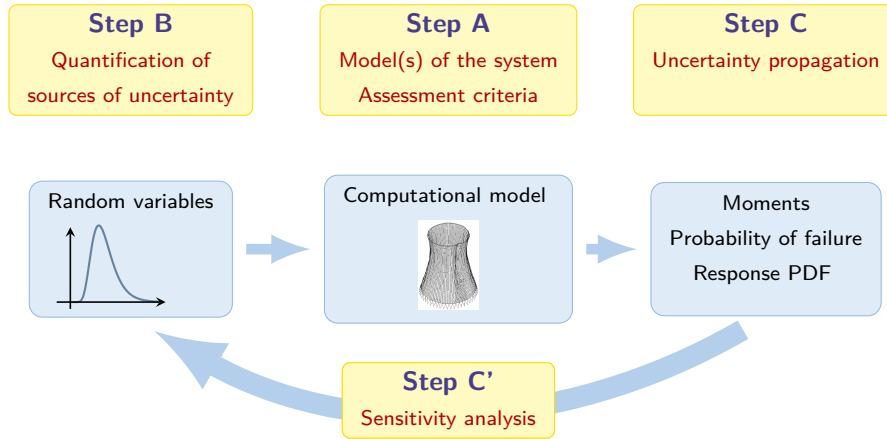


Figure 1: UQ framework

3 Sensitivity analysis

Sampling methods to input space

Sobol's indice

PC expansion-Sobol's indice

PCA-PCE-Sobol's indice

4 Uncertainty propagation

Prior predictive and posterior predictive distributions are very important indicators for assessing the predictive ability of the model and uncertainties propagation in output space.

The *prior predictive* distribution can be represented as:

$$\pi(\mathbf{y}) = \int_{\mathcal{D}_{\mathbf{x}}} \pi(\mathbf{x})\pi(\mathbf{y}|\mathbf{x})d\mathbf{x} \quad (1)$$

where \mathbf{y} is the simulator prediction, \mathbf{x} are the input parameters. It summarises the uncertainty about the model output considering the discrepancy model before calibration. In practice, it is used to determine whether the measured data

can be reproduced. It is essential for identifying and ruling out challenging inverse problems before proceeding with expensive calibration procedures.

The *posterior predictive* distribution can be similarly written as:

$$\pi(\mathbf{y}|\mathcal{Y}) = \int_{\mathcal{D}_{\mathbf{x}|\mathcal{Y}}} \pi(\mathbf{x}|\mathcal{Y})\pi(\mathbf{y}|\mathbf{x})d\mathbf{x} \quad (2)$$

It gives model predictions for future stages based on the current inferred parameters \mathbf{X} .

5 Model assessment

In the UQ setting, especially for geotechnical problems, engineers will use very complex and advanced constitutive models. Each run will cost hours to days, in this case, original simulator \mathcal{M} will and must be replaced by a cheap-to-evaluate surrogate $\tilde{\mathcal{M}}$ to accelerate the running speed.

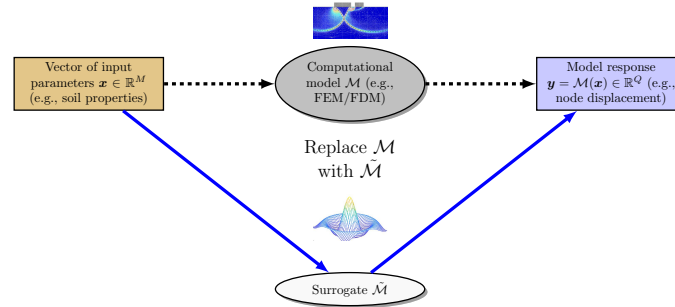


Figure 2: Surrogate to replace original simulator

Currently, ICSEP-UQ are using dimensionality reduction-based surrogates. Basically, *principal component analysis* is used for feature extraction and dimensionality reduction. Then a surrogate is built directly on these reduced-order spaces. And this will benefit in two ways: (1) accelerate the Bayesian inference process; (2) capture the connections among output points and give more accurate model predictions.

Dimensionality reduction techniques are only supporting:

- Principal component analysis

Some surrogates choices in ICSEP-UQ are:

- Polynomial chaos expansion
- Polynomial chaos kriging
- Stochastic spectral embedding

6 Model calibration

Bayesian inference or *model calibration* is the most challenging part in UQ. Because it requires designers having a high-level understanding of:

- Working flow of sequential Bayesian inference
- Error types from observations, model discrepancy and surrogates
- User-defined loglikelihood function

Connecting the reality with simulator $y \leftrightarrow \mathcal{M}(x)$ as:

$$y = \mathcal{M}(x) + \varepsilon$$

in which, y is the observation, $\mathcal{M}(x)$ is the model prediction and ε is the error term.

