Friday 9:00-10:30

Part 12

Hands-on examples of imprecise simulation in engineering (continued)

by Edoardo Patelli and Jonathan Sadeghi

Hands-on examples of imprecise simulation in engineering (continued)

Outline

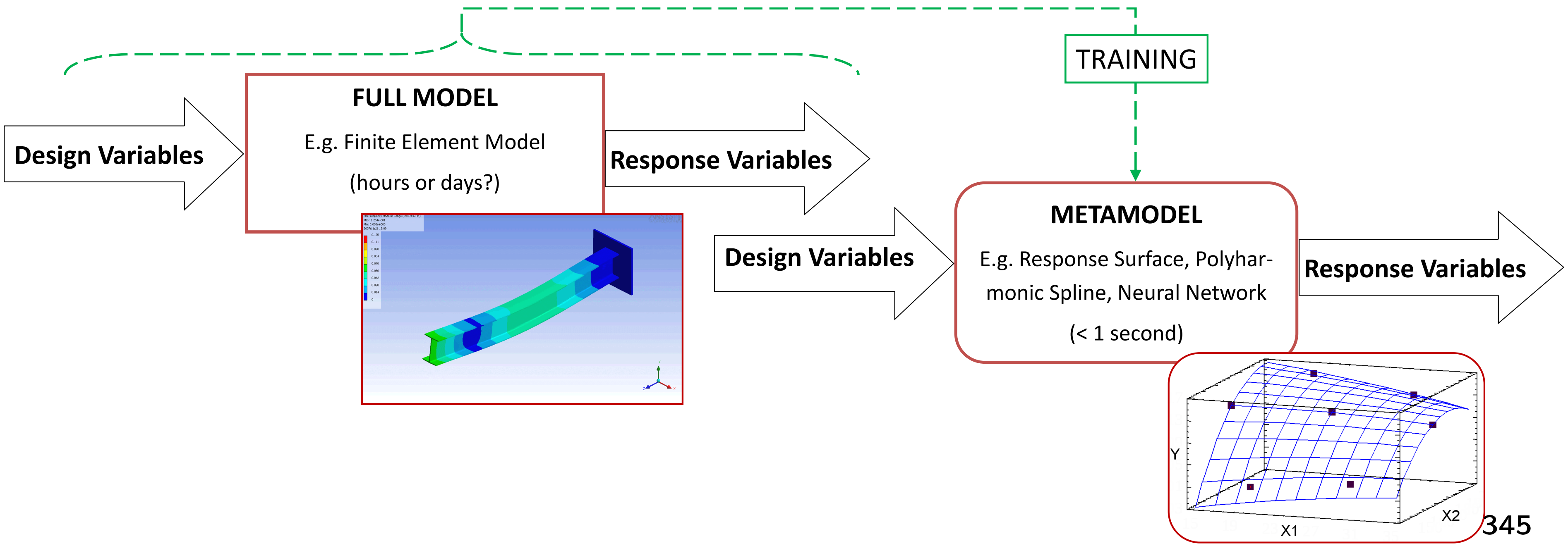
Hands-on examples of imprecise simulation in engineering (continued)

Outline

- Theory
  - Metamodels
  - Interval Predictor Models
  - Random Predictor Models
- Applications
  - History Matching
  - Simple Function
  - IC Fault Model
- Example

