Surrogate Modelling of the Tritium Breeding Ratio

Petr Mánek   Graham Van Goffrier

Centre for Doctoral Training in Data Intensive Science
University College London

22nd June 2020

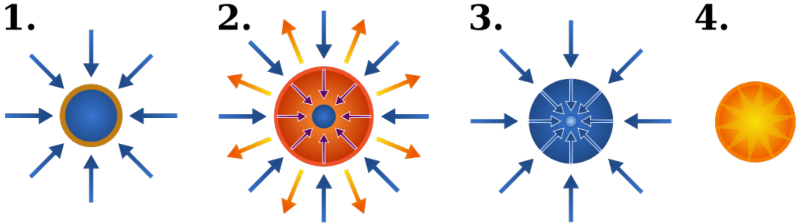
UCL CDT DIS

UK Atomic
Energy
Authority

Nuclear fusion – the energy of the future!

- Must produce and contain an extremely hot and dense plasma
  - Magnetic Confinement Fusion (MCF): toroidal circulation
  - Inertial Confinement Fusion (ICF): spherical compression

- Modern designs require enriched Hydrogen fuel of two varieties:
  - Deuterium (^{2}H) – abundant in naturally-sourced water
  - Tritium (^{3}H) – extremely rare, but can be produced *in-reactor*



Problem Description

Tritium breeding blankets convert neutron radiation from the fusion plasma into a steady supply of tritium fuel.

$$ {}^{1}_{0}n + {}^{6}_{3}Li \rightarrow {}^{3}_{1}H + {}^{4}_{2}He \qquad {}^{1}_{0}n + {}^{7}_{3}Li \rightarrow {}^{3}_{1}H + {}^{4}_{2}He + {}^{1}_{0}n $$

However, the tritium breeding ratio (TBR) depends on numerous geometric and material parameters.

TBR evaluation *Paramak* achieves very accurate results by OpenMC Monte Carlo neutronics simulation, but is computationally expensive.

Our Challenge:

Produce a fast TBR function which strongly approximates Paramak, making use of the latest in surrogate modelling techniques.

Data Generation

We designed the Approximate TBR Evaluator (ATE) package to generate training and test datasets from Paramak.

UCL's Hypatia cluster provided the multithreading power for us to produce one million TBR samples, representing 27 days of runtime.

These runs included full evaluations on the 18 continuous and discrete parameters of Paramak, and "slice" evaluations with all discrete parameters frozen.

	Parameter name	Domain
Blanket	Breeder fraction†	$[0, 1]$
	Breeder ^{6}Li enrichment fraction	$[0, 1]$
	Breeder material	$\{Li_2TiO_3, Li_4SiO_4\}$
	Breeder packing fraction	$[0, 1]$
	Coolant fraction†	$[0, 1]$
	Coolant material	$\{D_2O, H_2O, He\}$
	Multiplier fraction†	$[0, 1]$
	Multiplier material	$\{Be, Be_{12}Ti\}$
	Multiplier packing fraction	$[0, 1]$
	Structural fraction†	$[0, 1]$
	Structural material	$\{SiC, eurofer\}$
	Thickness	$[0, 500]$
First wall	Armour fraction‡	$[0, 1]$
	Coolant fraction‡	$[0, 1]$
	Coolant material	$\{D_2O, H_2O, He\}$
	Structural fraction‡	$[0, 1]$
	Structural material	$\{SiC, eurofer\}$
	Thickness	$[0, 20]$

Methodology

Conventional regression task – search for a cheap surrogate $\hat{f}(x)$ that minimizes dissimilarity with an expensive function $f(x)$:

- Regression performance: mean absolute error, σ of error, R^2, $R^2_{\text{adj.}}$
- Computational complexity: training & prediction time / sample.

2 approaches for surrogate training:

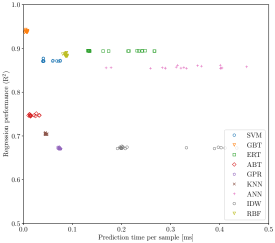
1. Decoupled – trains models from previously sampled datapoints.
2. Adaptive – repeats sampling & model training, increases sampling density in low-performance regions.

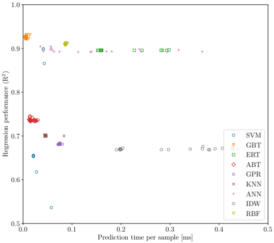
Decoupled Approach

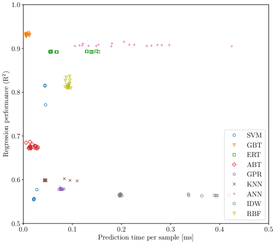
Outline

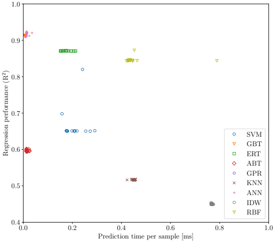
Compare 9 state-of-the-art surrogate families in 4 experiments:

1. Hyperparameter tuning (simplified) – Bayesian optimization, discrete features fixed & withheld.
2. Hyperparameter tuning – same as #1 but with all features.
3. Scaling benchmark – increase training set size.
4. Model comparison – train surrogates for practical use.





