Science and Technology Facilities Council

Scientific Computing

stfc.ukri.org

# Towards Smart Scientific Facilities: Uncertainty-Aware Modelling and Optimisation


**Xingchi Liu[1], Siu-lun Yeung[1], Jaehoon Cha[1], Boris Shustin[1], Kathryn Baker[2], Raunakk Banerjee[2], Gareth Nisbet[3], Richard Cunningham[3], Brett Sing Gen[3], Jeyan Thiyagalingam[1]**

1. Scientific computing department, Harwell Campus, Oxfordshire, OX11 0QX, UK
2. ISIS Neutron and Moun source, Harwell Campus, Oxfordshire, OX11 0QX, UK
3. Diamond Light Source, Harwell Campus, Oxfordshire, OX11 0QX, UK


diamond

ada lovelace centre

ISIS Neutron and Muon Source

## Introduction

Accelerators and beamline instruments require careful tuning to achieve reliable performance. This often relies on specialist expertise and repeated adjustments, while coupled parameters, changing conditions and costly evaluations make automation challenging.

We explore physics-aware surrogate modelling and few-shot learning to reduce this tuning burden. For the ISIS MEBT, Fourier feature embeddings and HyperGAN ensembles support beam prediction and uncertainty estimation, with latent calibration enabling adaptation to shifted beam conditions. We learn Gaussian process priors across related tasks to accelerate Bayesian optimisation of beam loss. Further studies investigate automated ISIS LEBT tuning and trust-region Bayesian optimisation for motorised-stage controllers at Diamond Light Source.

Our broader goal is to help scientific facilities respond to changing conditions with less manual intervention, freeing time and expertise for experimental science.

## ISIS MEBT Model

The MEBT model is a neural network with two shared layers of 512 nodes, followed by two branches, each with two layers of 512 nodes. One branch focuses on beam loss, while the other predicts beam descriptors using Fourier feature embeddings (FFE). The model is trained on 3,000 ASTRA simulations and optimised using Adam.

**Model inputs:**
- Incoming beam distribution
- Beam current
- 4x Cavity field strengths
- 8x Quadrupole strengths

**Model Outputs:**
- Envelope
- Position
- Emittance
- Bunch length
- Long. energy spread
- Beam loss

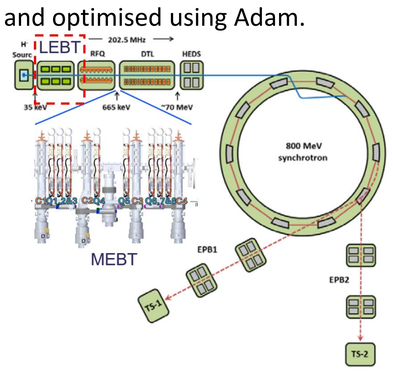



Figure 1:Schematic of the ISIS LEBT and MEBT

**Fourier Feature Embeddings:**

The 1D z-coordinate of the 200 discrete positions along MEBT is fed into the FF mapping and concatenated with the machine parameters before being passed to the NN.

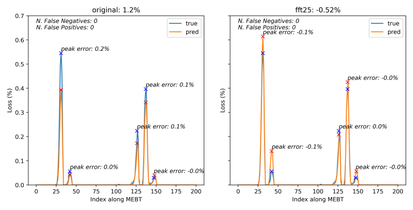



Figure 2: Use of FF mappings with embedding dimension m=25 produces outputs that match the beam loss profile more closely.

## Hyper-GAN for Uncertainty aware MEBT Surrogate

Hyper-GAN is a hypernetwork-based generative adversarial network (HGSM) that learns a compact latent space to generate diverse neural network architectures. Instead of directly mapping inputs to outputs, it generates weights for a target network, providing two main benefits: Adaptive Model Generation and Uncertainty Estimation.