Normally, it is assumed that ε is following a zero-mean Gaussian distribution because of its simple and popular form $\varepsilon \in \mathcal{N}(\varepsilon|0, \Sigma)$. The covariance matrix Σ with N outputs points can be simple as an identical matrix like:

$$\Sigma = \sigma^2 \mathbf{I} = \sigma^2 \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \quad (3)$$

or holding a more general form:

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1N} \\ \sigma_{12} & \sigma_{22} & \dots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \dots & \sigma_{NN} \end{bmatrix} \quad (4)$$

The discrepancy ε follows a Gaussian distribution with unknown residuals

variance Σ . Likelihood function \mathcal{L} can be expressed as:

$$\begin{aligned}\mathcal{L}(\mathbf{x}|\mathcal{Y}) &= N(\mathbf{y}|\mathcal{M}(\mathbf{x}), \Sigma) \\ &= \frac{1}{\sqrt{(2\pi)\det(\Sigma)}} \exp\left(-\frac{1}{2}(\mathbf{y} - \mathcal{M}(\mathbf{x}))^\top \Sigma^{-1}(\mathbf{y} - \mathcal{M}(\mathbf{x}))\right)\end{aligned}\quad (5)$$

Note: \mathbf{y} is an observation vector, which is different from FE prediction $\mathbf{y} = \mathcal{M}(\mathbf{x})$.

In Equation (3), ε is non-informative, thus, its prior for the variance σ follows uniform distribution $\sigma \sim \mathcal{U}(0, \sigma_0^2)$, in which σ_0^2 should be related to the observation values (e.g., mean or maximum). ε now can be considered as some additional priors \mathbf{x}^ε along with forward parameters $\mathbf{x}^\mathcal{M}$ shown as Figure 3. UQlab default setting only supports *iid* assumption as shown Equation (3), which is very unrealistic for most engineering problems. Because *iid* assump-

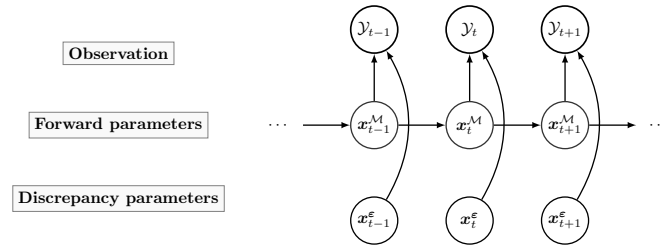


Figure 3: Sequential parameter updating

tion will lose all connections inside the discrepancy term. *Kernel methods* are good ways to capture the connections within the covariance matrix Σ in \mathbf{x}^ε , such as:

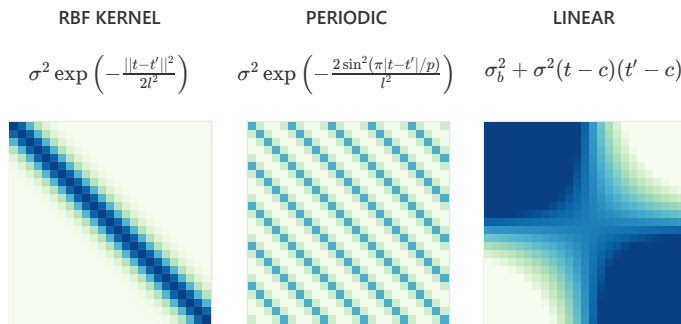


Figure 4: Some kernel choices

Things are not done yet. Even we know the the connection of discrepancy \mathbf{x}^ε can be reproduced by *kernel methods*. But we still need to give values

for initial prior distributions to x^ε . *Kernel methods* can only reproduce the connections within the x^ε rather than the absolute magnitude. This is a tricky part, because x^ε discrepancy is actually coming from the instrument accuracy, which we don't know. If we assign a very small value, it means there is smaller error and high confidence in the observation data we obtained. On the contrary, if we give a big value, we have less confidence in the data source. Obviously, the discrepancy term x^ε will influence the calibration process.

7 Example to get started

Problem statement

The problem involved the 3D simulation of laterally loaded piles in overconsolidated glacial clay till at the Cowden test site. All numerical analyses were conducted using the ICFEP, a deterministic forward simulator. The model's accuracy was verified, with the soil strength profile s_u and staged pile outputs y as illustrated in Figure 5. The FE pile model employed in this study conforms to the CL2 type outlined, featuring a diameter of 2.0 meters and a length ratio of $L/D = 5.25$. Given the symmetrical nature of the model, only half of the pile is modelled.

