

Using Sparse Gaussian Processes for Predicting Robust Inertial Confinement Fusion Implosion Yields

Peter Hatfield*, Steven Rose^{†*}, Robbie Scott[‡], Ibrahim Almosallam[§], Stephen Roberts[¶], Matt Jarvis^{|| **}

* Clarendon Laboratory, University of Oxford, Parks Road, Oxford OX1 3PU, UK

[†]Blackett Laboratory, Imperial College, London SW7 2AZ, UK

[‡]Central Laser Facility, STFC Rutherford Appleton Laboratory, Harwell Oxford, Didcot OX11 0QX, UK

[§]King Abdulaziz City for Science and Technology, Riyadh 11442, Saudi Arabia

[¶]Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ, UK

^{||}Astrophysics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford, OX1 3RH, UK

^{**}Department of Physics, University of the Western Cape, Bellville 7535, South Africa

Abstract—Here we present the application of an advanced Sparse Gaussian Process based machine learning algorithm to the challenge of predicting the yields of inertial confinement fusion (ICF) experiments. The algorithm is used to investigate the parameter space of an extremely robust ICF design for the National Ignition Facility, the ‘Simplest Design’; deuterium-tritium gas in a plastic ablator with a Gaussian, Planckian drive. In particular we show that i) GPz has the potential to decompose uncertainty on predictions into uncertainty from lack of data and shot-to-shot variation, ii) permits the incorporation of science-goal specific *cost-sensitive learning* e.g. focussing on the high-yield parts of parameter space and iii) is very fast and effective in high dimensions.

I. INTRODUCTION

Inertial confinement fusion (ICF), in which deuterium-tritium (DT) fuel is compressed to temperatures and densities exceeding that found in the Sun, is one of the main potential pathways to nuclear fusion as a source of energy. The world’s leading ICF facility is the National Ignition Facility (NIF) at Lawrence Livermore National Laboratory (LLNL). NIF uses *indirect drive* in which lasers first hit a hohlraum (typically a gold can), which then emits a thermal radiation field which drives the implosion, in contrast to *direct drive* ICF in which laser beams themselves drive the implosion.

Although huge progress has been made, NIF has been unable to reach the yields originally hoped for. This has led to an interest in using modern machine learning techniques to produce new designs and quantify uncertainties on predictions. [1] presented an early ensemble of thousands of ICF implosions, and used Gaussian Processes (GPs) to model the parameter space. [2] developed a novel neural network (NN) machine learning algorithm called DJINN (‘deep jointly-informed neural networks’) that used random forests to construct appropriate neural network architectures with relatively little human input, which was used in [3] to identify a novel non-spherically symmetric design for NIF. [4] used *transfer*

*learning*¹ with DJINN to update the machine learning predictions based on seeing real experiments. Finally [5] used genetic algorithms in an even wider parameter space to produce ICF designs almost from scratch.

In this work we use a GP based machine learning algorithm GPz, [6], [7] to build surrogate models of a robust ICF design. GPz has i) a flexible *cost-sensitive learning* feature that permits optimisation for the specific science goal at hand, ii) models *heteroscedastic noise*, permitting the uncertainty to be decomposed into uncertainty from shot-to-shot variation and uncertainty from lack of data, and iii) uses a sparse framework that lets it run quickly even in high dimensions. We compare it to DJINN’s performance on the same data, and discuss future approaches to building ICF surrogates.

II. PROBLEM FORMULATION AND METHODOLOGY

1) *The Simplest Design*: The point design for NIF is an indirect drive design, with a capsule of DT gas inside a ablator shell (plastic, CH or a few other options) with a thin layer of solid DT ice on the inside. In the conventional implosion the ablation compresses the DT gas to a relatively high convergence ratio (the ratio of the DT radius at the start of the experiment to the radius at peak compression), and extremely high temperatures. If certain criteria are met, sufficient nuclear reactions take place that alpha heating (heating of the DT plasma by alpha particles produced in the nuclear reactions) dominates, and the gas ignites. This then starts a burn wave propagating through the DT ice layer, from which most of the neutron yield comes from. Most designs use a series of shockwaves to fine-tune the implosion, with several variants studied.

The Simplest Design removes some of the more complex/challenging aspects of ICF where the physics is more uncertain: i) delicate pulse timing, ii) the burn wave through the DT ice and iii) high convergence ratios. This design simply

¹Transfer learning is a family of machine learning methods that seek to let an algorithm apply information/knowledge gained from one problem, to the task of solving a similar but different problem.

has DT gas (but of a much higher density), with a CH ablator, and a Gaussian drive (sketch of capsule design shown in figure 1). It is unlikely that the Simplest Design will be able to lead to ignition at NIF, but it does represent a pathway to very predictable robust 1D implosions. We use a thermal x-ray drive with a Gaussian temperature time dependence.

