Comparison of Robust Artificial Neural Network and Gaussian Process Model for Interval Regression

Bright Uchenna Oparaji & Edoardo Patelli

Institute for Risk and Uncertainty

UNIVERSITY OF LIVERPOOL

REC2018

Context

Context

Context

UNIVERSITY OF
LIVERPOOL

Context

Context

Real World

Represent

Mathematical World

Mathematical Models

As time passes

Use model to predict future, verify the past

Compare

Observe Future data

Predict Behaviour

UNIVERSITY OF LIVERPOOL

Context

The **big problem** is that the model are expensive to run !!

- Model Complexity.
- The governing equations used to describe the system.
- Computing rare events that exist at the tail of the distribution.

Problem Statement

Surrogate Model

High Fidelity Model

$$f(x)$$

Fast-running Surrogate

$$y = \mu_y(x) + \varepsilon(x)$$

Problem Statement

Surrogate Model

High Fidelity Model

$$f(x)$$

Fast-running surrogate

$$y = \mu_y(x) + \varepsilon(x)$$

Surrogate output

Problem Statement

Surrogate Model

Simulator (Expensive Code)

$$f(x)$$

Fast-running surrogate

$$y = \mu_y(x) + \varepsilon(x)$$

Surrogate output

Mean of surrogate

Problem Statement

Surrogate Model

Simulator (Expensive Code)

$$f(x)$$

Emulator (fast running)

$$y = \mu_y(x) + \varepsilon(x)$$

Simulator Output

Emulator

Noise

Artificial Neural Network as a Surrogate Model

What do we want from an ANN?

Artificial Neural Network as a Surrogate Model

What do we want from an ANN?

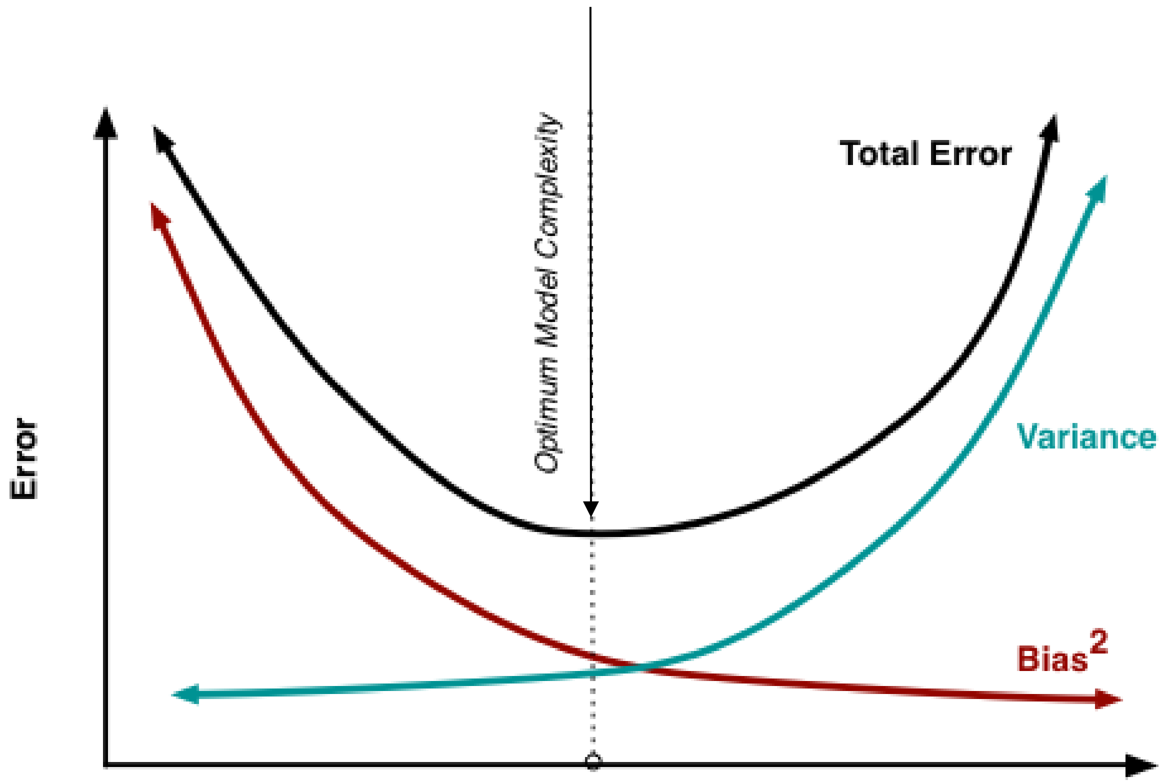
Robustness!!!

