Emulation of the Availability of an Offshore Windfarm

Jack Kennedy, Kevin Wilson, Daniel Henderson

Newcastle University

j.c.kennedy1@ncl.ac.uk

5/3/19

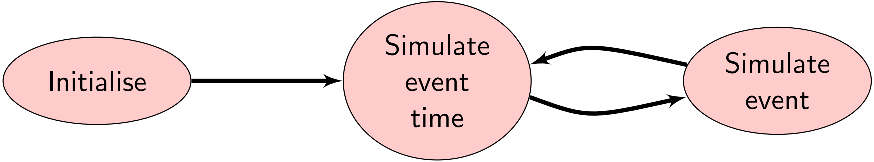
Introduction

- Mathematical modelling used across all of science
- We use simulators to perform "computer experiments" - vital in (risky!) engineering contexts
- Sophisticated simulators often computationally intensive
- **Example**: stochastic windfarm simulation, used to model large, offshore windfarms [1]
  - Walney Extension - 87 turbines / 659 MW capacity
  - London Array - 175 turbines / 630 MW capacity

[1] Zitrou et al. [2013]

Simulator Description

- Simulator developed as decision support tool to understand how to get the most out of a windfarm
- Windfarm represented by complex Monte Carlo simulation

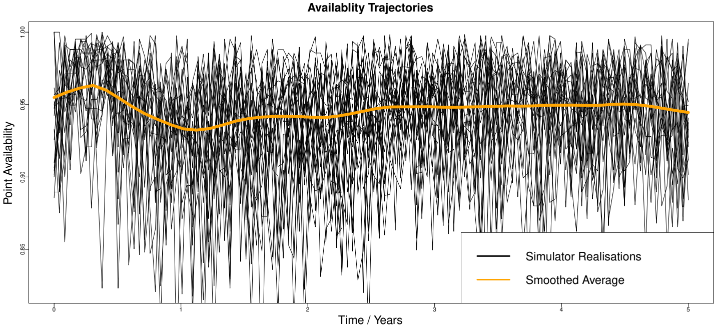


- Inputs: Hundreds - some known and some prone to uncertainty

Fixed, Known	Fixed, Unknown
Windfarm topology	Cable failure rates
Simulation length	Initial hazard function parameters
	Learning Rate

Simulator Description

- Output: Availability trajectories over time.
- Point Availability Informed Capacity: $C(t;\boldsymbol{x}) = \frac{\sum_{i=1}^{n} OP_i(t;\boldsymbol{x})}{\sum_{i=1}^{n} IP_i(t;\boldsymbol{x})}$



- Average over time period gives "fixed term availability" or $A(\boldsymbol{x})$.
- A profitable windfarm satisfies $A(\boldsymbol{x}) > 97\%$

Simulator Description

Simulator prone to uncertainty

- **Epistemic uncertainty**: simulator inputs are unknown
  $\rightarrow$ quantify uncertainty and propagate
- **Aleatory uncertainty**: simulator is stochastic
  $\rightarrow$ investigate with many simulations

No. Turbines	1 Run (s)	10^6 Runs (years)
9	15	0.5
200	120	3.8

- Performing an uncertainty analysis is seemingly impossible!

Emulators

However ... there is a solution!

- Utilise Gaussian Process (GP) emulators [2] facilitate computation
- An emulator is a statistical (Bayesian) **surrogate** model of the simulator or **“model of the model”**.
- Crucially: emulators are incredibly cheap to evaluate

[2] Kennedy and O'Hagan [2001]

Emulators

Basic idea of emulation:

- Simulator viewed as an unknown function $y = \eta(\boldsymbol{x})$

Prior specification:

$$\eta(\cdot) \sim \mathcal{GP}\{m_0(\cdot), C_0(\cdot,\cdot)\}$$

$$C_0(\boldsymbol{x}_i, \boldsymbol{x}_j) = \sigma^2 \exp\left\{ -\sum_{k=1}^{K} \left(\frac{x_i^k - x_j^k}{\theta_k}\right)^2 \right\} + \lambda^2 \delta_{ij}$$

Emulators

- We observe training data $(\boldsymbol{Y}_t, \boldsymbol{x}_t)$; want to infer $\boldsymbol{Y}_p = \eta(\boldsymbol{x}_p)$

$$\begin{pmatrix} \boldsymbol{Y}_t \\ \boldsymbol{Y}_p \end{pmatrix} \sim \mathcal{N}\left\{ \begin{pmatrix} \boldsymbol{m}_0(x_t) \\ \boldsymbol{m}_0(x_p) \end{pmatrix}, \begin{pmatrix} \Sigma_{TT} & \Sigma_{TP} \\ \Sigma_{PT} & \Sigma_{PP} \end{pmatrix} \right\}$$