We parametrise the Simplest Design with 5 parameters and investigate the parameter space within the following limits:

- $50\text{eV} < T_{\text{peak}} < 400\text{eV}$ - the peak temperature of the drive
- $0.01\text{ns} < \sigma < 5\text{ns}$ - the standard deviation of the time dependence of the temperature of the drive
- $0.1\text{mm} < r_1 < 1.5\text{mm}$ - the radius of the DT gas
- $0.05\text{mm} < r_2 < 1\text{mm}$ - the radius of the CH ablator
- $10\text{mg/cc} < \rho < 200\text{mg/cc}$ - the density of the DT gas (mg/cc)

N.B. the gas fill is much higher than typical for ICF (normally closer to $\sim 0.1\text{mg/cc}$) - meaning that the implosions have a much lower convergence ratio, but that we would not expect the design to be able to reach yields comparable to that which are in principle possible with the point design on NIF. This is because the point design has low density gas that is (comparatively) easily compressed. At peak compression, the hotspot (the low density gas that has been compressed) and DT ice are approximately isobaric, but the density is highly non-uniform - resulting in a small region of high temperature in the hotspot, with low temperature in the ice. The temperatures in the low density hotspot are in principle sufficient to initiate significant burn within the hotspot (which hopefully drives a burn wave through the DT ice, to give extremely high yields). For an equivalent DT mass arranged uniformly (as a function of radius), the energy required to compress it sufficiently to heat it to temperatures high enough to initiate fusion would be far higher than that available on NIF. The design space is chosen to roughly correspond to what is achievable with a gold hohlraum on NIF; the main restrictions are total capsule radius and total energy in drive.

The Simplest Design shares some design philosophy with the 2-shock design of [8] and [9]. The design of shot N161004 described in [9] still has a DT ice layer like more typical designs, but has a relatively high gas fill density of 5mg/cc (much closer in log-space to the densities we consider than to the point design), and a correspondingly more modest convergence ratio. They also have a relatively simple 2-shock drive. These features are experimentally observed to make the implosion much more 1D, and much closer to the predictions of simulations. The Simplest Design goes a step further, with an even higher gas fill, even lower convergence ratio and an even simpler 1-shock drive - and thus also has the potential to give a robust 1D implosion that closely matches simulation.

2) *Data*: Typically machine learning methods give better predictions in parts of parameter space with lots of data, and vice-versa. We use here a Monte Carlo sampling, but rather than sampling uniformly in the parameter under consideration,

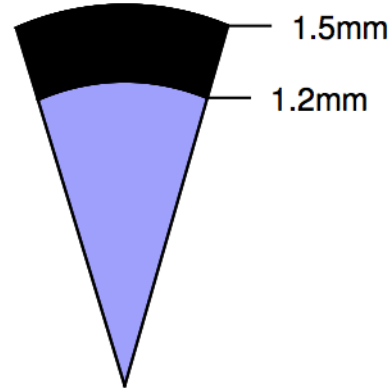


Fig. 1. Example Simplest Design diagram; black shows CH, pale blue DT gas. Scale and of proportions of the 4×10^{15} design of Section III-1 shown, labels are radii from the centre of the capsule.

points are sampled from a multi-variate Gaussian² centred on where a preliminary estimate for an optimal design was³. [3] conversely use latin hypercube sampling (LHS) to achieve good coverage of the parameter space, which may be a valuable alternative sampling approach for future work. This base design was $T_{\text{peak}} = 300\text{eV}$, $\sigma = 2\text{ns}$, $r_1 = 1\text{mm}$, $r_2 = 0.15\text{mm}$ and $\rho = 55\text{mg/cc}$. 5000 simulations were run and we divided it into 30% training, 30% validation and 40% testing data. The training data is used to infer the large number of parameters that make up the machine learning predictive model, the validation data is used to infer model hyper-parameters (essentially the complexity of the model), and the test data is held back to use as a final test of performance. We simulate shot-to-shot variation by artificially adding Gaussian scatter to the calculated log-yields⁴ in number of neutrons. This is chosen to be ρ dependent; the scatter is chosen to have standard deviation $0.1 \times \rho / (55\text{mg/cc})$ dex.

Our implosion simulations are performed using the HYADES[10] radiation-hydrodynamics simulation code, which is well benchmarked and used widely for the simulation of inertial fusion and high energy density physics applications[11], [12], [13], [14]. HYADES models hydrodynamics within a Lagrangian framework. Electron and ion thermal energy transport is described by a flux-limited Spitzer-Härm thermal conductivity model. Equations-of-state either use the Los Alamos SESAME tables[15] or QEOS[16]. Ionization levels come from a hydrogenic average-atom model or self-consistently

²In log-space

³Optimal sampling in the parameter space likely depends on the end goal e.g. if planning to eventually implement transfer learning it may still be valuable to sample the low-yield parts of parameter space.

⁴Log-yields always measured in logarithms of base ten of number of neutrons

from QEOS. Radiation transport uses the multi-group diffusion approximation; here we use 60 groups. A 1D spherically symmetric geometry is employed.