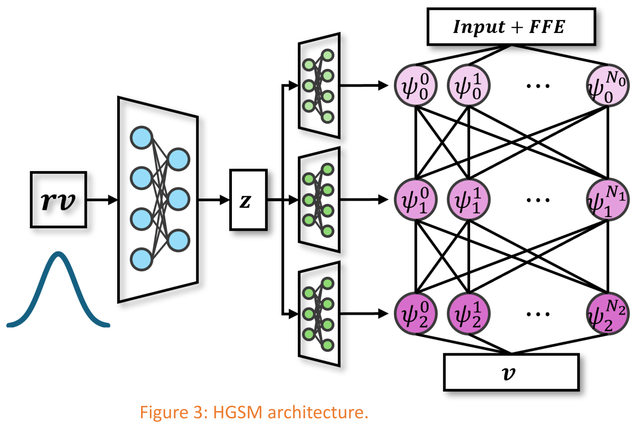



Figure 3: HGSM architecture.

Hyper-GAN offers the following key advantages:
- Generating **Multiple Predictions from a Single Model**: By sampling from a learned latent space, Hyper-GAN mimics an ensemble without the overhead of maintaining multiple models.
- **Dynamic Adaptability**: The method can dynamically generate models tailored to specific beam conditions, providing a flexible solution that traditional static ensembles cannot offer.

## HGSM Results

To estimate uncertainty, 100 neural networks are generated using Hyper-GAN and applied to both the training and validation datasets. From the same input, these networks produce 100 different outputs. By taking the minimum and maximum values across these outputs, uncertainty can be quantified and visualised as shaded areas.

To further illustrate the results, prediction versus true values with uncertainty are presented for the training, validation, and test sets. Each image contains 12 subplots representing loss percent, $mean_x$, $mean_y$, $sigma_x$, $sigma_y$, normalised $emittance_x$, normalised $emittance_y$, $sigma_z$, sigma of energy, $cor_{x_xp}$, $cor_{y_yp}$, and cumulative loss.

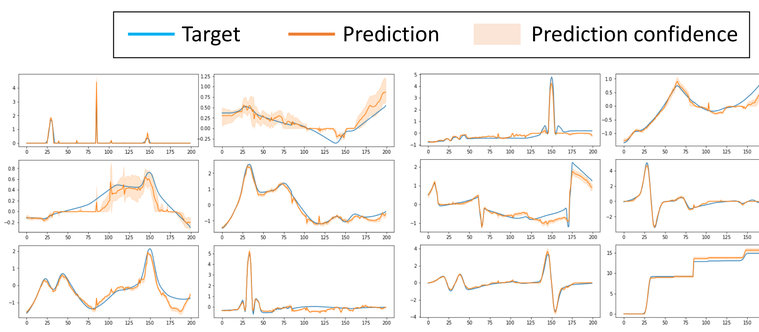



Figure 4: Predictions of beam descriptors using Foureir Features and using 100 models from HyperGAN

We tested whether the HGSM can adapt to shifted beam conditions without retraining the Mixer or Generator.
Calibration was performed using nested subsets of 64, 128, 256, 512 and 1024 samples, and all calibrated models were evaluated on the same fixed 128-sample held-out test set

| Samples | HGSM calibration | | | Trained from scratch | | |
|---|---|---|---|---|---|---|
| | $\mathcal{L}_{\text{loss}}$ | $H$ | Time | $\mathcal{L}_{\text{loss}}$ | $H$ | Time |
| Before | 0.0433 | 1.2481 | – | – | – | – |
| 64 | **0.0288** | **1.1868** | 7.41 s | 0.0871 | 1.7673 | 4.87 s |
| 128 | **0.0298** | **1.1907** | 6.61 s | 0.0778 | 1.6419 | 9.96 s |
| 256 | **0.0293** | **1.1868** | 7.48 s | 0.0808 | 1.4151 | 19.77 s |
| 512 | **0.0297** | **1.1848** | 10.04 s | 0.0751 | 1.2647 | 38.94 s |
| 1024 | **0.0310** | 1.1855 | 18.05 s | 0.0702 | **1.1753** | 78.62 s |

Figure 5: Few-shot latent calibration under shifted beam conditions

## MEBT few shot optimisation via meta learning

To minimise the final **beam loss at the end of MEBT efficiently,** we learn a prior by construct meta-learning tasks via exploiting the structure of the surrogate model.