Artificial Neural Network as a Surrogate Model

What do we want from an ANN?

Robustness!!!

Low bias and variance



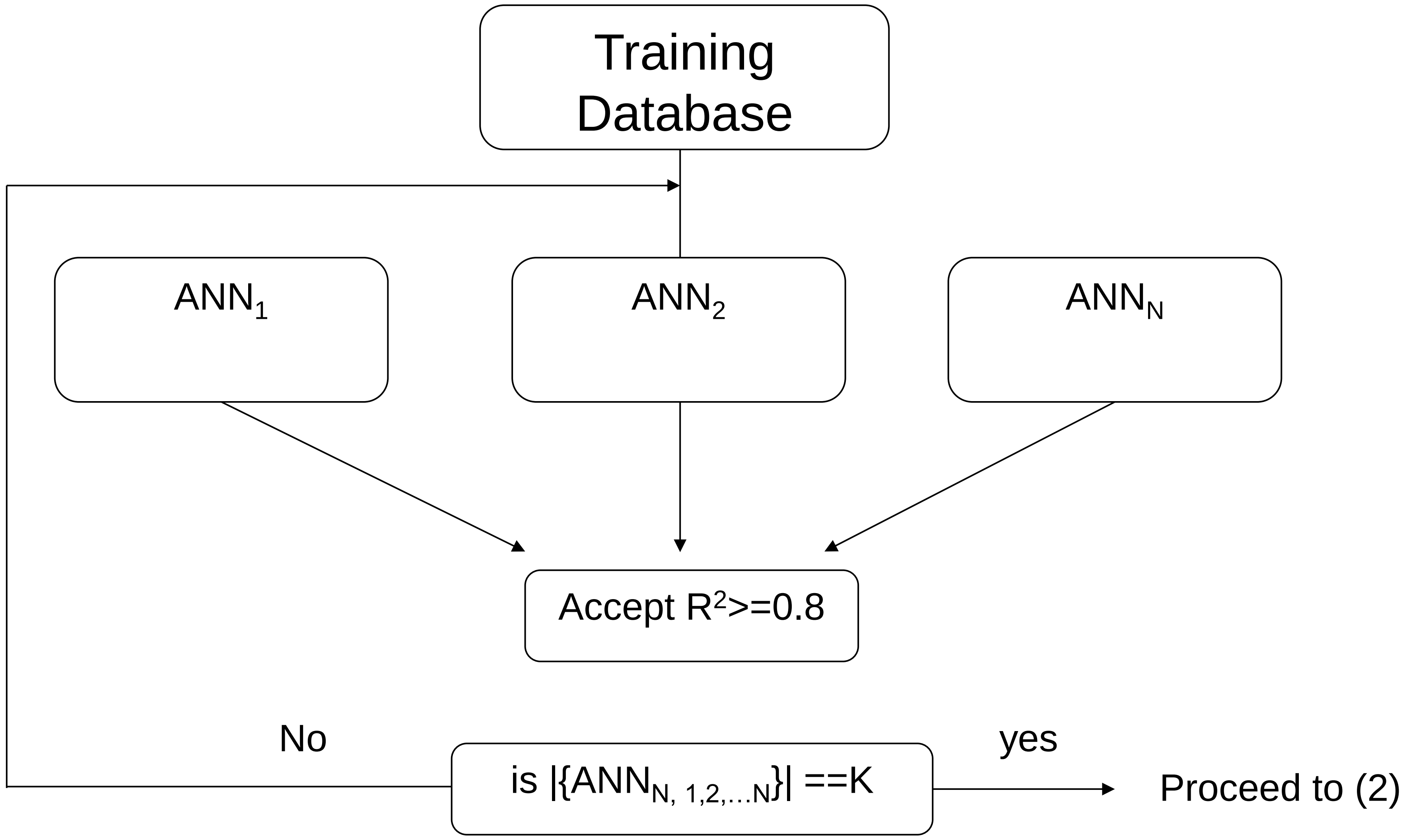
Artificial Neural Network as a Surrogate Model

The big problem are the **uncertainties** introduces on top of the surrogate model

- Random model parameters.
  - Train a large set of models,
  - Select the best model based on a performance metric (i.e. R^2).
  - Noise in data could falsify results.