Neither laser-plasma interactions nor hohlraum physics are modelled; instead we use an incoming x-ray drive imposed at the outside of the grid. The capsules are modelled within a 5mm helium container. Simulations start in cryogenic conditions at $1.551 \times 10^{-3} \text{eV} = 18\text{K}$. The HYADES runs were performed on SCARF at the Central Laser Facility at Rutherford Appleton Laboratory using 500 CPUs. The modelling used here is likely appropriate for the design investigated here⁵, but the next level of sophistication of modelling would be to a) move to 2D/3D simulations (e.g. so Rayleigh-Taylor instabilities are incorporated) and b) start to include laser/hohlraum physics and non-Planckian drives (e.g. simulate the laser light being converted to x-rays, rather than just assuming an incoming x-ray drive). See [17], [18] for overviews of the wide range of physics involved in ICF, and associated issues concerning which physics to include in simulations etc.

3) *GPz*: GPz is a machine learning regression algorithm originally developed for the problem in astrophysics of calculating the photometric redshifts of galaxies; the details of the algorithm and the key developments in machine learning (ML) theory are described in [6], [7], and applied to photometric redshift calculation in [19], [20], and to orbital dynamics in [21]. The algorithm is ‘GP’ based; a Gaussian Process is a stochastic process with a random variable defined at each point in a space of interest, such that any finite subset of the random variables has a multivariate normal distribution (equivalently any linear combination of random variables from different points has a normal distribution). A GP is essentially an un-parametrised continuous function defined everywhere with Gaussian uncertainties. A GP based ML algorithm will typically take a set of data over the parameter space of interest and in some sense try and find the function in the function space defined by the Gaussian process that was most likely to have produced the data - and then make predictions for other parts of parameter space based on that.

GPz is a sparse Gaussian process based code, a fast and a scalable approximation of a full Gaussian Process [22], with the added feature of being able to produce input-dependent variance estimations (heteroscedastic noise). For the full details of the algorithm see [6], [7], [23], but we summarise the main details here. The model assumes that the probability of the observing a target variable y given the vector input \mathbf{x} is $p(y|\mathbf{x}) = \mathcal{N}(\mu(\mathbf{x}), \sigma(\mathbf{x})^2)$. The mean function, $\mu(\mathbf{x})$, and the variance function $\sigma(\mathbf{x})$ are both linear combinations⁶ of ‘basis

functions’ that take the following form:

$$\mu(\mathbf{x}) = \sum_{i=1}^m \phi_i(\mathbf{x}) w_i \quad (1)$$

$$\sigma(\mathbf{x}) = \exp\left(\sum_{i=1}^m \phi_i(\mathbf{x}) v_i\right) \quad (2)$$

where $\{\phi_i(\mathbf{x}), w_i, v_i\}_{i=1}^m$ are sets of m basis functions and their associated weights respectively⁷. Basis function models (BFM), for specific classes of basis functions such as the squared exponential, have the advantage of being universal approximators, i.e. there exist a function of that form that can approximate any function, with mild assumptions, to any desired degree of accuracy; i.e. a one size fits all function. BFM are a form of sparse Gaussian Processes [23]. The most general form of the squared exponential is:

$$\phi_i(\mathbf{x}) = \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^T \Lambda_i^{-1}(\mathbf{x} - \boldsymbol{\mu}_i)\right) \quad (3)$$

The goal of GPz essentially is to find the optimal parameters, $\{\boldsymbol{\mu}_i, \Lambda_i, w_i, v_i\}_{i=1}^m$ such that the mean and variance functions, $\mu(\mathbf{x})$ and $\sigma(\mathbf{x})$, are the most likely functions to have generated the data - using Bayesian inference. These optimal parameters are found with a Limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm [24] (a hill-climbing optimisation algorithm for differentiable non-linear problems). During the training stage of the algorithm GPz is inferring the locations and spreads that describe the basis functions; during the validation stage it is inferring the appropriate complexity of the model, essentially how many basis functions to use.

The key features introduced by GPz include a) implementation of a sparse GP framework, allowing the algorithm to run in $O(nm^2)$ instead of $O(n^3)$ (where n is the number of samples in the data and m is the number of basis functions), b) a ‘cost sensitive learning’ framework where the algorithm can be tailored for the precise science goal, and c) properly accounts for uncertainty contributions from both variance in the data as well as uncertainty from lack of data in a given part of parameter space (by marginalising over the functions supported by the GP that could have produced the data).

Unless otherwise stated, we use the settings in Table I (see [6], [7] for precise definitions and interpretations). GPz requires very little fine-tuning. The most important parameter is m , the number of basis functions. A higher m corresponds to higher model complexity and longer training times. Figure 2 shows algorithm performance as a function of m ; best performance is achieved for $m \approx 10 - 100$, in line with findings in [6], [7].

GPz permits custom cost-sensitive learning e.g. a specific science goal to specified to the algorithm. In the application of calculating galaxy redshifts this typically might be something like specifying that low-redshift galaxies are a low priority. In the ICF case, we can also set particular parts of parameter space of interest e.g. areas around a cliff, parts of parameter

⁵It is in some sense hard to definitively know this is the case until a real shot has been performed, so this statement must remain conditional without experimental data. However the aforementioned design of [9] finds that 1D models are in reasonable agreement with the data; and the implosions described here should be ‘even more 1D’.

⁶N.B. This is a different sense of linear combination to that which can be used in the definition of a Gaussian Process. That definition of a Gaussian Process involves the addition of random variables at different points; the linear combination discussed here is an addition of multivariable functions to build up a representation of $\mu(\mathbf{x})$ and $\sigma(\mathbf{x})$.