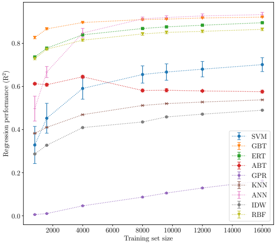


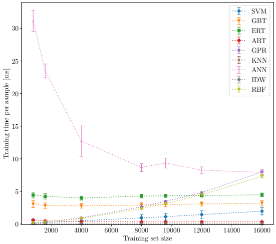


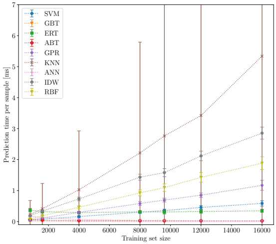
- Experiment 1 (3× top), Experiment 2 (left). Plots show $\bar{t}_{\text{pred.}}$ vs. R^2 (best = upper left).
- Showing the best 20 surrogates per family.
- Omitting discrete features yields only a negligible improvement in performance.
- Overall dominated by tree-based surrogates (GBTs, ERTs) and neural networks.

Experiment 3: Scaling Benchmark

- We observe a hierarchy.
- Best-performing families from the previous experiments also scale the best in $\bar{t}_{\text{pred.}}$.
- More samples: neural networks outperform tree-based models.
- Instance-based surrogates (KNN, IDW) train trivially but have complex lookup.
- Neural networks show inverse scaling due to parallelization.

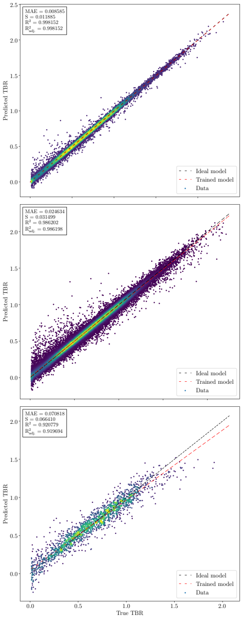






Experiment 4: Model Comparison

- Trained 8 models for practical use.
- Plots show true vs. predicted TBR by Models 1, 2 & 4, coloured by density.
- Best regression performance:
  - Model 1: ANN, 500K samples.
  - $R^2 = 0.998$, $\sigma = 0.013$,
  - $\bar{t}_{\text{pred.}} = 1.124\,\mu\text{s}$, $6\,916\,416\times$ faster.
- Fastest prediction:†
  - Model 2: ANN, 500K samples.
  - $R^2 = 0.985$, $\sigma = 0.033$,
  - $\bar{t}_{\text{pred.}} = 0.898\,\mu\text{s}$, $8\,659\,251\times$ faster.
- Smallest training set:†
  - Model 4: GBT, 10K samples.
  - $R^2 = 0.913$, $\sigma = 0.072$,
  - $\bar{t}_{\text{pred.}} = 6.125\,\mu\text{s}$, $1\,269\,777\times$ faster.



† with acceptable regression performance.

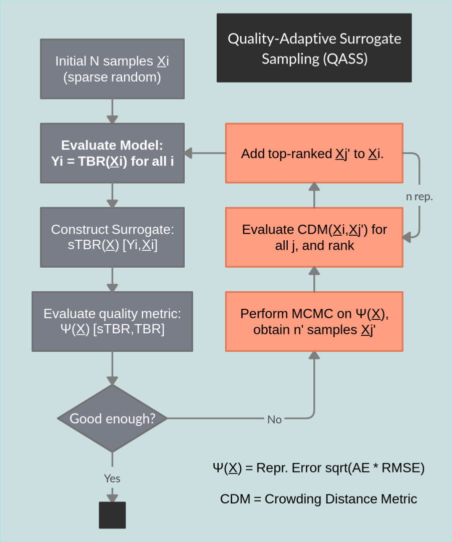
Adaptive Approach

Adaptive Sampling: Theory

Adaptive sampling takes advantage of surrogate information content *during training* to reduce sample quantity.

We developed a technique which is novel in the literature:

- Construct surrogate quality distribution by nearest-neighbour interpolation.
- Draw candidate samples by quality using MCMC.
- Include samples with high crowding distance.
- Repeat!



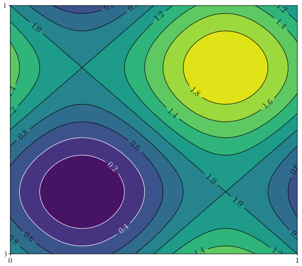
Application on Toy Theory

Toy functional TBR theory with wavenumber n, and similar ANN performance to Paramak:

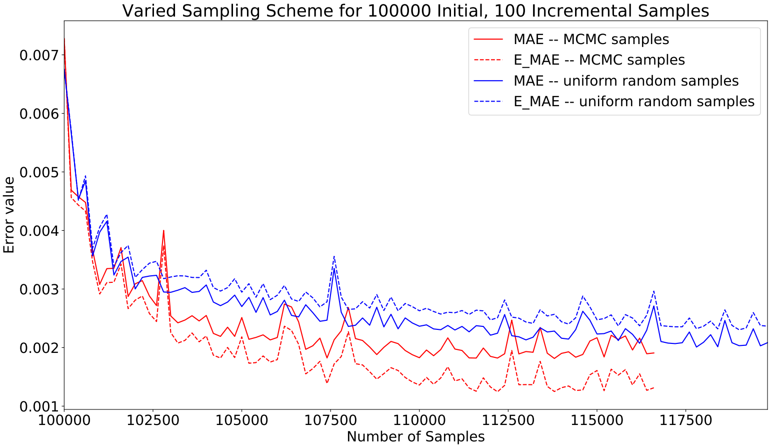
$$\mathrm{TBR} = \frac{1}{|C|} \sum_{i \in C} \left[1 + \sin(2\pi n(x_i - 1/2))\right]$$

Evaluation was performed both on the adaptively-sampled dataset and on a baseline uniform-random dataset.

An baseline scheme with incremental but uniform-random samples was used as a placebo comparison.



Adaptive Sampling: Results



Conclusion