Proposed Method

(1) Construction of a set of K Identical Networks



UNIVERSITY OF LIVERPOOL

Proposed Method

(2) Rank the ANNs in the set

$$P(N_k|D_{train}) = \frac{P(D_{train}\{x,y\}|N_k)P(N_k)}{\sum_{q=1}^{M} P\left(D_{train}\{x,y\}\middle|N_q\right)P(N_q)}$$

$P(D_{train}\{x,y\}|N_k)$ ← Likelihood of the *k^th^* network in the set

$P(N_k)$ ← Prior Probability of *k^th^* network in the set

$P(N_k|D_{train})$ ← Posterior Probability of *k^th^* network in the set

- Due to the difficulty of estimating $P(N_k)$ uniform probability *1/M* is assigned

$$\sigma_k^2 = \frac{1}{N}\sum_{i=1}^{N} \varepsilon_{ki}^2 \qquad \varepsilon_{ki} = D_{train}\{y_i\} - \hat{y}_i, \varepsilon_{ki} \sim N\left(0, \sigma_k^2\right), i = 1, \ldots, \mathrm{N}$$

Maximum Likelihood Estimation Approach

$$P(D_{train}|N_k) \approx \frac{1}{\sqrt{2\pi\sigma_k^2}} \frac{1}{N}\sum_{i=1}^{N} exp\left\{\frac{-[y_i - \hat{y}_{ki}]^2}{2\sigma_k^2}\right\}$$

UNIVERSITY OF LIVERPOOL

Proposed Method

(3) Device Robust prediction and quantify of model uncertainties

$$y_{robust} = \hat{y}^* + A_f$$

y_{robust} ⟵ robust prediction (i.e. best prediction considering model uncertainties)

$\hat{y}^*$ ⟵ Prediction from best Neural Network

A_f ⟵ Adjustment factor assumed to be a normally distributed *N(μ, σ)*

$$E[A_f] = \sum_{i=1}^{M} P(N_i|D_{train})(\hat{y}_i - \hat{y}^*)$$

$$Var[A_f] = \sum_{i=1}^{M} P(N_i|D_{train})\,(\hat{y}_i - E[y_{robust}])^2$$

$$E[y_{robust}] = \hat{y}^* + E[A_f]$$

$$Var[y_{robust}] = Var[A_f]$$

Confidence Interval ⟶ $[\overline{y}_{robust}, \quad \underline{y}_{robust}]$

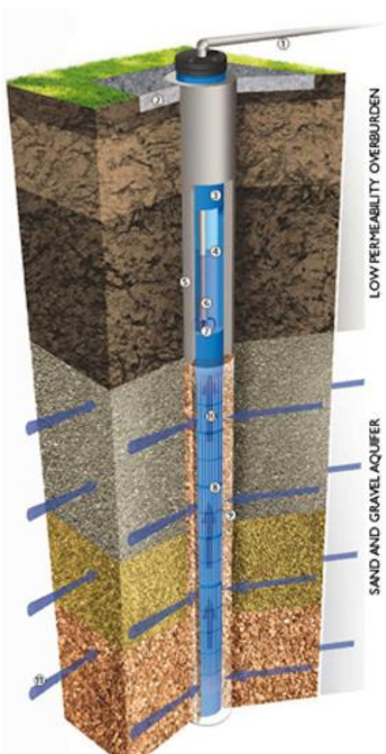
$$\bar{y}_{robust} = E(y_{robust}) + 1.96\sqrt{Var(y_{robust})}$$