⁷In practice GPz actually uses $\sigma^2 = \exp(\sum_{i=1}^m \phi_i(\mathbf{x}) v_i + b)$ where addition of the bias term ‘ b ’ is used for practical reasons but is not required theoretically.

TABLE I
PARAMETER SETTING OF GPZ.

| Parameter | Value | Description |
|-----------------|-------|---|
| m | 100 | Number of basis functions (the ϕ_i); complexity of GP |
| maxIter | 500 | Maximum number of iterations (comparisons with the validation data) permitted |
| maxAttempts | 50 | Maximum number of iterations to attempt if there is no progress on the validation set |
| method | GPVC | Type of bespoke covariances (the Λ_i in equation 3) used on each basis function (see [7] for the different options) |
| normalize | True | Pre-process the input by subtracting the means and dividing by the standard deviations |
| joint | True | Jointly learn a prior linear mean-function (learn the function means and variances jointly) |
| heteroscedastic | True | Model noise as well as point estimates |

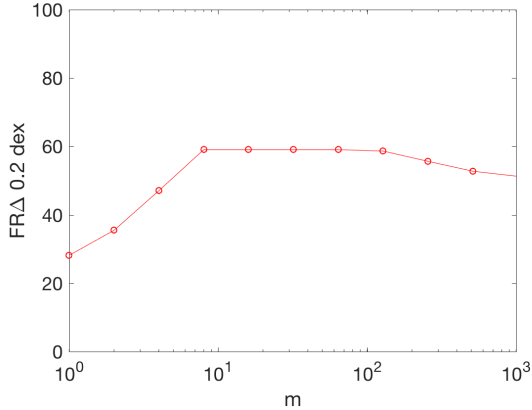


Fig. 2. The performance of the algorithm as a function of the number of basis functions used. The y-axis is the fraction of test data within 0.2 dex of the prediction.

space that are possible with a particular facility etc. In this work we will use the weighting $w = \sqrt{Y}$ (where Y is the simulated yield in number of neutrons), up-weighting the cost of getting predictions wrong in the high-yield part of parameter space, and assuming that we are relatively insensitive to getting predictions wrong in low-yield parts of parameter space. The choice of $w = \sqrt{Y}$ was motivated by requiring a function that was a) monotonically increasing with Y (so that higher yield parameter space has a higher weight than lower yield parameter space), b) a power law (so that the weighting is invariant under rescaling of Y) and c) has a ‘reasonable’ dynamical range for the design space considered (here yield spans $\sim 10^{10} - 10^{16}$, so the ratio of weightings of different parts of parameter space can reach up to $\sim \sqrt{\frac{10^{16}}{10^{10}}} = 1000$).

The 5D space considered here is of relatively modest dimensionality, but GPz has been shown to be effective and fast running in $\gtrsim 10D$, see [19] and subsection III-5.

We also trialled our data with DJINN [2]⁸. The DJINN solution to difficulties in designing neural network structure is to use a novel mapping from decision tree to network structure, giving a very user-friendly algorithm that works with very little human input in a wide variety of circumstances. Parameters used for DJINN used are shown in table II, with choices motivated by [2] (improved performance may be achievable with further parameter fine-tuning).

⁸Taken from the version on <https://github.com/LLNL/DJINN>

III. RESULTS

1) *Predictions and Optimal Design*: Figure 3 shows the yield predicted by the machine learning algorithms for the test data, compared with the Hyades yields. GPz is largely able to correctly predict the yield for most of the test data, with reasonably realistic uncertainties. Figure 4 shows the yield as a function of two of the parameters and the corresponding uncertainty on the predictions. The part of parameter space with maximal yield within design constraints (e.g. what is feasible for a given facility) can easily be extracted, or a similar stability test as in [3] can be used to find the design with the best combination of yield and stability etc. For example, say we fix the drive at $T_{\text{peak}} = 300\text{eV}$ and $\sigma = 1.5\text{ns}$, and restrict interest to capsules smaller than 1.5mm (e.g. $r_1 + r_2 < 1.5\text{mm}$). The capsule with highest predicted yield is easily found to be $r_1 \approx 1.2\text{mm}$, $r_2 \approx 0.3\text{mm}$ and $\rho \approx 25\text{mg/cc}$, giving a yield of $Y \sim 4 \times 10^{15}$. However suppose we were only interested in making sure that the design robustly had a yield above $Y = 10^{15}$, and instead wanted to minimise the capsule radius. We define $P(Y > 10^{15} | \delta = 0.2)$ as the fraction of designs that still have a yield above $Y = 10^{15}$ when each parameter is perturbed by an amount sampled from a Gaussian with a standard deviation of 20% (c.f. [3]). We can find the design that minimises $r_1 + r_2$ with $P(Y > 10^{15} | \delta = 0.2) > 0.9$, leading to a slightly different design, $r_1 \approx 0.95\text{mm}$, $r_2 \approx 0.3\text{mm}$ and $\rho \approx 25\text{mg/cc}$ ($Y \sim 2.5 \times 10^{15}$).