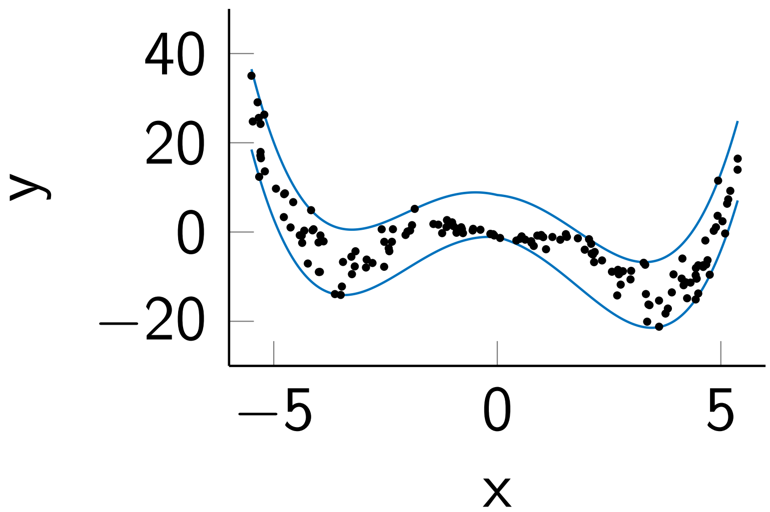
Metamodels

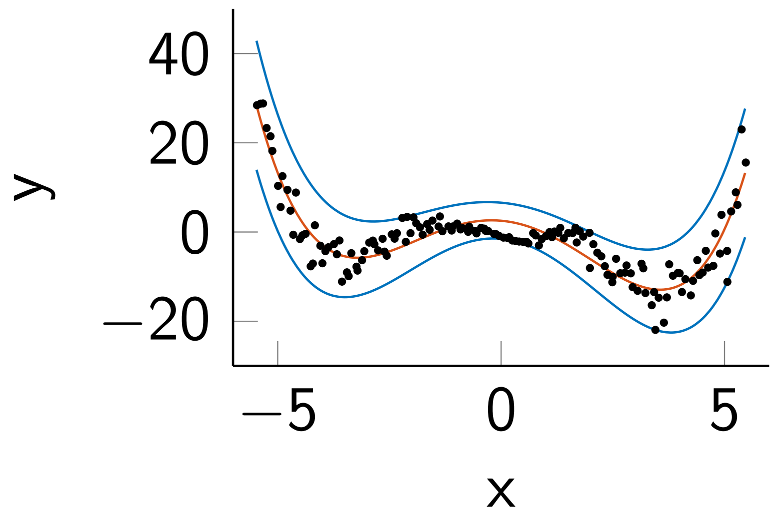
- If the full model is too computationally expensive to do many simulations, or we have simulation results (or real data!) already available we can replace the full model with an approximation:
- Response Surfaces, Polyharmonic Spline, Neural Networks...
- Interval Predictor Models and Random Predictor Models.
- A good approximation should fit existing data well and generalise well to new data



Interval/Random Predictor Model

- IPMs and RPMs are new types of metamodel with favourable properties for dealing with scarce/limited data.
- The variance in the data can be robustly estimated without making unjustified assumptions (distribution of noise, for example).
- The reliability of the metamodel can be bounded (more on this later).





Hands-on examples of imprecise simulation in engineering (continued)

Outline

Interval Predictor Models - Mathematics

- An IPM is defined as a function returning an interval for each vector $x \in X$
- i.e.

$$I_y(x, P) = \{y = M(x, p), p \in P\} \quad (52)$$

- Crespo (2016) considers for example:

$$I_y(x, P) = \left\{y = p^T \phi(x), p \in P\right\} \quad (53)$$

- p is a member of the hyper-rectangular uncertainty set:

$$P = \left\{p : \underline{p} \leq p \leq \bar{p}\right\} \quad (54)$$

- IPM

$$I_y(x, P) = [\underline{y}(x, \bar{p}, \underline{p}), \bar{y}(x, \bar{p}, \underline{p})] \quad (55)$$

How to train a type 1 IPM

- $$\underline{y}(x, \bar{p}, \underline{p}) = \bar{p}^T \left(\frac{\phi(x) - |\phi(x)|}{2} \right) + \underline{p}^T \left(\frac{\phi(x) + |\phi(x)|}{2} \right) \tag{56}$$
- $$\bar{y}(x, \bar{p}, \underline{p}) = \bar{p}^T \left(\frac{\phi(x) + |\phi(x)|}{2} \right) + \underline{p}^T \left(\frac{\phi(x) - |\phi(x)|}{2} \right) \tag{57}$$
- Can use polynomial or radial basis
- To find a good model attempt to minimise (expected value of):
  $$\delta_y(x, \bar{p}, \underline{p}) = (\bar{p} - \underline{p})^T |\phi(x)| \tag{58}$$
  with the constraints that all data points to be fitted lie within these bounds and that the upper bound is greater than the lower bound
- i.e. we solve a linear optimisation program
- These constraints give a type 1 IPM

Outliers

- Two criterion are used to find outliers:
- We can find a CDF for the distance of each p from the centre of the uncertainty set and then identify a fraction λ_p of points which prevent the interval being further minimised
- We can find the fraction λ_e of points with the furthest squared distances from the LS fit
- Points satisfying both criterion can be disregarded as outliers - then we can retrain with the new subset of points
- The analyst must make a sensible choice of λ_p and λ_e

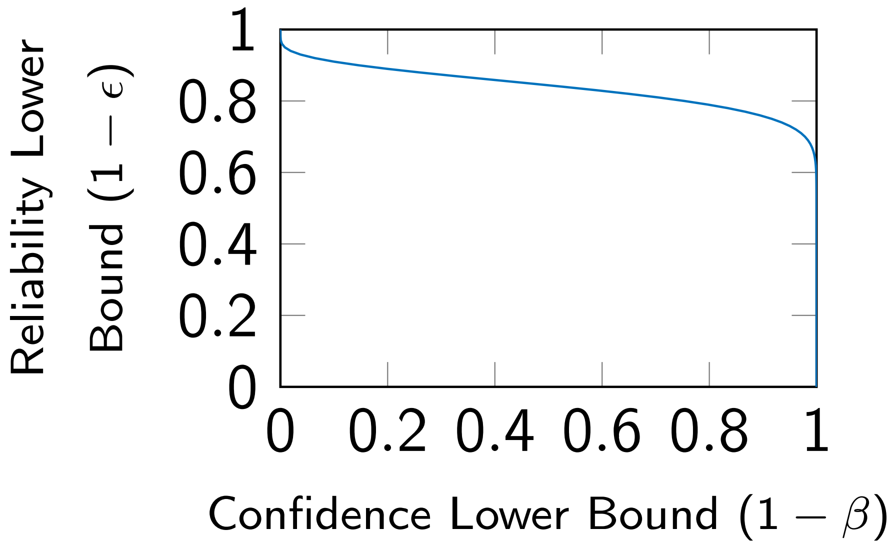
Reliability

- For reliability parameter ϵ and confidence parameter β satisfying

$$\binom{k+d-1}{k} \sum_{i=0}^{k+d-1} \binom{N}{i} \epsilon^{i} (1-\epsilon)^{N-i} \leq \beta, \qquad (59)$$