$$\text{Beam loss} := f(\mathbf{x}_{\text{in}},\ \mathbf{x}_{\text{mag}}), \quad \mathbf{x}_{\text{in}} \in \mathbb{R}^9,\ \mathbf{x}_{\text{mag}} \in \mathbb{R}^{12},$$
$$g_i(\mathbf{x}_{\text{mag}}) = f_{\text{HGSM}}(\mathbf{x}_{\text{in}}^{(i)},\ \mathbf{x}_{\text{mag}})$$

We generate different tuning tasks by randomly sampling $\mathbf{x}_{\text{in}}$, then for each fixed $\mathbf{x}_{\text{in}}$, we sample control parameters $\mathbf{x}_{\text{mag}}$ to generate task-specific datasets

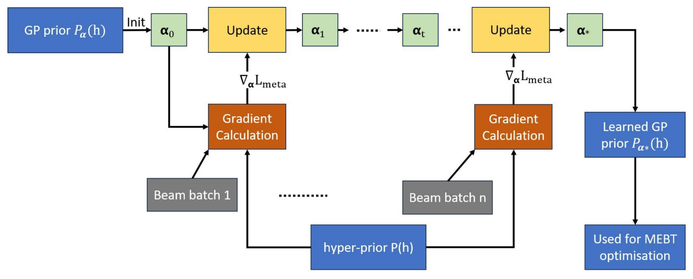



Figure 6: Illustrations of meta-training process

Parameterise the prior as a deep kernel, enabling the prior to model complex, input-dependent patterns that standard GPs with fixed stationary kernels cannot capture.
The parameters of the prior are learned by minimising the following objective:

$$\mathcal{L}_{\text{meta}}(\boldsymbol{\alpha}) = \frac{1}{N}\sum_{i=1}^{N}\Big[-\frac{1}{T_i}\log P_{\boldsymbol{\alpha}}\big(y^{\mathcal{D}_i} \mid \mathbf{X}^{\mathcal{D}_i}\big) + \lambda_i\,\mathrm{KL}\Big(P_{\boldsymbol{\alpha}}\big(h^{\widetilde{\mathbf{X}}_i}\big) \parallel P\big(h^{\widetilde{\mathbf{X}}_i}\big)\Big)\Big].$$

The first encourages the learned GP prior to explain the meta-training observations, whereas the second term regularizes its finite-dimensional function-space distributions towards those induced by the hyper-prior.

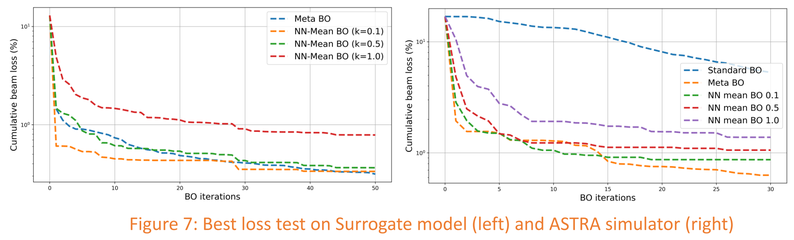



Figure 7: Best loss test on Surrogate model (left) and ASTRA simulator (right)

## Controller Parameter Tuning for Motorised Stage at DLS

There are more than three thousand motorised stages at Diamond Light Source (DLS) with a substantial fraction of these operating with PID controllers. Tuning these stages can be very time consuming and relies heavily on the experience of the motion .

$$u_k = \underbrace{K_p e_k + K_i \sum_{j=1}^{k} e_j - K_v \Delta x_k}_{\text{feedback}} + \underbrace{K_{\text{vff}}\Delta r_k + K_{\text{aff}}\Delta^2 r_k + K_{\text{pff1}} r_k + K_{\text{pff0}}|r_k|r_k}_{\text{feedforward}}.$$

$$\boldsymbol{\theta} = [\texttt{kp}, \texttt{kv}, \texttt{ki}, \texttt{kvff}, \texttt{kaff}, \texttt{kpff1}, \texttt{kpff0}].$$

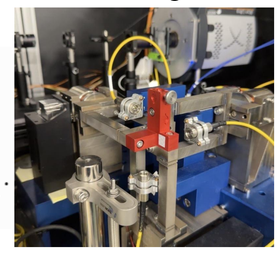


Figure 8: Prototype Delta Robot Sample Scanner System

The good tuning locates in highly concentrated areas, so we proposed to use **trust region Bayesian optimisation with input warping** to solve the problem.