$$\underline{y}_{robust} = E(y_{adj}) - 1.96\sqrt{Var(y_{robust})}$$

UNIVERSITY OF LIVERPOOL

Simulation of water through a bore hole

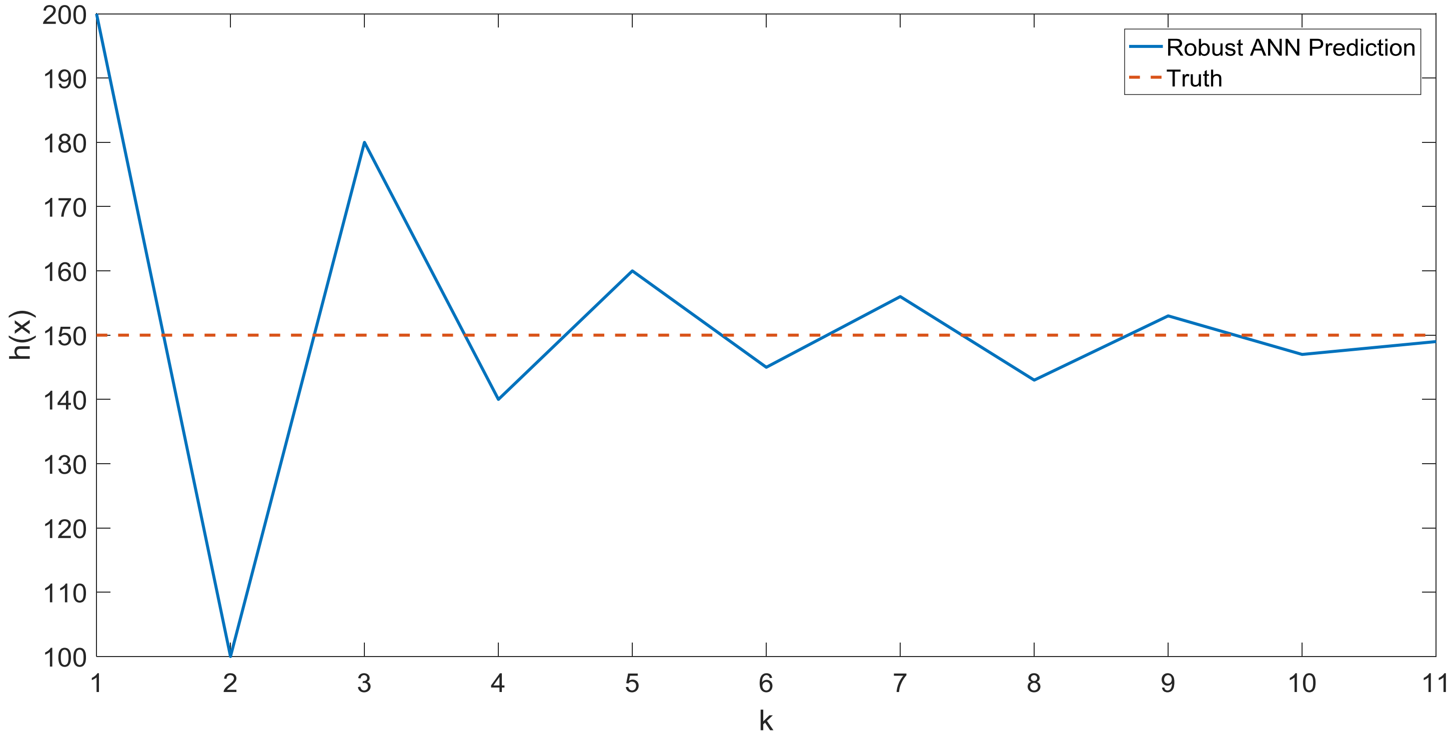
$$h(x) = \frac{2\pi T_u (H_u - H_l)}{\ln(r/r_w)\left(1 + \frac{2LT_u}{\ln(r/r_w) r_w^2 K_w} + \frac{T_u}{T_l}\right)}$$



Parameters	Distribution	Values
r_w	Uniform	[0.05 0.15]
r	Log-Normal	[7.71 1.0056]
T_u	Uniform	[63070 115600]
H_u	Uniform	[990 1110]
T_l	Uniform	[63.1 116]
H_l	Uniform	[700 820]
L	Uniform	[1120 1680]
K_w	Uniform	[9855 12045]