2) *Cost-Sensitive Learning*: Figure 3 shows a comparison of GPz $w = 1$ (‘normal’ in [6]) and using cost-sensitive learning ($w = \sqrt{Y}$). Figure 5 and 6 show the bias (mean of $Y_{\text{test}} - Y_{\text{prediction}}$) and root-mean-squared-error (RMSE) on the predictions, illustrating that performance in higher yield parts of parameter space ($\log Y \sim 15 - 17$) is indeed improved at the cost of performance in lower yield regions. In general CSL can be linked to specific science goal of a study e.g. for the goal of achieving ignition at the NIF we probably have particular interest in having low RMSE close to cliff edges, but are relatively insensitive to RMSE or bias both far above and far below this boundary. We would note that CSL is a method that can in some circumstances obtain slightly better statistical properties in certain parts of parameter space [20]; but it cannot extract information from the data that simply isn’t there (e.g. a really extreme weighting scheme will still fail to improve predictions in parts of parameter space with almost no data).

TABLE II
PARAMETER SETTING OF DJINN.

| Parameter | Value | Description |
|-------------|-------|---------------------------------------|
| ntrees | 10 | Number of neural nets in ensemble |
| maxdepth | 5 | Maximum depth of tree |
| dropoutkeep | 0.95 | Dropout |
| niters | 100 | Number of times network evaluated for |

3) *Comparison with DJINN*: We also applied (Bayesian) DJINN to our data, also shown in Figures 3, 5 and 6. The intention is not to do a rigorous code comparison, as the codes were developed for different machine learning goals and it is non-trivial to directly compare model complexity, but simply to illustrate that Gaussian processes are comparably viable tools for building ICF surrogates.

DJINN also performs well in producing a surrogate model. The algorithm trained and predicted in ~ 80 s and ~ 2 s respectively, compared with ~ 5 s and ~ 0.02 s for GPz on a laptop⁹.

Some authors have found that the best results are achieved using a committee of a variety of machine learning methods, so it is possible best results could be achieved using a combination of neural nets and Gaussian processes with an ensemble averaging [25] or hierarchical Bayesian approach [20]. This would also start a move towards non-Gaussian pdfs, as predictions near cliffs are likely to be multi-modal (an experiment either ignites or it doesn't). Greater precision in key parts of parameter space can of course also be achieved by doing more simulations in that part of parameter space, but that requires advance knowledge of which part of parameter space is interesting. Future work could couple the process of sampling parameter space and building the surrogate. It might also be interesting to consider how CSL-like methods might be implemented within DJINN e.g. loss-calibrated learning as per [26].

4) *Uncertainty Decomposition*: The right subplot of Figure 4 shows $\sqrt{\nu/\beta_\star^{-1}}$, the ratio of uncertainty from lack of data to uncertainty from intrinsic variation. ν is the variance from lack of data, defined as $\nu = \phi(\mathbf{x})\Sigma^{-1}\phi(\mathbf{x})^T$, where ϕ is the vector of non-linear basis functions that the prediction mapping is constructed from, \mathbf{x} is the test data, and Σ is the covariance matrix of uncertainty on the weights applied to the basis functions when constructing the posterior mean. It is essentially the uncertainty on the $\mu(x)$ in equation 1. β_\star^{-1} is the input-dependent noise variance. It is essentially the $\sigma(\mathbf{x})^2$ in equation 2. The total variance is $\sigma_{\text{Total}}^2 = \nu + \beta_\star^{-1}$. See Equations 3.13 and 5.10 in [7] for a more in depth explanation of the calculation and interpretation of these quantities, and [27] for a more general background. The plot illustrates that GPz correctly identifies that for $\rho < 10$ mg/cc more of the uncertainty is coming from lack of data rather than shot-to-shot variation (as there were no simulations done in that part of parameter space). This shows that GPz can help understand what is the dominant source of uncertainty in different

parts of parameter space. This uncertainty decomposition does however come with the caveat that uncertainty is still likely underestimated at the edge of the domain, and anywhere far from data, due to the use of only a finite number of basis functions. Estimates of shot-to-shot variation can also be incorporated into design optimisation e.g. suppose we wanted to find the design with a yield above $Y = 10^{15}$, but with minimal shot to shot variation, we can find the design with smallest β_\star^{-1} that meets this yield criterion, which is $r_1 \approx 1.05$ mm, $r_2 \approx 0.15$ mm and $\rho \approx 15$ mg/cc (for the same drive as in Section III-1).

The shot-to-shot variation considered here is not quite identical to the real problem in ICF; here we added scatter to the thousands of simulations, whereas more realistically we might have a large number of simulations with no scatter/uncertainty, and then just a few experiments with some shot-to-shot variation. Nonetheless an approach similar to that detailed here could be effective in folding shot-to-shot variation into surrogate building. For example, one might simulate a large number of shots varying both parameters the designer controls (e.g. capsule shell thickness) as well as ones they don't (e.g. imperfections on the surface of the shell). The outputs from these simulations could then be used to build a 'noisy' surrogate as a function of the controlled parameters. Another useful feature is that understanding the shot-to-shot variation permits the model to say that there is substantial amount of uncertainty in the prediction of the yield for an individual shot, but also to say that there is no need to do any further experiments in that part of parameter space as all the uncertainty is from shot-to-shot variation. Features of GPz in development that could be useful for ICF surrogate building in future include i) incorporating uncertainty on predictions due to uncertainty on input parameters (e.g. if there are error bars associated with the gas fill densities and shell thicknesses etc.) and ii) coping with incomplete data (e.g. training on data where a subset of the shots don't have some of the experimental properties recorded).