- the confidence and reliability parameters of the IPM are bounded by

$$Prob_{P^n}[R \geq 1-\epsilon] \geq 1-\beta. \qquad (60)$$



Hands-on examples of imprecise simulation in engineering (continued)

Outline

Random Predictor Models

- A function returning a random variable for each vector $x \in X$ Crespo (2015) considers for example:

$$R_y(x, P) = \left\{ y = p^T \phi(x), p : F_p(p), p \in P \right\} \tag{61}$$

- it can be shown that:

$$\underline{p} \leq \mu \leq \bar{p} \qquad 0 \leq \nu \leq (\mu - \underline{p}) \odot (\bar{p} - \mu) \qquad -1 \leq c \leq 1 \tag{62}$$

$$C(\nu, c) \succeq 0 \tag{63}$$

- σ surface connects all outputs τ standard deviations from μ

$$I_\sigma(x, \mu, -\tau, \nu) = [I(x, \mu, -\tau, \nu, c), I(x, \mu, \tau, \nu, c)] \tag{64}$$

-

$$\nu_y(x, \nu, c) = \phi(x) C(\nu, c) \phi(x) \tag{65}$$

$$\mu_y(x, \mu) = \mu^T \phi(x) \tag{66}$$

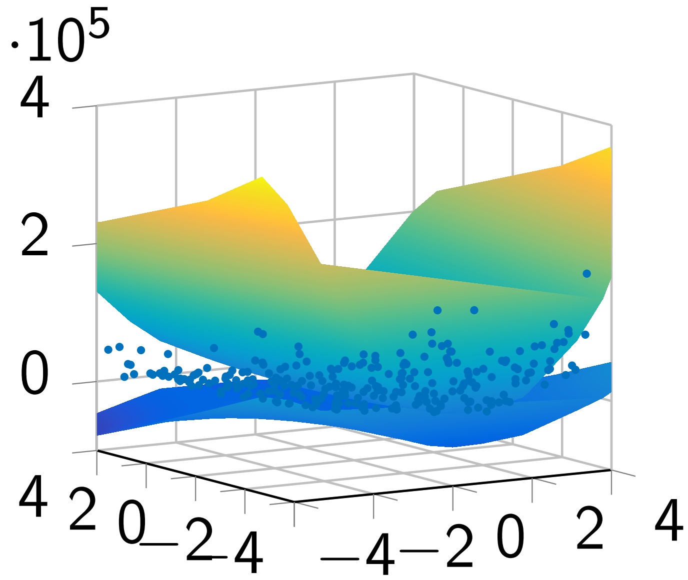
Type 1 RPM - Optimisation program

- $$I(x, \mu, \tau, \nu, c) = \mu^T \phi(x) + \tau \sqrt{\nu_y(x, \nu, c)} \tag{67}$$
- μ is found by any means - least squares is commonly used.
- $$\hat{\nu} = \underset{\nu \geq 0}{\text{argmin}} \left\{ \mathbf{E}[\nu_y(x, \mu)] : I(x_i, \mu, -\sigma_{max}, \nu, c) \leq y_i \leq I(x_i, \mu, \sigma_{max}, \nu, c) \text{ for } i = 1, ..., N \right\} \tag{68}$$
- σ_{max} is chosen by analyst to decide number of standard deviations from mean containing all data points.
- Reliability assessment from IPM applies to $I_\sigma = [I(x_i, \mu, -\sigma_{max}, \nu, c), I(x_i, \mu, \sigma_{max}, \nu, c)]$ also.
- Similar outlier removal algorithm possible (distance from mean, normalised by variance).
- We can also use Type 2 RPMs (chance constrained formulation where constraint violation is allowed).

Implementation

- Implemented a class to construct IPMs/RPMs in generalized uncertainty quantification software OpenCOSSAN
- Training, Reliability evaluation, Outlier removal are all performed automatically in OOP framework, with choice of optimisers/basis type/additional constraints and more

COSSAN



Hands-on examples of imprecise simulation in engineering (continued)

Outline

Hands-on examples of imprecise simulation in engineering (continued)

Outline

What is History Matching?

- A type of model calibration
- If we have some real data and a model with some free parameters which we wish to tune to reproduce the data
- Many methods
- Bayesian Inversion is popular
- See Tarantola, Inverse Problem Theory or Carter, J. N. "Using Bayesian statistics to capture the effects of modelling errors in inverse problems."
- Usually use least squares objective function between data and model output - and a clever optimisation algorithm!