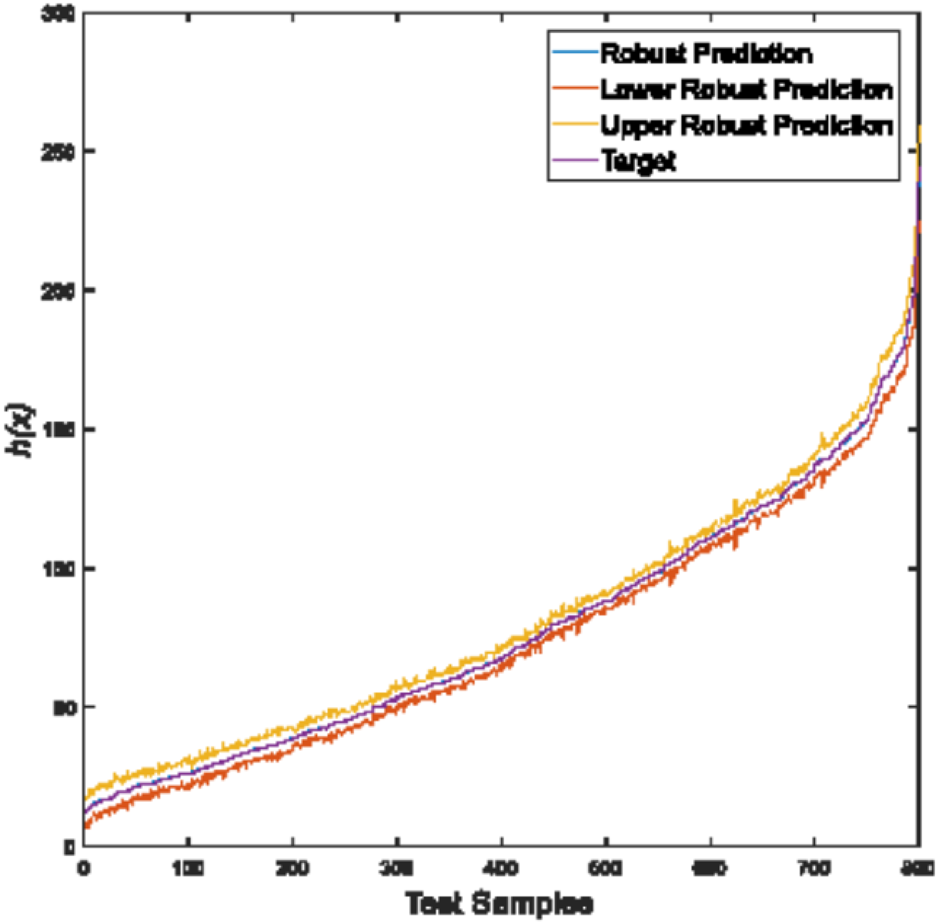
Case Study: Experimental settings

- Design of experiment via LHS (1000 training samples)
- Train ANN
  - 1 hidden layer ANN architecture [8 5 1]
  - Adam optimizer (Tensorflow)