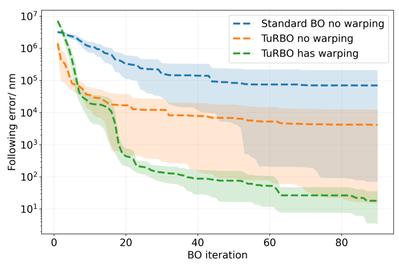



Figure 9: Best stage following error comparisons

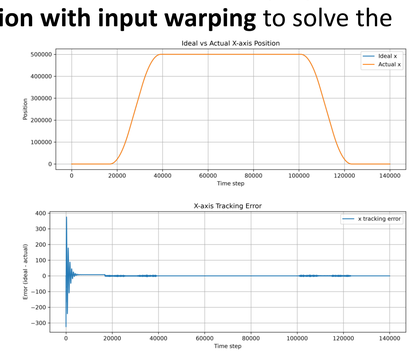



Figure 10: Best tuning's performance

## Tuning the ISIS LEBT with Bayesian Optimisation

ISIS runs in ~12-week cycles, interrupted by one maintenance day. In maintenance periods, operators are responsible for tuning the machine after each shut down to ensure that the required currents are being delivered to the synchrotron. Currently this is done manually, but this work highlights the first attempts at tuning the accelerator automatically using EPICS and Bayesian Optimisation.

- Ion source performance can be recovered by tuning parameters in the Low Energy Beam Transport (LEBT):

**Variables:**
- 3x LEBT Dipoles
- 3x LEBT Solenoids strength

**Objectives:**
- Maximise Pre-Injector efficiency

**Constraints:**
- IHT5 above ~23mA

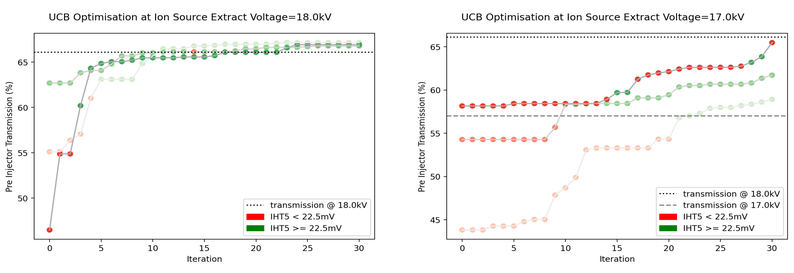



Figure 11: Nov 2024 – Reduced Ion Source Extract Voltage

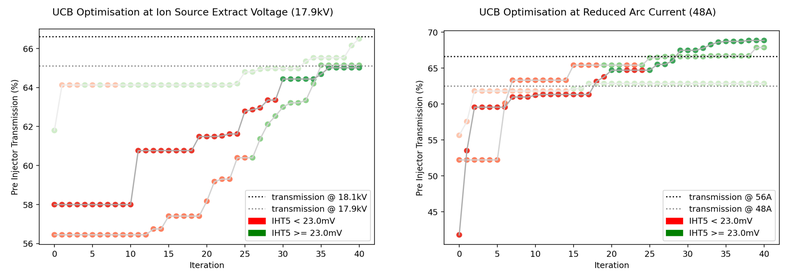



Figure 12: Dec 2024 – Reduced Ion Source Extract Voltage & Reduced Arc Current

- BO can reproduce the performance achieved with manual tuning. In some cases, it was able to outperform manual tuning.
- BO was able to achieve these results in runs lasting around 20 mins. This provides greater consistency and reliability in time than manual tuning, which is reported to vary from 15 mins to a couple of hours.

## To discuss

How can we **optimise safely** with only a **few** evaluations?
How can **simulation-trained models** adapt to real machines?
How can we learn joint **6D beam distributions** from limited simulations?