Beyond Research on Neural Networks: Data-Driven Robust Optimisation of Combined Cycle Gas Power Plant

Waqar Muhammad Ashraf, Vivek Dua

The Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, University College London, UK

UCL

Problem Statement

- ✓ Neural networks are universal function approximators but black-box [1]
- ✓ Neural networks are point-predictors; do not provide prediction intervals
- ✓ How to carry out data-driven robust optimisation integrating the neural networks for multi-level operation of combined cycle gas power plant ?

Objectives

- ✓ Introduce data-information to infuse interpretability for neural networks
- ✓ Compute prediction intervals using neural network parameters space
- ✓ Present two-stage robust-optimisation framework embedding the neural networks

Data Information Integrated Neural Network (DINN)

- ✓ Feature association is computed by Pearson Correlation Coefficient (PCC)
- ✓ The loss function of DINN is customized to include PCC information for updating the parameters:

$$\mathcal{L} = \left(\frac{\lambda}{1+\lambda}\right).\frac{\sum(D-Z)^2}{N} + \left(\frac{1}{1+\lambda}\right).\frac{\sum_{i=1}^{m}(r_{X_i|D} - r_{X_i|Z})^2}{N} \quad (1)$$

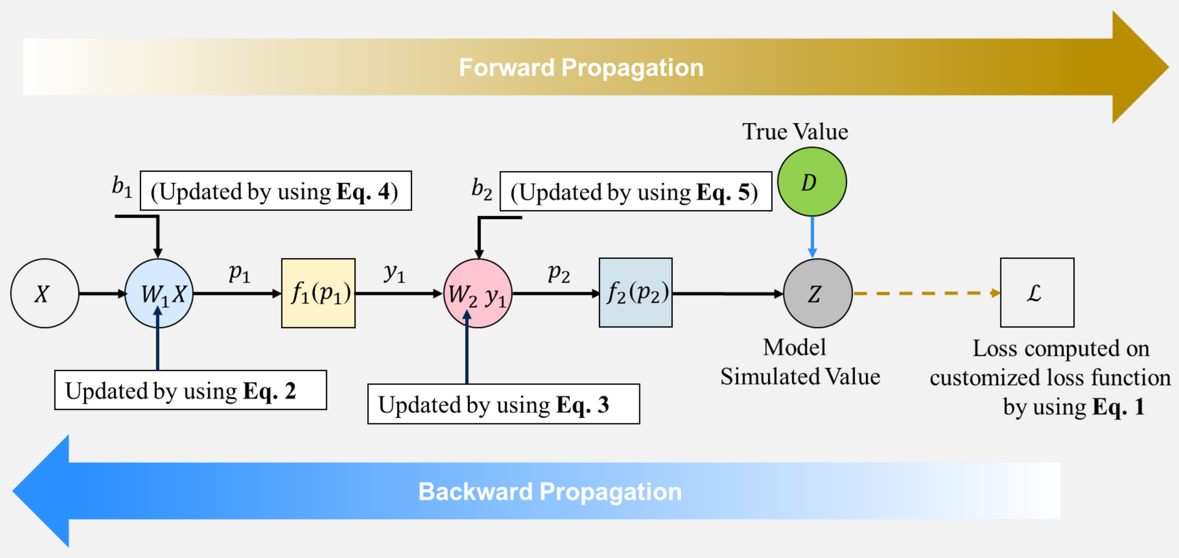
- ✓ Gradient descent with momentum algorithm updates the parameters:

$$W_1^{new} = W_1 + \eta\left(\beta V_{W_1} + (1-\beta)\left(\frac{2(D-Z)}{N} + r_{X_i|Z}.\left(\frac{X_i^{\mu}}{\mathcal{B}_i} - \frac{Z^{\mu}}{\mathcal{M}}\right)\right)W_2^T(1-y_1^2)X^T\right. \quad (2)$$

