

# Surrogate Modelling of the Tritium Breeding Ratio

Petr Mánek   Graham Van Goffrier

Centre for Doctoral Training in Data Intensive Science  
University College London

8th October 2020



UK Atomic  
Energy  
Authority

## Nuclear fusion – the energy of the future!

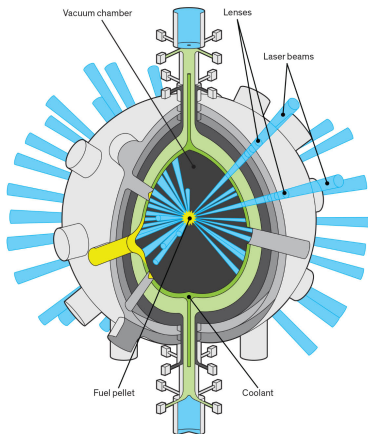
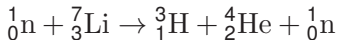
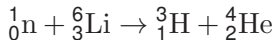


Illustration by Chris Philpot, courtesy of IEEE Spectrum.

- Designing next-generation Inertial Confinement Fusion (ICF) facility.
- Search for optimal reactor design.
- **Fueling** – important for viability.
- Require fuel of 2 varieties:
  - Deuterium  $^2\text{H}$  – abundant in naturally-sourced water.
  - Tritium  $^3\text{H}$  – **extremely rare**.
- Modern reactors can generate tritium during operation.

**Tritium breeding blankets** convert neutron radiation to  $^3\text{H}$  fuel:



$^3\text{H}$  balance described by **Tritium Breeding Ratio (TBR)** =  $\frac{\text{fuel bred}}{\text{fuel consumed}}$

- Depends on numerous geometric and material parameters.
- Evaluated by *Paramak* – OpenMC neutronics simulation.
- Slow . . . we want to consider as many reactor designs as possible!

## Our Challenge

Produce a *fast* TBR surrogate that strongly approximates Paramak.

Produced datasets by sampling Paramak outputs over its 7 discrete and 11 continuous input parameters at random.

Deployed at UCL's

Hypatia cluster:

- Created **1M points**.
- **27 days** of runtime.

2 classes of runs:

- All parameters free.
- Discrete fixed, continuous free.

	Parameter name	Domain
Blanket	Breeder fraction <sup>†</sup>	[0, 1]
	Breeder <sup>6</sup> Li enrichment fraction	[0, 1]
	Breeder material	{Li <sub>2</sub> TiO <sub>3</sub> , Li <sub>4</sub> SiO <sub>4</sub> }
	Breeder packing fraction	[0, 1]
	Coolant fraction <sup>†</sup>	[0, 1]
	Coolant material	{D <sub>2</sub> O, H <sub>2</sub> O, He}
	Multiplier fraction <sup>†</sup>	[0, 1]
	Multiplier material	{Be, Be <sub>12</sub> Ti}
	Multiplier packing fraction	[0, 1]
	Structural fraction <sup>†</sup>	[0, 1]
	Structural material	{SiC, eurofer}
First wall	Thickness	[0, 500]
	Armour fraction <sup>†</sup>	[0, 1]
	Coolant fraction <sup>†</sup>	[0, 1]
	Coolant material	{D <sub>2</sub> O, H <sub>2</sub> O, He}
	Structural fraction <sup>†</sup>	[0, 1]
	Structural material	{SiC, eurofer}
	Thickness	[0, 20]

Groups of parameters marked<sup>†‡</sup> are required to sum to 1.

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,  $\sigma$  of error,  $R^2$ ,  $R_{\text{adj}}^2$ .
- Computational complexity: training & prediction time / sample

2 approaches to solution:

## Decoupled Approach

- 1 Collect training dataset.
- 2 Use data to train a surrogate.

## Adaptive Approach

- 1 Collect initial training dataset.
- 2 Use data to train a surrogate.
- 3 Collect more data in regions where surrogate performed poorly.
- 4 Repeat steps 2 and 3.

# Decoupled Approach

Compared 9 state-of-the-art surrogate families:

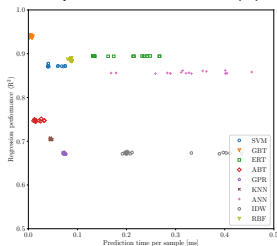
- Support vector machines,
- Gradient boosted trees,
- Extremely randomized trees,
- AdaBoosted decision trees,
- Gaussian process regression,
- $k$  nearest neighbors,
- Artificial neural networks (MLP),
- Inverse distance weighting,
- Radial basis functions.

Performed 4 experiments:

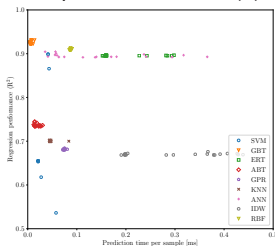
- 1 Hyperparameter tuning (simplified) – Bayesian optimization, discrete features fixed & withheld.
- 2 Hyperparameter tuning – same as #1 but with all features.
- 3 Scaling benchmark – increase training set size.
- 4 Model comparison – train surrogates for practical use.