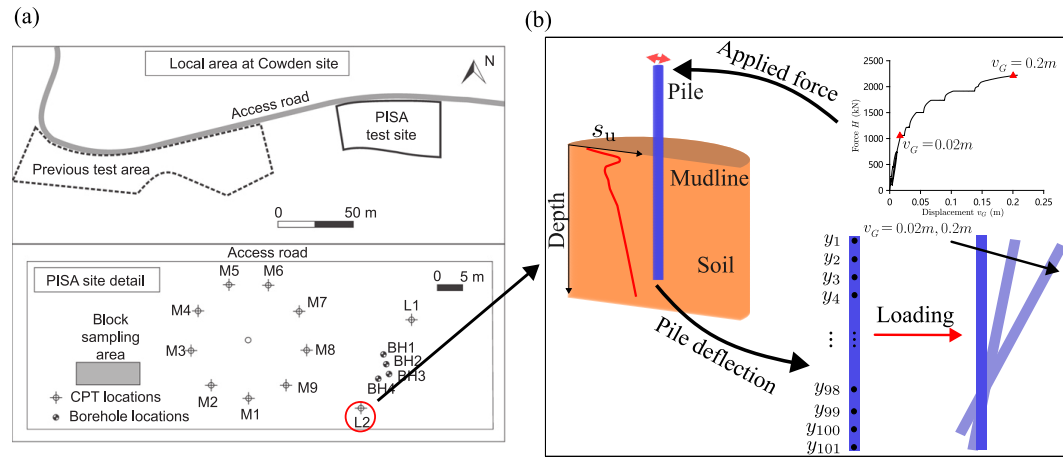


Figure 5: Loading pile at Cowden site

Similar to the original investigation, laboratory data was interpreted to constructed the prior set of best-estimate parameters. This study however con-

siders a number of parameter of the forward model as a random variables sampled from the input domain $\mathcal{D}_{x\mathcal{M}}$ through a distribution defined by the prior $\pi(x)$. Monotonic lateral loading was imposed at the pile head as shown in Figure 5. Two separate measurements of deflection along the length of the pile were collected: (1) pile deflection and (2) reaction force near mudline.

- Two sets of output fields: deflection and force
- Three parameters to be validated: K_0 , G_0 and OCR
- Nine loading stages in real life/70 loading stages in FE modelling
- deflection output is a vector and force output is a scalar

Data preparation

If this is a synthetic problem, a 'GroundTruth.csv' should be provided:

```
... \RUN\input\csv\ICFEP\Pisa_pile\GroundTruth.csv
```

Two output fields observations : (1) pile deflection and (2) reaction force and their coordination

```
... \RUN\input\csv\ICFEP\Pisa_pile\Observation1.csv
```

```
... \RUN\input\csv\ICFEP\Pisa_pile\Observation2.csv
```

Priors for forward parameters: K_0 , G_0 and OCR

```
... \RUN\input\csv\ICFEP\Pisa_pile\Prior_forward.csv
```

Priors for discrepancy hyperparameters: w and h

```
... \RUN\input\csv\ICFEP\Pisa_pile\hyper_obs_discrepancy1.csv
```

```
... \RUN\input\csv\ICFEP\Pisa_pile\hyper_obs_discrepancy2.csv
```

Experimental design range: for surrogate construction

```
... \RUN\input\csv\ICFEP\Pisa_pile\Prior_EoD.csv
```

Loading stage time: FE stage time/observation time (try to make these at same coordination)

```
... \RUN\input\csv\ICFEP\Pisa_pile\StageTime.csv
```

```
... \RUN\output\mat\EoD_FE_summary\xxx.mat\FE_stagetime
```

FE model evaluation: LHS input/output experimental design realization and its coordination

```
... \RUN\output\mat\EoD_FE_summary\xxx.mat\FEcell_output  
... \RUN\output\mat\EoD_FE_summary\xxx.mat\EDarray  
... \RUN\output\mat\EoD_FE_summary\xxx.mat\coord
```

Code setting

Working path must be in 'xxx/Run' file

```
... \SBI\RUN
```

Add subroutine path and UQlab core path

```
addpath("... \UQlab\UQLab_Rel2.1.0\UQLab_Rel2.1.0\core"); %UQLAB core  
addpath("... \SBI\Main\subroutines"); %subroutine
```

Add noise to the observation data

```
Noise.scalar = 0; % Value as a percentage (e.g. 0.01-1.0)
```

Input the FE loading stages number

```
AnParam.N_FE_stage = 70; %number of FE loading stages
```

Switches to control whether save/show the predictive distributions for the output

```
AnParam.Switch_Predictive = "off";  
% switch for prior/posterior predictive for real outputs
```

Validation runs setting: Keep some runs for validation

```
AnParam.TestDataRun = 10; %Test FE runs number
```

Switches to surrogate plotting figures to see the fitting degree

```
AnParam.Switch_surrogatePlot = "on";
```

Switches to corner figure plotting

```
AnParam.Switch_corner_scatter = "on";
```

Switches to Sobol figure plotting

```
AnParam.Switch_Sobol = "on";
```

Switches to threeSigma figure plotting

```
AnParam.Switch_threeSigma = "on";
```

Switches to show figures in matlab live script or save outside the matlab

```
AnParam.figureDisplay = "Show-in-mlx";
```

All other switches not mentioned dont suggest using

Code running