$$W_2^{new} = W_2 + \eta\left(\beta V_{W_2} + (1-\beta)\left(\frac{2(D-Z)}{N} + r_{X_i|Z}.\left(\frac{X_i^{\mu}}{\mathcal{B}_i} - \frac{Z^{\mu}}{\mathcal{M}}\right)\right)y_1\right. \quad (3)$$

$$b_1^{new} = b_1 + \eta\left(\beta V_{b_1} + (1-\beta)\left(\frac{2(D-Z)}{N} + r_{X_i|Z}.\left(\frac{X_i^{\mu}}{\mathcal{B}_i} - \frac{Z^{\mu}}{\mathcal{M}}\right)W_2^T(1-y_1^2)\right)\right. \quad (4)$$

$$b_2^{new} = b_2 + \eta\left(\beta V_{b_2} + (1-\beta)\left(\frac{2(D-Z)}{N} + r_{X_i|Z}.\left(\frac{X_i^{\mu}}{\mathcal{B}_i} - \frac{Z^{\mu}}{\mathcal{M}}\right)\right)\right. \quad (5)$$



Storage of Weights And Retrieval Method (SWARM)

- ✓ Online-mode of training is implemented for SWARM approach
- ✓ The loss function of neural network incorporates standard deviation term:

$$\mathcal{L} = \frac{(D-Z)^2}{2} + \frac{|D-Z|}{\sqrt{2}} \quad (6)$$

- ✓ Mean absolute difference based non-conformity score for D_1 at each epoch:

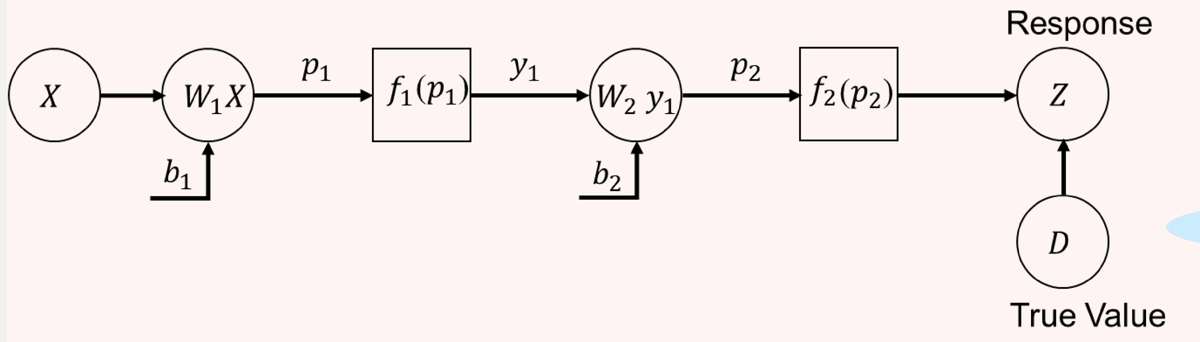
$$E_{1_{epoch}} = |D_1 - (Z_1)_{epoch}| \quad (7)$$

- ✓ SWARM based prediction interval (PI) on quantile value of $\hat{q}_{1-\alpha}$:

$$PI(Z_1)_{SWARM} = [Z_1 - \hat{q}_{1-\alpha}(E_{1_{epoch}}), Z_1 + \hat{q}_{1-\alpha}(E_{1_{epoch}})] \quad (8)$$

- ✓ The inductive conformal prediction (ICP) technique-based PI for test dataset on quantile value of $\hat{q}_{1-\alpha}$:

$$PI(Z_{test})_{ICP} = [Z_{test} - \hat{q}_{1-\alpha}(E_{cal}), Z_{test} + \hat{q}_{1-\alpha}(E_{cal})] \quad (9)$$