5) *Scalability*: Thus far we have explored the applicability of GPz to building ICF surrogates in the context of the Simplest Design, as a possible pathway to extremely robust implosions. Other ICF design space data sets in general however might have more complex features, in particular i) sharp ignition cliffs, ii) higher dimensionality and iii) larger numbers of simulations. Future work will investigate more fully the performance of GPz in more complex design spaces. Here however we briefly consider a simple analytic model, to see if GPz is likely to have the capacity to scale well (c.f.

⁹16 GB RAM, 3.1 GHz Intel Core i5

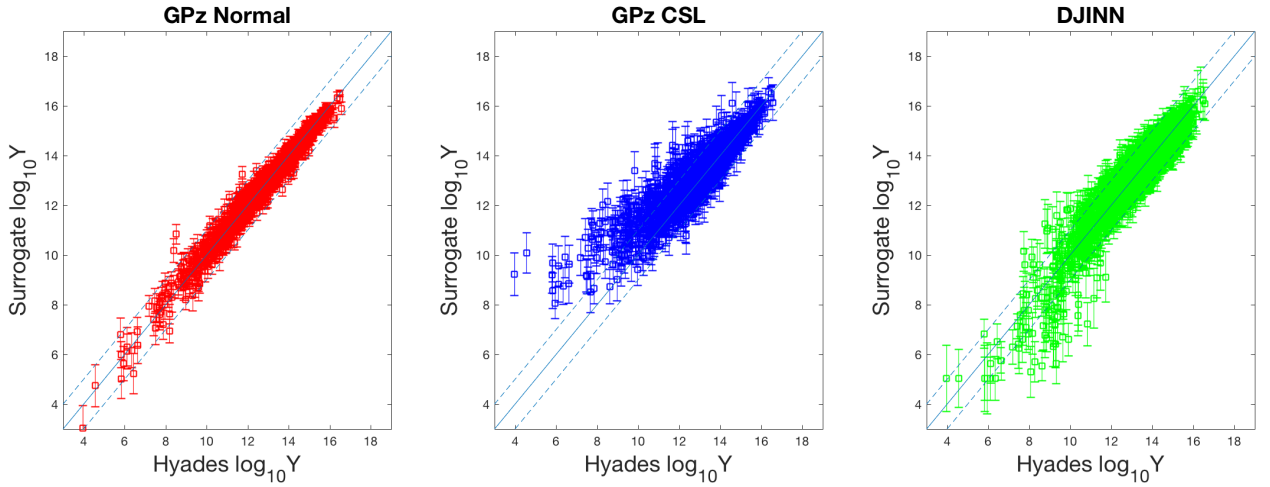


Fig. 3. Surrogate yield and Hyades yield compared for different surrogate set-ups.. GPz-Normal is shown in red (left), GPz-CSL in blue (centre), DJINN in green (right). The diagonal filled line shows equality, and the dashed lines show one dex discrepancies

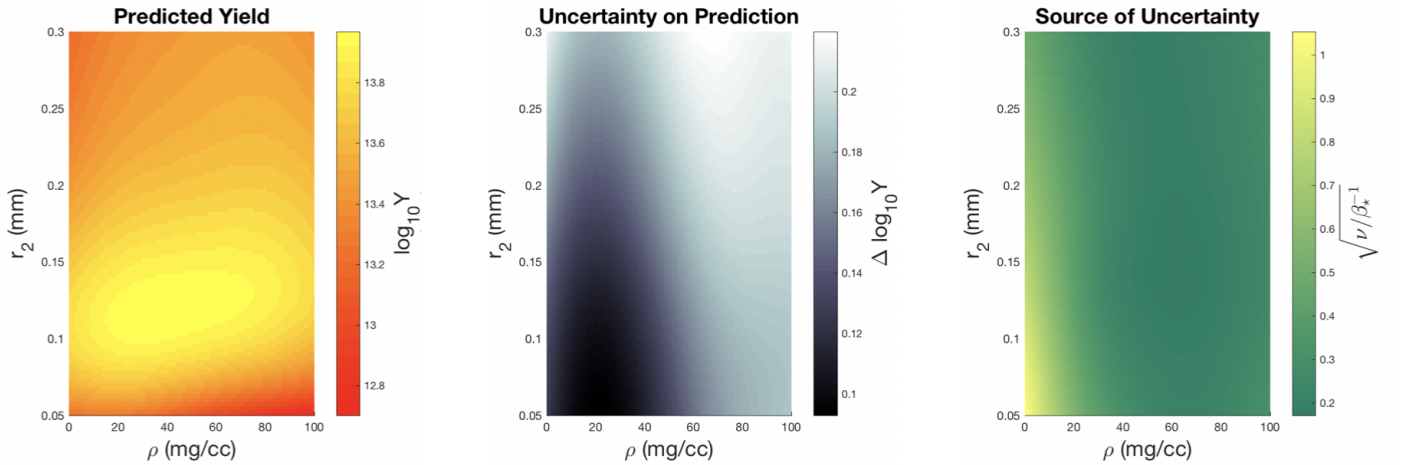


Fig. 4. A 2D slice through the 5D parameter space (with the other three parameters set to $r_1 = 1\text{mm}$, $T_{\text{peak}} = 300\text{eV}$, $\sigma = 0.5\text{ns}$), showing the surrogates yield (left), the uncertainty on the prediction (centre) and the uncertainty from lack of data divided by uncertainty from shot-to-shot variation (right).