# Experiments 1 & 2: Hyperparameter Tuning

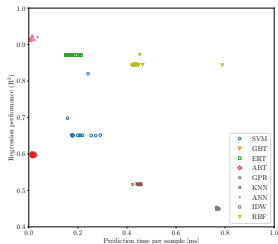
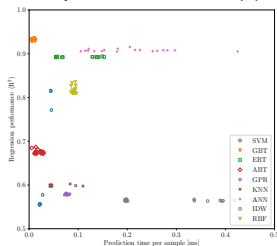
Experiment 1, slice (a)



Experiment 1, slice (b)



Experiment 1, slice (c)



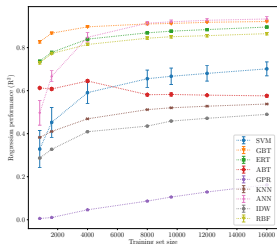
Experiment 2

- Showing  $\bar{t}_{\text{pred.}}$  vs.  $R^2$  for the 20 best surrogates per family (top left  $\Leftrightarrow$  fastest, most accurate).
- Omit discrete features  $\rightarrow$  negligible performance improvement.
- Dominated by **trees** (GBTs, ERTs) and **neural networks**.

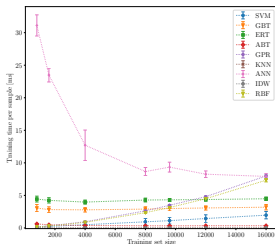


# Experiment 3: Scaling Benchmark

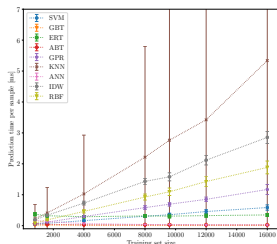
- We observe a hierarchy.
- **Trees** and **neural networks** scale the best in  $\bar{t}_{\text{pred.}}$ .
- Maximizing training set size, **neural networks** dominate.
- Instance-based surrogates (KNN, IDW) train trivially but have slow lookup.
- Neural networks show inverse scaling due to parallelization.



Regression performance



Training time / sample



Prediction time / sample

# Experiment 4: Model Comparison

Trained **8 models** for practical use.

## Model 1, best regression performance

- ANN (4-layer MLP), 500K samples.
- $R^2 = 0.998$ ,  $\sigma = 0.013$ ,
- $\bar{t}_{\text{pred.}} = 1.124 \mu\text{s}$ , **speedup 6 916 416 $\times$**

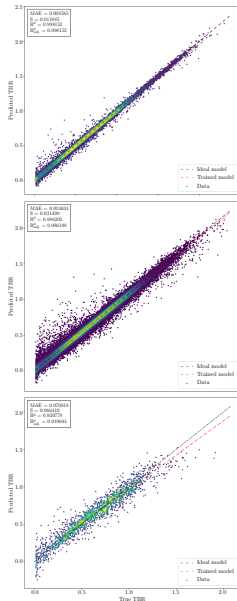
## Model 2, fastest prediction<sup>†</sup>

- ANN (2-layer MLP), 500K samples.
- $R^2 = 0.985$ ,  $\sigma = 0.033$ ,
- $\bar{t}_{\text{pred.}} = 0.898 \mu\text{s}$ , **speedup 8 659 251 $\times$**

## Model 4, smallest training set<sup>†</sup>

- GBT, **10K samples**.
- $R^2 = 0.913$ ,  $\sigma = 0.072$ ,
- $\bar{t}_{\text{pred.}} = 6.125 \mu\text{s}$ , **speedup 1 269 777 $\times$**

<sup>†</sup> with acceptable regression performance.

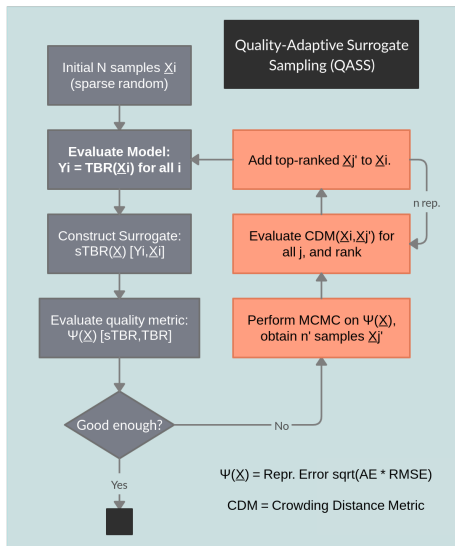


# Adaptive Approach

How to take advantage of surrogate information content *during training* to reduce sample quantity?