Hands-on examples of imprecise simulation in engineering (continued)

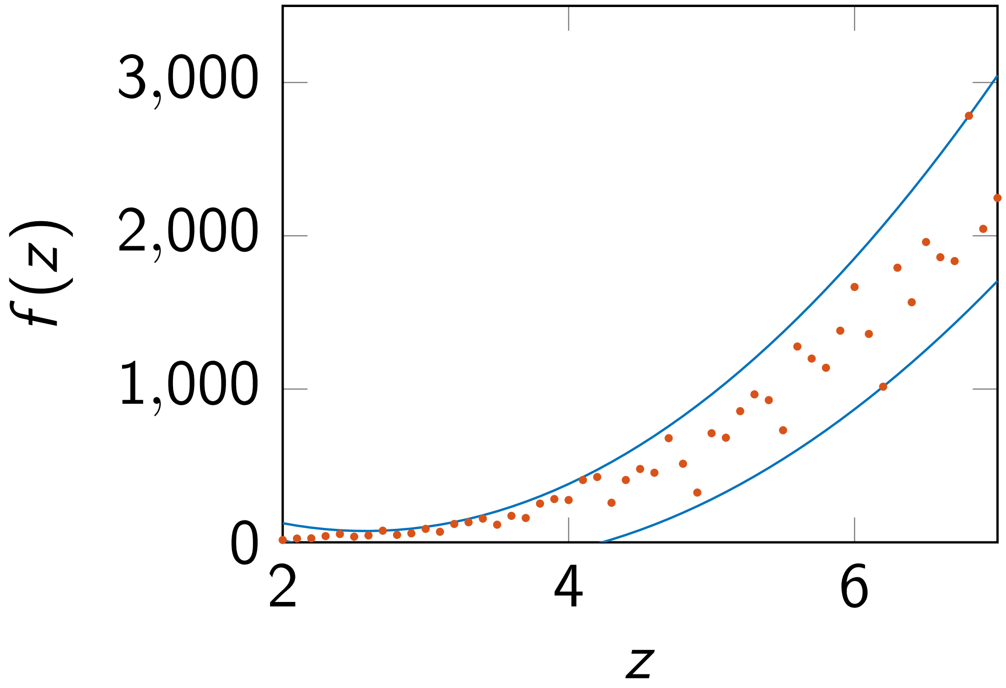
Outline

Simple Example

- As in Carter (2004), the following function will be taken as a black box

$$f(z) = (z^2 + 0.1z)^2 + \eta_1, \qquad (69)$$

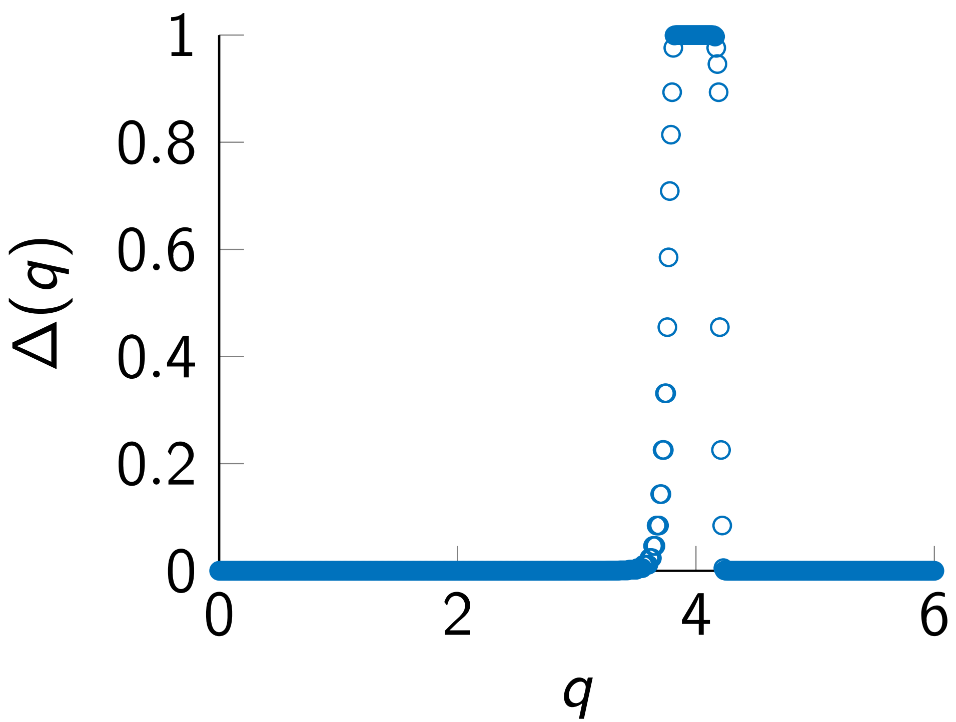
- Data provided is for $z = 2$ to $z = 7$ - challenge is to predict $z = 10$
- The 'model' we have to match is $g(q, z) = z^q$