the simple analytic model considered in Section 2 of [4]). We consider a 10-dimensional hypercube with sides going from 0 to 1, and sample 100,000 points randomly in this volume (each dimension sampled uniformly independently). We calculate a mock-yield of $Y = r^5(1 + 100000 \times (1 + \text{erf}(10 \times (r - 2))))$, where r is the radius from the origin in this 10D space (the error function giving an ignition-like cliff). We then train GPz (with $m = 100$ basis functions; same settings as Table I used except heteroscedastic=false) and DJINN on 90% of this data (90,000 points) and test on the other 10% (10,000 points); results shown in figure 7. Training time was 415s for GPz and 450s for DJINN. Predicting time was 0.1s for GPz and 7s for DJINN. Both capture the design space well, with each doing slightly better or worse at different aspects of the prediction e.g. GPz has a few outliers which DJINN doesn't, but it is less biased for high and low r . Which statistical properties are

most desirable are in general likely to depend on the specific science goal at hand. In summary it appears that GPz should be able to perform well, and run in reasonable amounts of time, for higher dimensional spaces, with larger quantities of data, when there are steep cliffs in the design space.

IV. CONCLUSIONS

Our results show that Gaussian processes have the potential to be useful ICF surrogates, and that in particular GPz is shown to be effective for the task as an easy-to-use algorithm that can cope with huge amounts of data in a high number of dimensions, with realistic uncertainties. In particular GPz may be useful for modelling shot-to-shot variation alongside uncertainty from lack of data in integrated experiments.

Future work will seek to combine experiment and simulation, either through transfer learning as per [4], or possibly

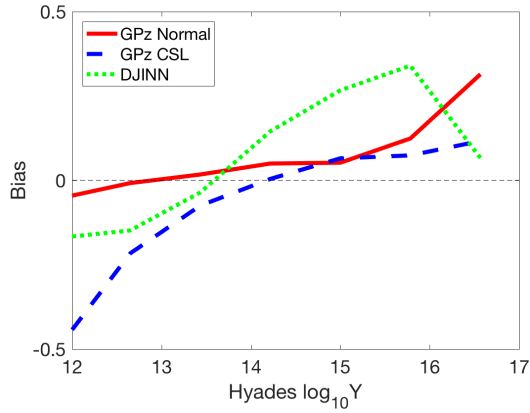


Fig. 5. The bias on the predictions, GPz-Normal shown in red (solid line), GPz-CSL in blue (dashed), DJINN in green (dotted).

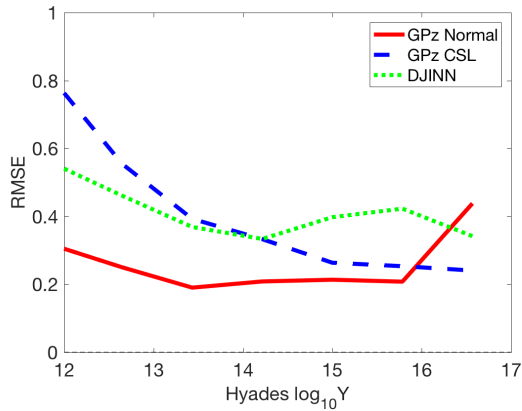


Fig. 6. The RMSE on the predictions, GPz-Normal shown in red (solid line), GPz-CSL in blue (dashed), DJINN in green (dotted).

through scaling the surrogate as per [28], taking care to understand biases induced by differences between the target distribution and the training set.

Key findings:

- GPz can make highly effective surrogate models for predicting the outcomes ICF experiments
- Cost-sensitive learning can help improve the statistical properties of predictions in the most important parts of parameter space
- It is possible to distinguish between uncertainty from shot-to-shot variation and uncertainty from lack of data
- The Simplest Design should be able to produce yields of order 10^{15} neutrons extremely robustly on NIF

ACKNOWLEDGMENT

P.W.H. and S.J.R. acknowledge funding from the Engineering and Physical Sciences Research Council. I.A.A acknowledges the support of King Abdulaziz City for Science and Technology. Many thanks to Warren Garbett at the Atomic Weapons Establishment, Aldermaston, Kris McGlinchey and Jeremy Chittenden at Imperial College, London, Jim Gaffney,