Data-Driven Robust Optimisation

- ✓ The multi-objective optimisation function is defined:

$$\min_x f(x) = (f_1(x) + f_2(x) + \cdots + f_n(x)) \quad (10)$$

subject to:

$$h(x) = 0$$
$$x^L \leq x \leq x^U \quad (11)$$

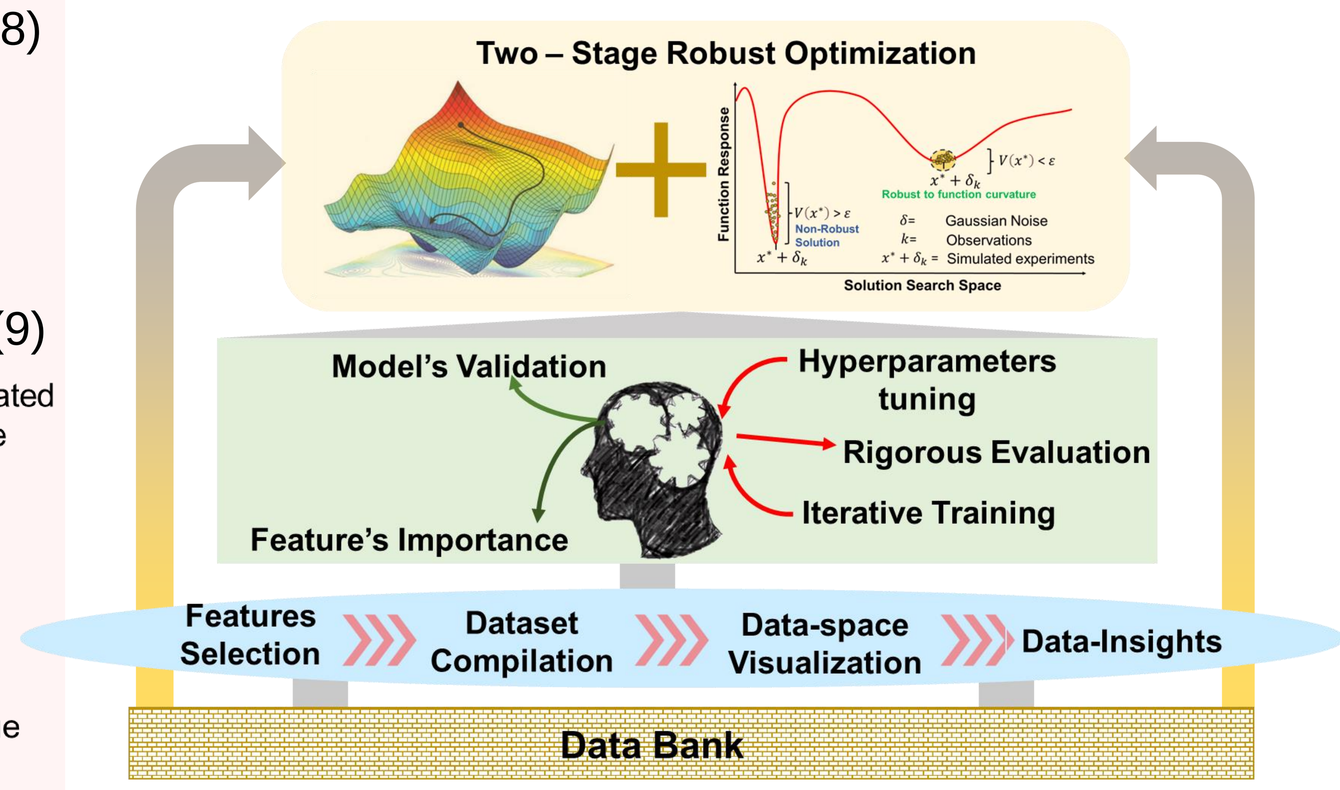
- ✓ The mean and variance in $f(x^*)$ for Monte Carlo simulations are computed as:

$$F(x^*) = \frac{\sum_{k=1}^{H} f(x^* + \delta_k)}{H} \quad (12)$$

$$V(x^*) = \frac{\|F(x^*) - f(x^*)\|}{\|f(x^*)\|} < \varepsilon \quad (13)$$

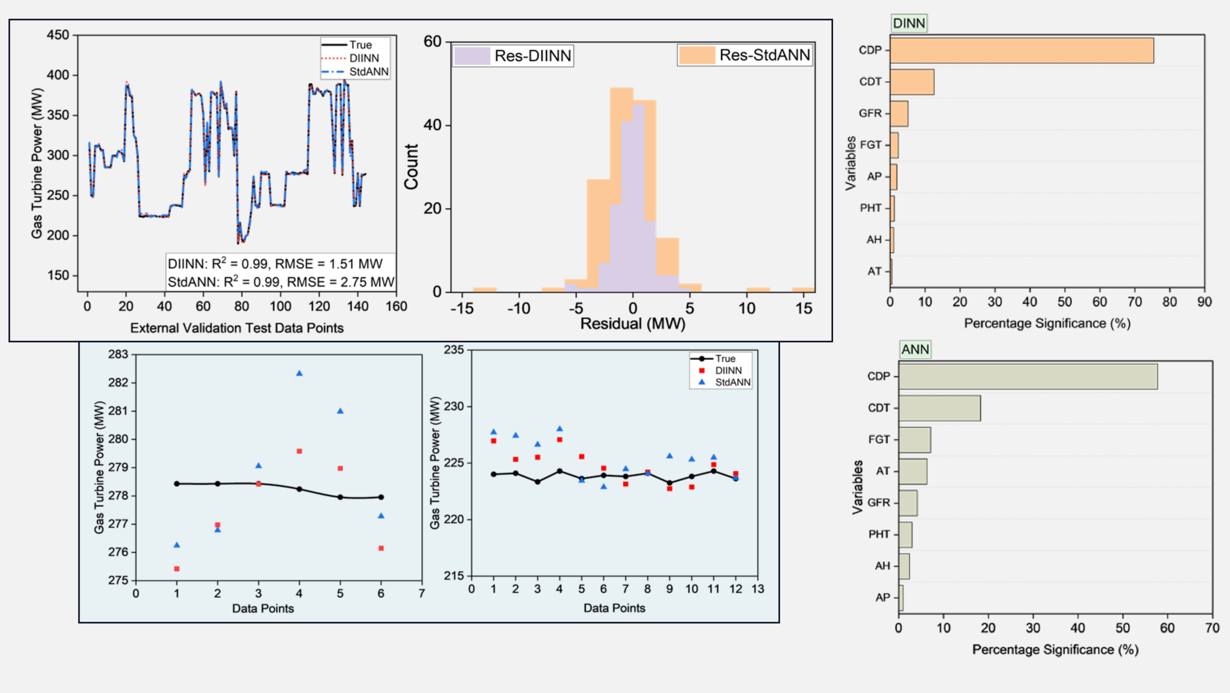
- ✓ The solution is "non-robust" when:

$$V(x^*) > \varepsilon \quad (14)$$

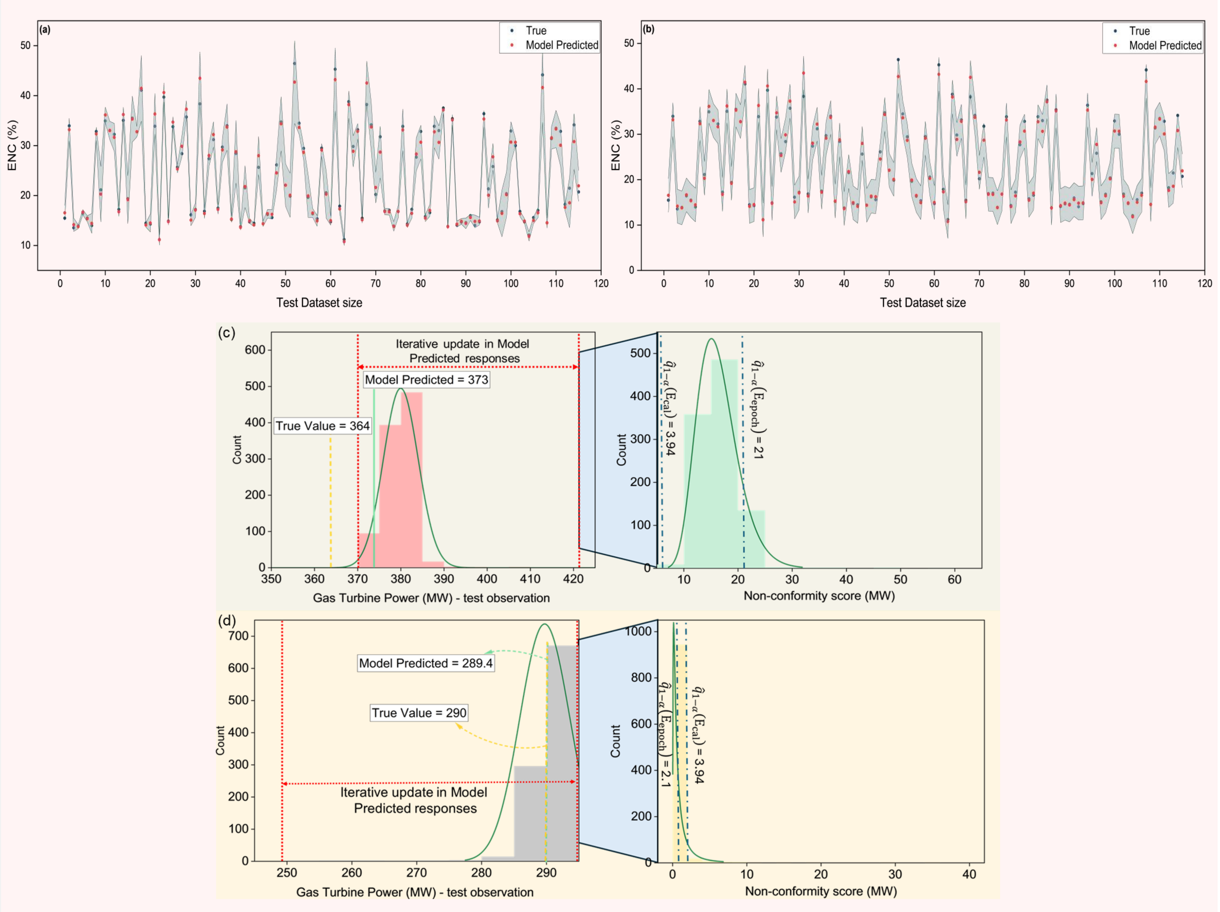


Case Studies

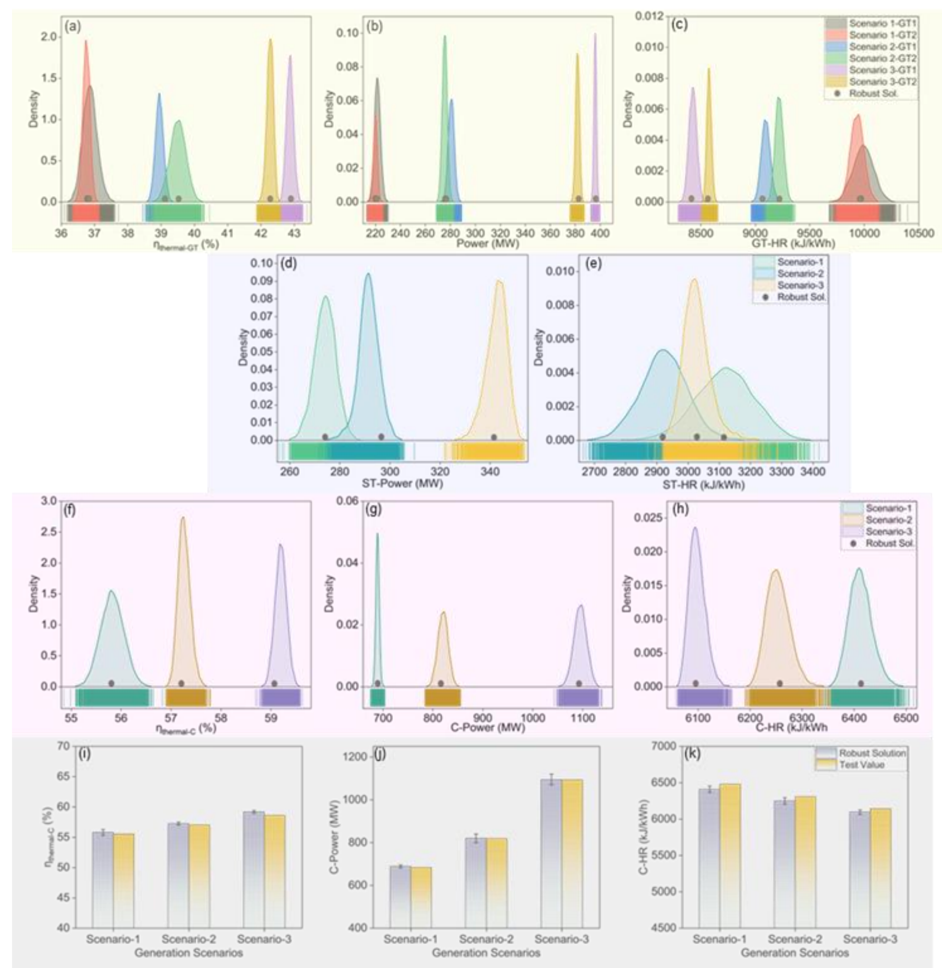
- ✓ DINN and ANN are trained to predict power generation from 395 MW capacity gas turbine system
- ✓ The feature importance for the power generation is established for well-trained DINN and ANN models [2]



- ✓ The PI are constructed by SWARM and ICP techniques
- ✓ Energy efficiency cooling data is used [3]



- ✓ Data-driven robust optimisation of 1180 MW capacity combined cycle gas power plant is carried out under top-down approach [4]



Conclusions & Future Work

- ✓ Expressions for the tuning of the model parameters are derived
- ✓ DINN based feature importance order better complies with domain knowledge than those of ANN
- ✓ SWARM based prediction intervals are better local–compliant than the fixed–width prediction intervals of ICP
- ✓ Top-down robust optimisation of combined cycle gas power plant reduces 62 ± 20 kt of CO_2 annually
- ✓ The future work will focus on bi-level optimisation of combined cycle gas power plant with neural networks

References

[1] Benítez, José Manuel, Juan Luis Castro, and Ignacio Requena. "Are artificial neural networks black boxes?." *IEEE Transactions on neural networks* 8, no. 5 (1997): 1156-1164

[2] Ashraf, Waqar Muhammad, and Vivek Dua. "Data Information integrated Neural Network (DINN) algorithm for modelling and interpretation performance analysis for energy systems." Energy and AI 16 (2024): 100363

[3] Ashraf, Waqar Muhammad, and Vivek Dua. "Storage of weights and retrieval method (SWARM) approach for neural networks hybridized with conformal prediction to construct the prediction intervals for energy system applications." *International Journal of Data Science and Analytics* (2024): 1-15.

[4] Ashraf, Waqar Muhammad, and Vivek Dua. "Driving towards net-zero from the energy sector: Leveraging machine intelligence for robust optimization of coal and combined cycle gas power stations." *Energy Conversion and Management* 314 (2024): 118645.

Contact Details

Waqar Muhammad Ashraf: Waqar.ashraf.21@ucl.ac.uk

Vivek Dua: v.dua@ucl.ac.uk

Affiliation

Sargent Centre for Process Systems Engineering

Funding Source