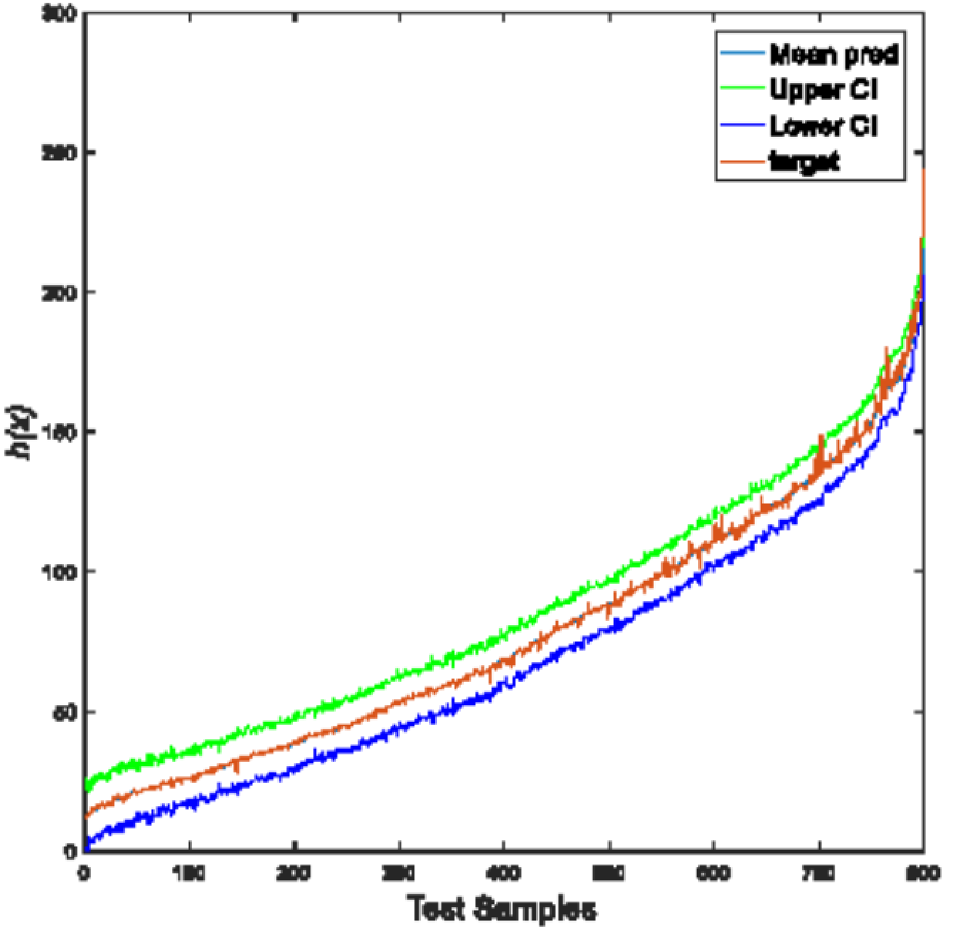


Results

Robust ANN Prediction



Prediction from GP Model

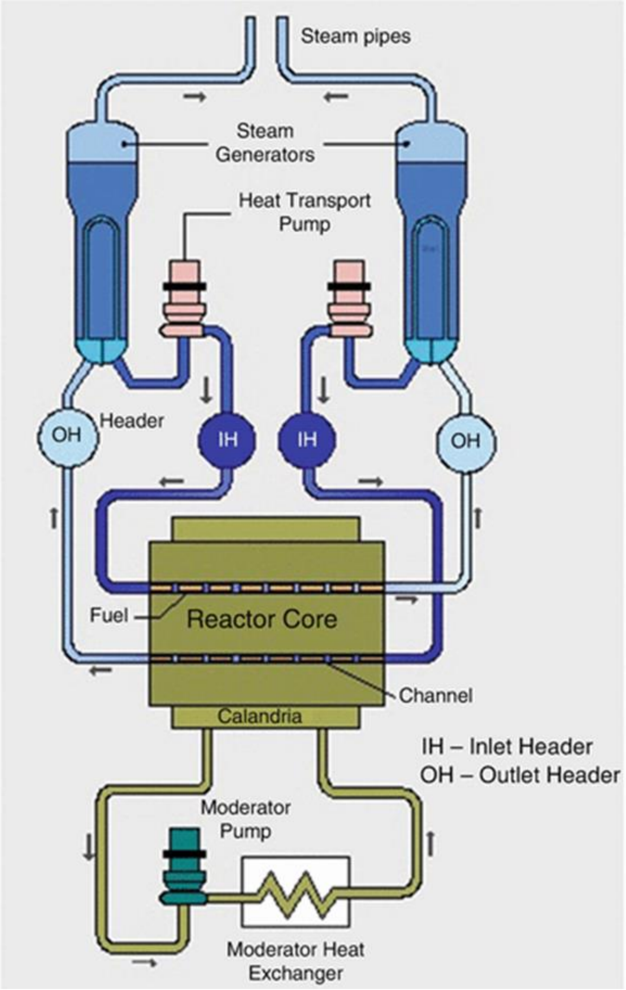


Case Study 2

- ➢ Presents a Nuclear Engineering problem (T V Santhosh, et al (2011)
- ➢ Task is to predict the break size
  - ▪ Pipe transporting fluid
  - ▪ Loss of coolant accident
  - ▪ Pattern recognition tool