We developed **a new technique**:

- 1 Construct surrogate quality distribution by nearest-neighbour interpolation.
- 2 Draw candidate samples by quality using MCMC.
- 3 Include samples with greatest separation from neighbours.
- 4 Repeat!



Toy functional TBR theory with wavenumber  $n$ , and qualitatively comparable ANN performance to Paramak:

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

(where  $C$  enumerates all continuous variables)

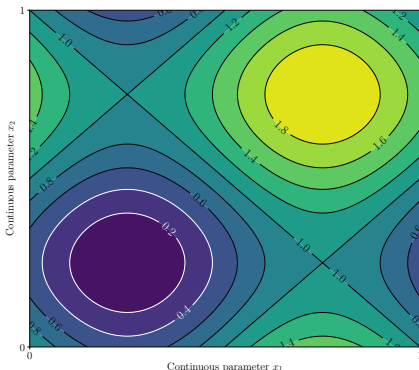
Evaluation set:

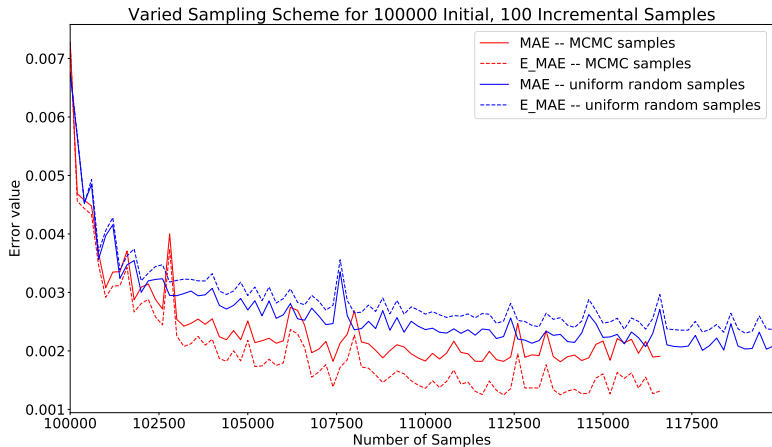
- Adaptive samples
- Generated during runtime

Validation set:

- Uniform random samples
- Generated independently

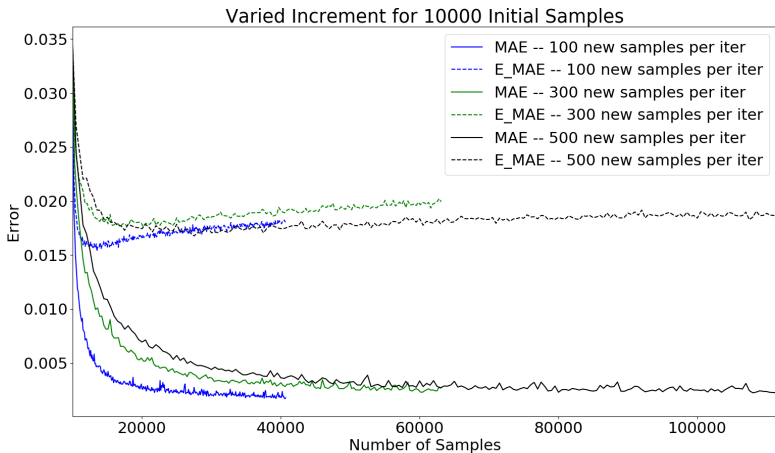
Placebo comparison – incremental uniform-random samples, no MCMC.





60% decrease in MAE for validation set (dashed)

Equivalently, 6% decrease in samples needed for same accuracy



Fewer incremented samples can lead to better accuracy!

But depends on initial samples, specific model – further study needed.

## Decoupled Approach

- Heuristic: **GBTs** for  $< 10^4$  samples, **ANNs** for  $\geq 10^5$  samples.
- Fastest found surrogate evaluates TBR in  $0.898 \mu\text{s}$  with error 0.033. This is roughly  $8 \cdot 10^6 \times$  **faster** than Paramak.
- Found surrogates with comparable properties with  $\approx 10\text{K}$  **samples**.

## Adaptive Approach

- New theoretical approach **QASS** developed, based on MCMC.
  - **60% decrease** in evaluation MAE demonstrated.
  - **6% decrease** in expensive TBR samples needed.
- 
- All methods **portable**  $\rightarrow$  cheap approximation of any simulation.
  - Article in IOP **Journal of Nuclear Fusion** (pending).
  - Included as a benchmark in the **SciML Collaboration**.



# Thank you for listening!

Petr Mánek ..... petr.manek.19@ucl.ac.uk  
Graham Van Goffrier ..... graham.vangoffrier.19@ucl.ac.uk

## Further Reading

- Single page abstract (available online).
- Journal article, currently in internal pre-submission review (available online):  
*Fast Regression of the Tritium Breeding Ratio in Fusion Reactors.*
- Industry group project final report (available online).
- All models, plots, training data, source code and technical documentation.  
<https://github.com/ucl-tbr-group-project>