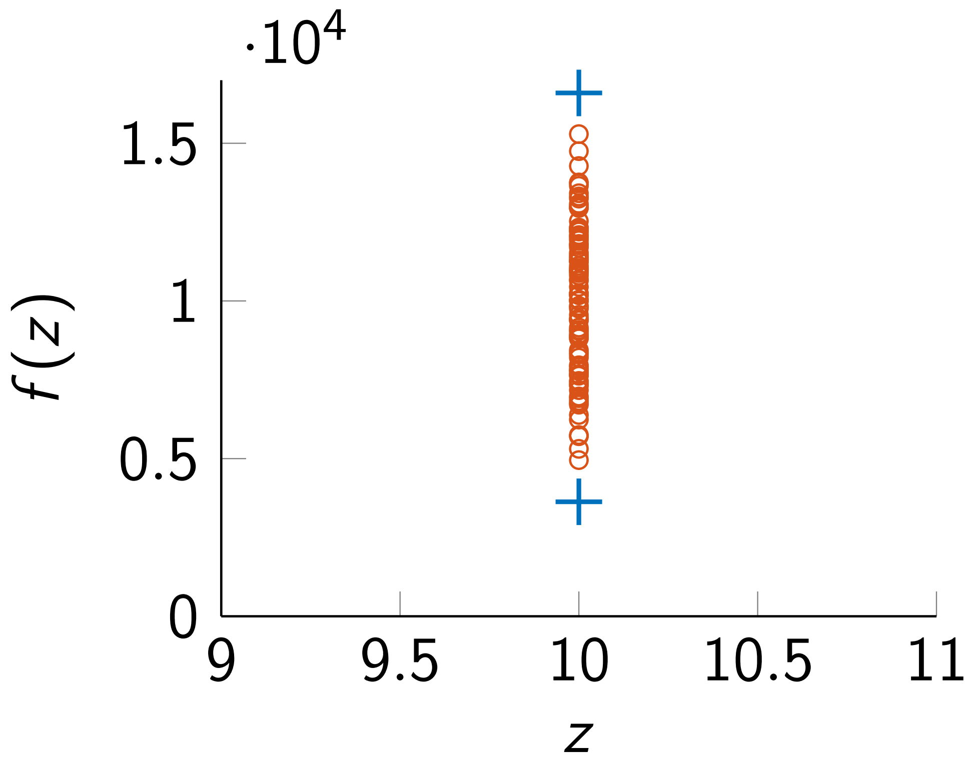
- As you can see I fitted an IPM to the data. New objective function for simulations:

$$\Delta(q) = \sum_{i=0}^{C(q)} \binom{D}{i} {R^*}^i (1 - R^*)^{D-i}, \tag{70}$$

- Then find feasible q:



- Which enables us to make predictions...

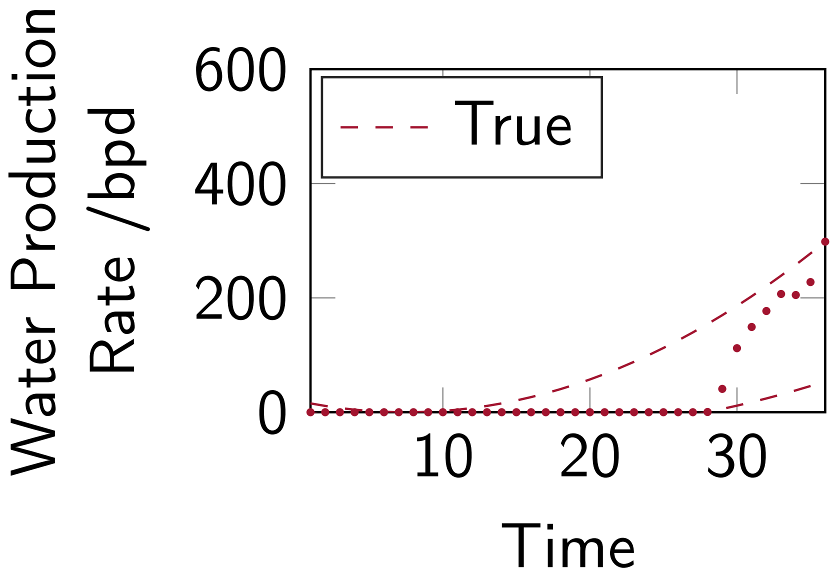


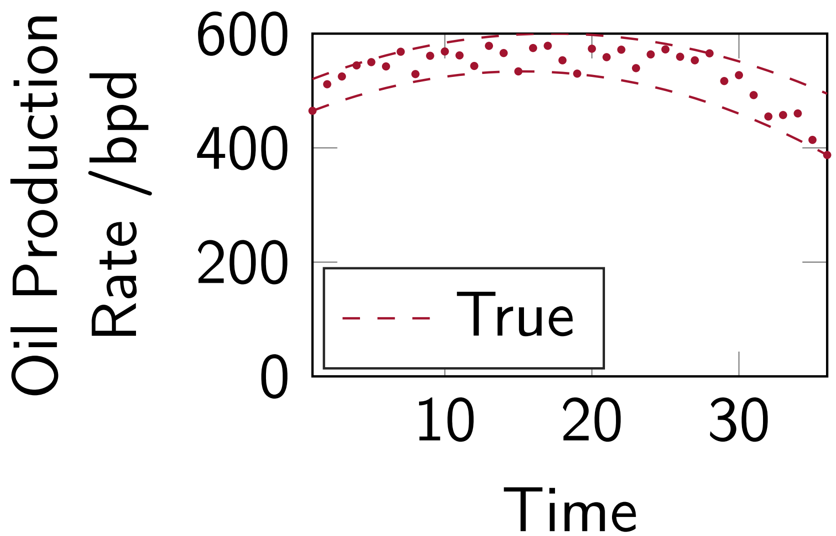
Hands-on examples of imprecise simulation in engineering (continued)