Then

$$\boldsymbol{Y}_p | \boldsymbol{Y}_t = \boldsymbol{y}_t \sim \mathcal{N}\{\boldsymbol{m}^*(\boldsymbol{x}), \Sigma^*\}$$

Where

$$\boldsymbol{m}^*(\boldsymbol{x}) = \boldsymbol{m}_0(\boldsymbol{x}_p) + \Sigma_{PT}\Sigma_{TT}^{-1}(\boldsymbol{y}_t - \boldsymbol{m}_0(\boldsymbol{x}_t))$$

$$\Sigma^* = \Sigma_{PP} - \Sigma_{PT}\Sigma_{TT}^{-1}\Sigma_{TP}$$

Fitting the Emulator

- Since think of the simulator as a computer experiment $\rightarrow$ need to think about the design of this experiment
- Experimental Design: Latin Hypercube over 6 inputs, 50 datapoints:

Farm Characteristics	Initial Hazard Function Parameters
Learning rate	Generator Wearout Onset
Cable Failure rate	Gearbox Wearout Onset
Cable repair rate	Frequency Converter Wearout Onset

- Experiment result: 50000 realisations of $A(\boldsymbol{x})$; 1000 for each $\boldsymbol{x}$
- Will emulate the mean and variance of $A(\boldsymbol{x})$

Fitting the Emulator

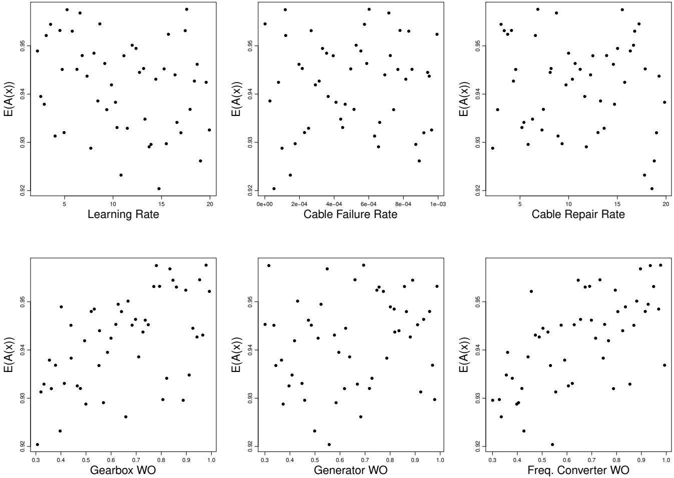


Figure 1: Mean of $A(\boldsymbol{x})$ against inputs

Fitting the Emulator

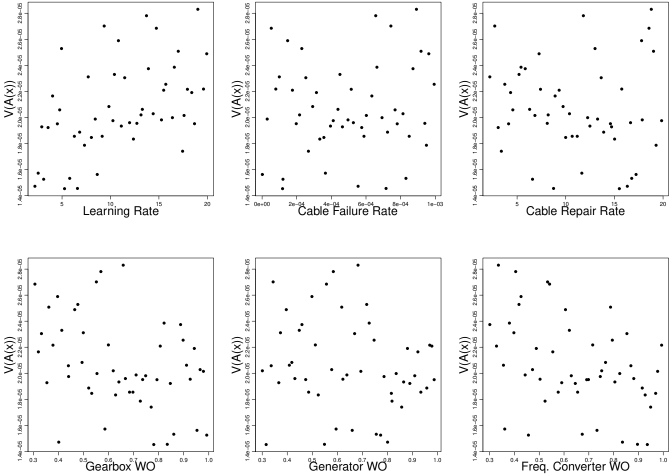


Figure 2: Variance of $A(\boldsymbol{x})$ against inputs

Fitting the Emulator

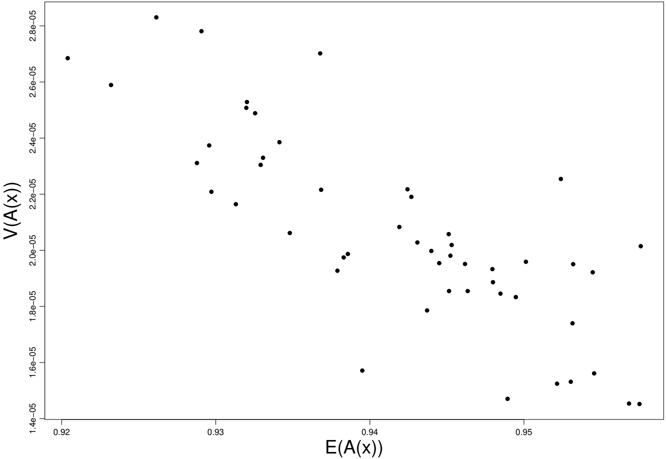


Figure 3: Variance of $A(\boldsymbol{x})$ against the mean of $A(\boldsymbol{x})$