Experimental settings

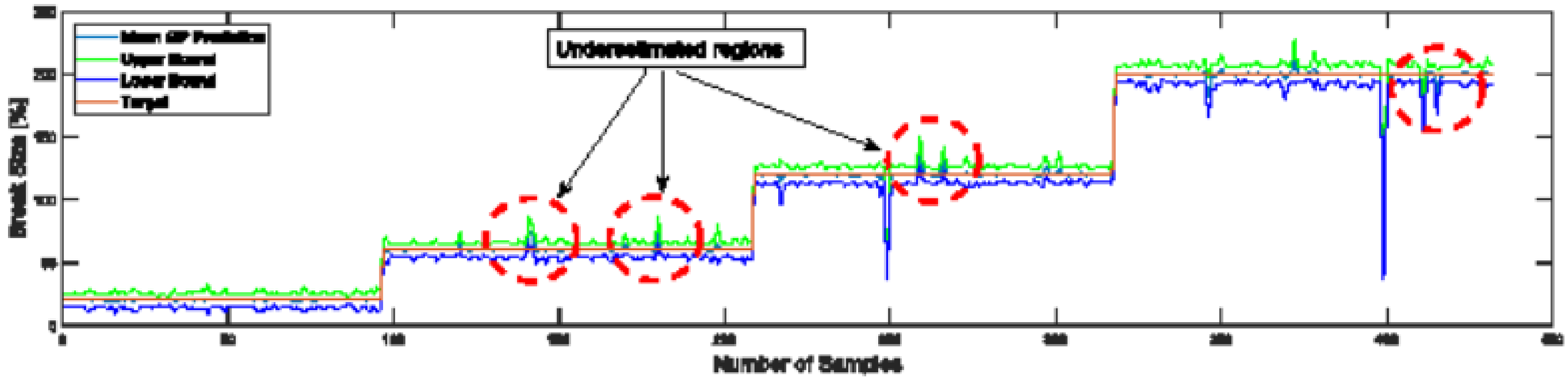
- ➢ Adopting an architecture [37 19 26 1] for k=10 iterations

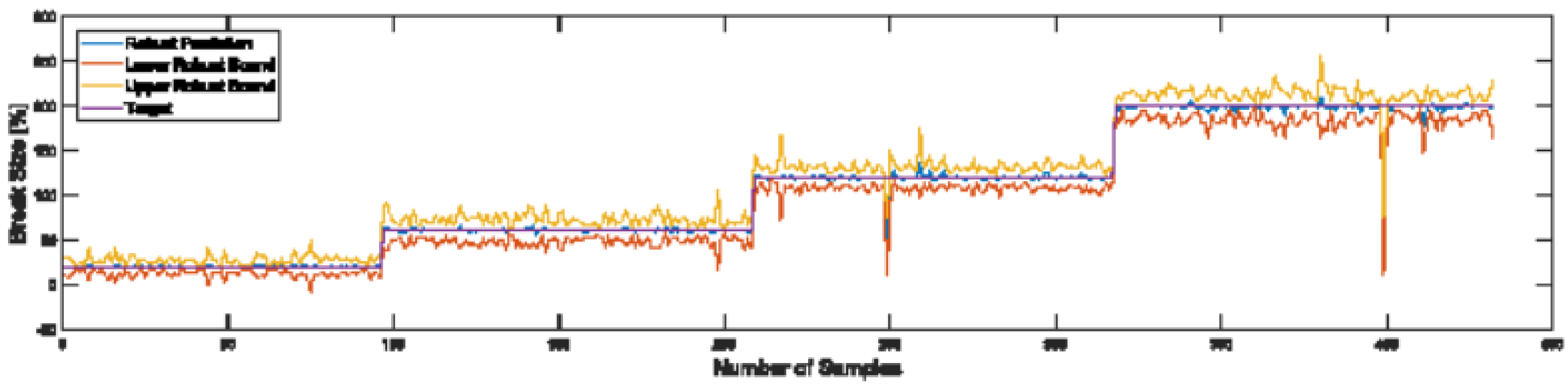


TV Santhosh, et al. A diagnostic system for identifying accident conditions in
a nuclear reactor. Nuclear engineering and design, 241(1):177{184, 2011.

UNIVERSITY OF LIVERPOOL

Results





UNIVERSITY OF LIVERPOOL

Conclusion

- Robust ANN performed better than GP model in terms of predicting tighter confidence bounds.

- Although for some cases the confidence bounds produced by robust ANN are larger than the bounds GP model, the reliability of robust ANN bounds is usually higher.

- For big data set, robust ANN will always perform better than GP model.