Outline

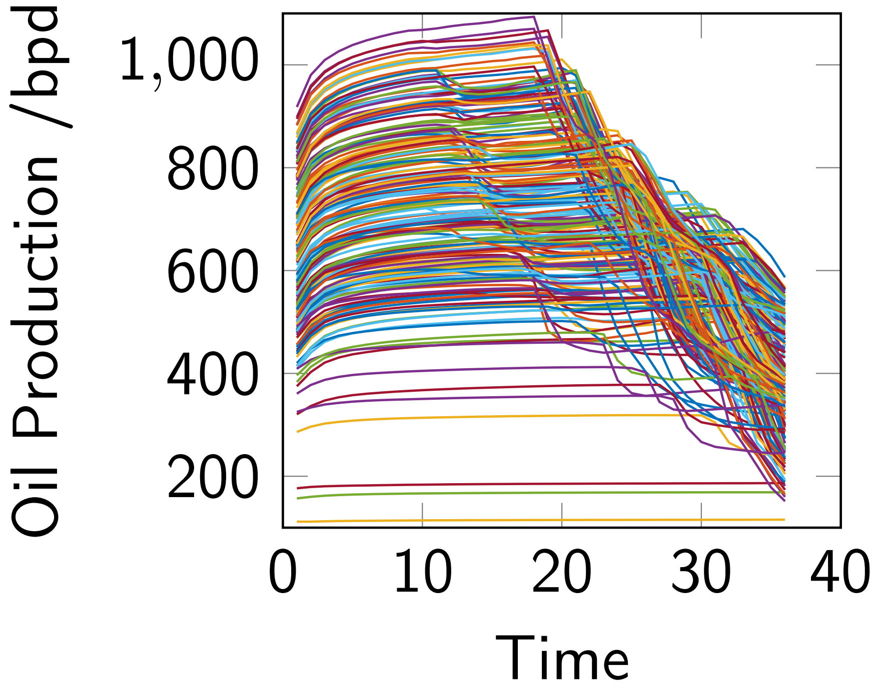
Imperial College Fault Model

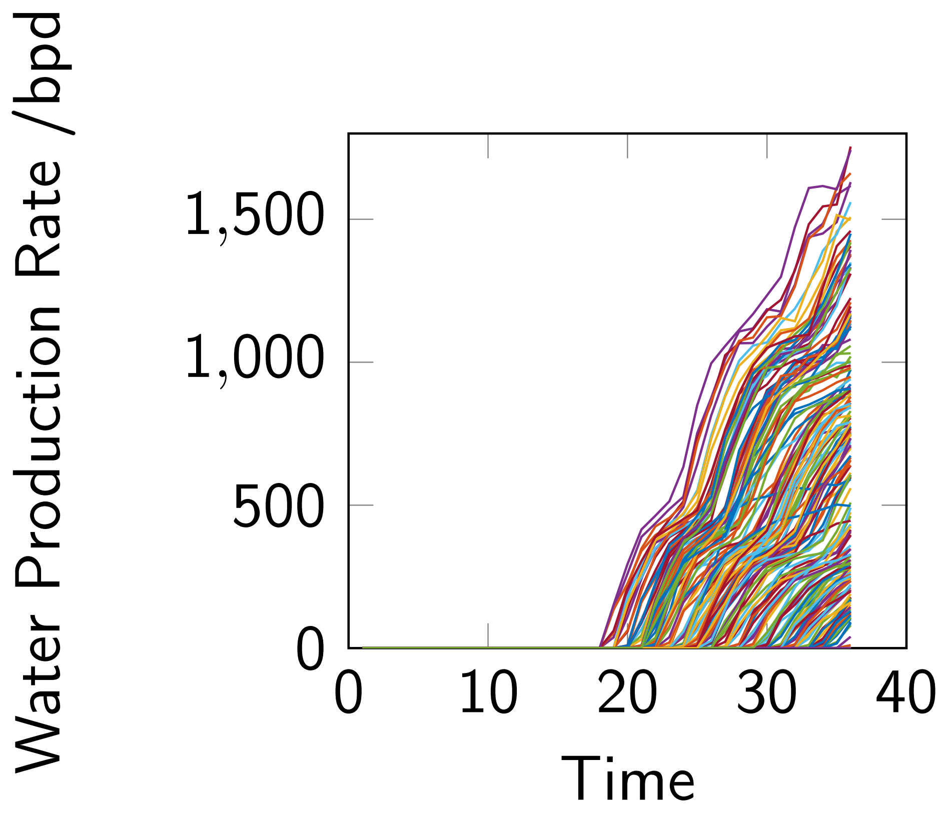
- Model of a reservoir which has been producing oil for 36 months and has now started producing water ('true' data was produced using a hetrogenous model with added noise (3%)).
- The challenge is to predict future production using a finite element model (homogenous)
- Good and low quality sand permeabilities and fault throw are unknown - to be determined by matching history data with the true data.
- Database with $\sim$ 160000 simulation results available online