Fitting the Emulator

We can emulate the windfarm simulator with a 4 stage approach

1. Generate training data: $(\boldsymbol{y}_t, \boldsymbol{x}_t)$; compute means and variances
2. Construct emulator for $\Phi\big(\mathrm{E}\{A(\boldsymbol{x})\}\big)$
3. Construct emulator for $\log\big(\mathrm{Var}\{A(\boldsymbol{x})\}\big)|\mathrm{E}\{A(\boldsymbol{x})\}$
4. Reintroduce the stochasticity: $A(\boldsymbol{x}) \sim Beta\big\{a(\boldsymbol{x}), b(\boldsymbol{x})\big\}$

Result: Super cheap way to obtain realisations of windfarm's fixed term availability

Bonus step : Test emulator predictions against **independently generated** validation data

Emulator Validation

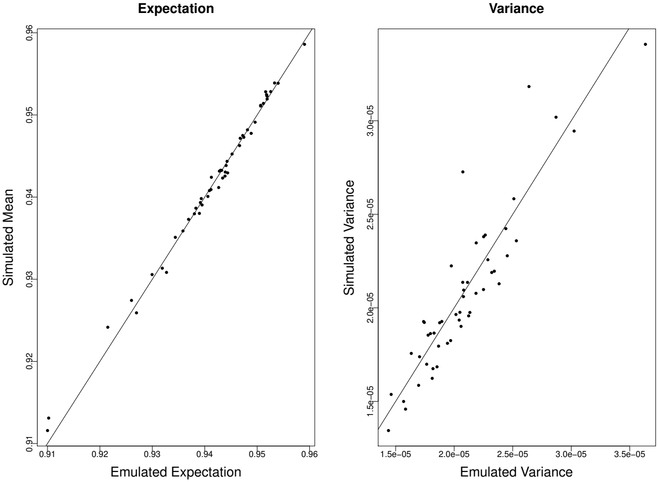


Figure 4: Observed vs predicted values of simulator mean and variance on independently generated validation data

Emulator Validation

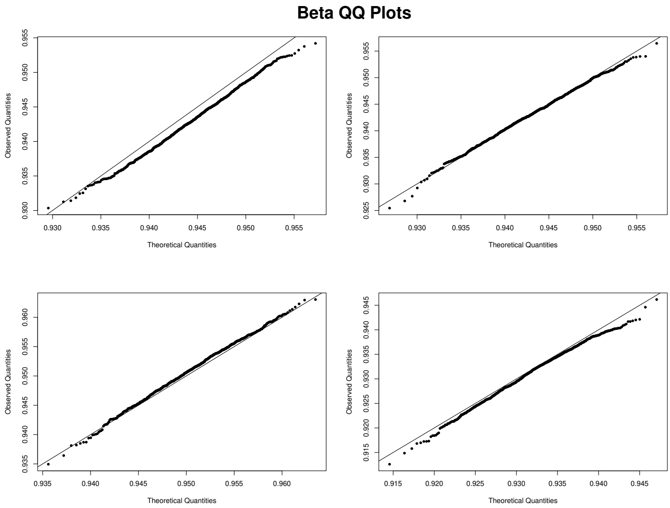


Figure 5: Beta QQ plots based on emulated mean and variance of validation data

Conclusions and Further Work

- Emulator takes ~ 0.5s to obtain mean and variance of $A(\boldsymbol{x})$ compared to simulator taking $\sim 4 - 5$ hours.
- Emulator still very accurate.
- Want to perform uncertainty quantification on input parameters to assess (say) $\mathrm{P}(A > 97\%)$
- In performing the UQ, we need to identify important inputs and elicit uncertainty over these inputs

Bibliography

MC Kennedy and A O'Hagan. Bayesian calibration of computer models. *Journal Of The Royal Statistical Society Series B-Statistical Methodology*, 63:425–450, 2001. ISSN 1369-7412.

Athena Zitrou, Tim Bedford, Lesley Walls, Kevin Wilson, and Keith Bell. Availability growth and state-of-knowledge uncertainty simulation for offshore wind farms. In *22nd ESREL conference 2013*, September 2013. URL `https://strathprints.strath.ac.uk/45377/`.