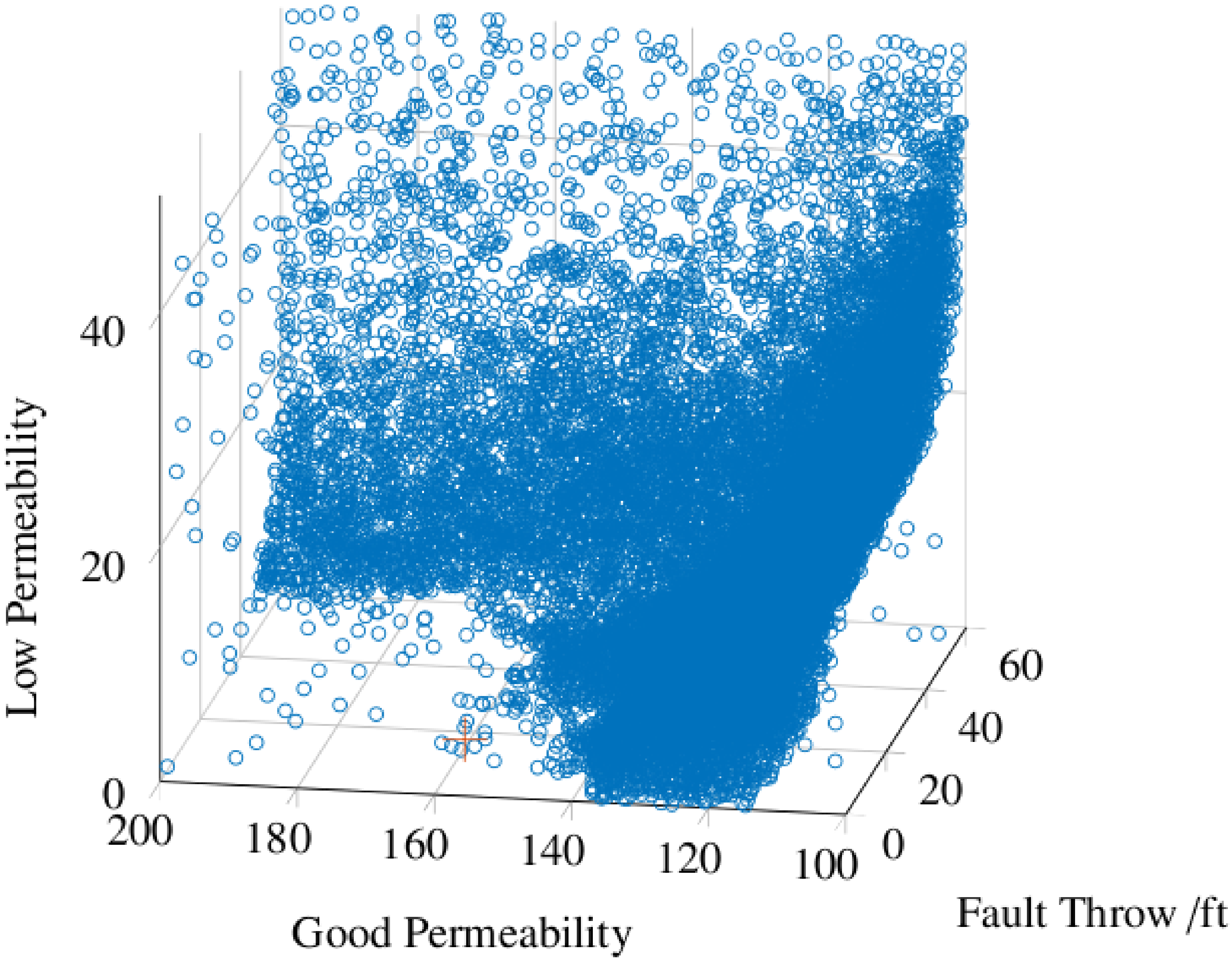
Simulations





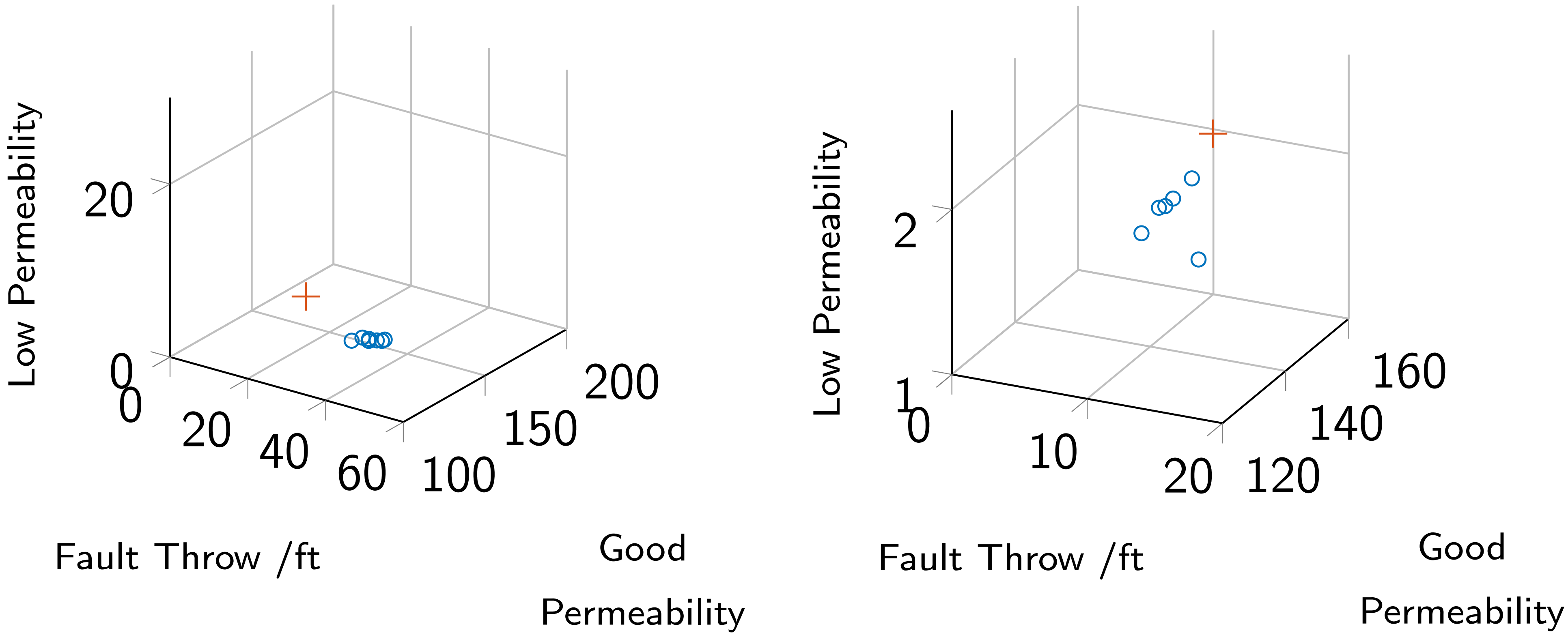
IC Fault

- Look for solutions with $\Delta(m) > 0.01$



Results

- Simulations close to minima of the objective function:

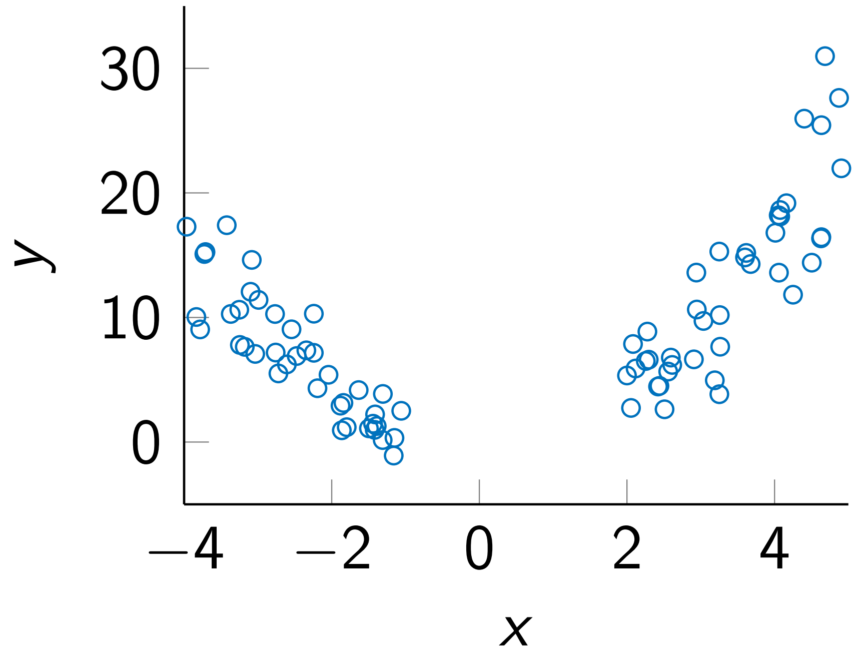


Hands-on examples of imprecise simulation in engineering (continued)

Outline

An example for you to try

- Please refer to your handouts
- Your friend at the University requires help with some data analysis.
- Use the programming language you prefer. I have provided instructions on a numerical method. I have prepared a solution in Matlab, and hence have provided some Matlab hints.
- Please give an interval for the value of y at $x = 1$ with a probability bound.



Questions?

- Thank you.
- Jonathan Sadeghi
- J.C.Sadeghi@liverpool.ac.uk