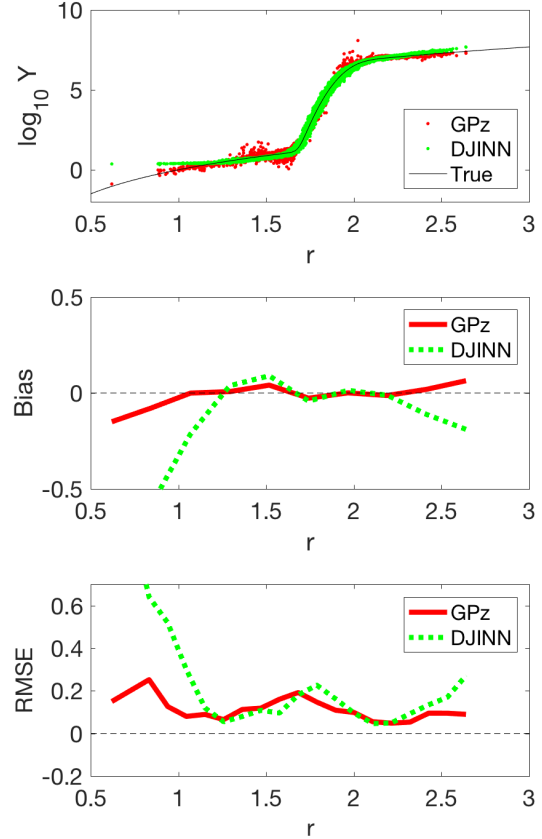


Fig. 7. Top plot: GPz (red) and DJINN (green) predictions as a function of radius for our analytic model, with the analytic result shown with a black line. NB. these are results in 10D; we plot the 1D r and Y values of the test data for ease of visualisation. Middle plot: RMSE on the predictions, GPz-Normal shown in red (solid line), DJINN in green (dotted). Bottom plot: RMSE on the predictions, GPz-Normal shown in red (solid line), DJINN in green (dotted).

Luc Peterson, Kelli Humbird and Brian Spears at Lawrence Livermore National Laboratory, and Chris Bridge at Massachusetts General Hospital for useful discussions.

REFERENCES

- [1] R. Nora, J. L. Peterson, B. K. Spears, J. E. Field, and S. Brandon, "Ensemble simulations of inertial confinement fusion implosions," *Statistical Analysis and Data Mining: The ASA Data Science Journal*, vol. 10, no. 4, pp. 230–237, aug 2017. [Online]. Available: <http://doi.wiley.com/10.1002/sam.11344>
- [2] K. D. Humbird, J. L. Peterson, and R. G. McClarren, "Deep Jointly-Informed Neural Networks," jul 2017. [Online]. Available: <http://arxiv.org/abs/1707.00784>
- [3] J. L. Peterson, K. D. Humbird, J. E. Field, S. T. Brandon, S. H. Langer, R. C. Nora, B. K. Spears, and P. T. Springer, "Zonal flow generation in inertial confinement fusion implosions," *Physics of Plasmas*, vol. 24, no. 3, p. 032702, mar 2017. [Online]. Available: <http://aip.scitation.org/doi/10.1063/1.4977912>
- [4] K. D. Humbird, J. L. Peterson, and R. G. McClarren, "Transfer learning to model inertial confinement fusion experiments," dec 2018. [Online]. Available: <http://arxiv.org/abs/1812.06055>
- [5] P. W. Hatfield, S. J. Rose, and R. H. H. Scott, "The blind implosion-maker: Automated inertial confinement fusion experiment design," *Physics of Plasmas*, vol. 26, no. 6, p. 062706, jun 2019. [Online]. Available: <http://aip.scitation.org/doi/10.1063/1.5091985>
- [6] I. A. Almosallam, S. N. Lindsay, M. J. Jarvis, and S. J. Roberts, "A sparse Gaussian process framework for photometric redshift estimation," *Monthly Notices of the Royal Astronomical Society*, vol. 455, no. 3, pp. 2387–2401, jan 2016. [Online]. Available: <http://arxiv.org/abs/1505.05489http://dx.doi.org/10.1093/mnras/stv2425https://academic.oup.com/mnras/article-lookup/doi/10.1093/mnras/stv2425>
- [7] I. A. Almosallam, M. J. Jarvis, and S. J. Roberts, "GPz: non-stationary sparse Gaussian processes for heteroscedastic uncertainty estimation in photometric redshifts," *Monthly Notices of the Royal Astronomical Society*, vol. 462, no. 1, pp. 726–739, oct 2016. [Online]. Available: <http://arxiv.org/abs/1604.03593http://dx.doi.org/10.1093/mnras/stw1618https://academic.oup.com/mnras/article-lookup/doi/10.1093/mnras/stw1618>
- [8] S. F. Khan, S. A. MacLaren, J. D. Salmonson, T. Ma, G. A. Kyrala, J. E. Pino, J. R. Rygg, J. E. Field, R. Tommasini, J. E. Ralph, D. P. Turnbull, A. J. Mackinnon, K. L. Baker, L. R. Benedetti, D. K. Bradley, P. M. Celliers, E. L. Dewald, T. R. Dittrich, L. Berzak Hopkins, N. Izumi, M. L. Kervin, J. L. Kline, S. R. Nagel, A. Pak, and R. E. Tipton, "Symmetry tuning of a near one-dimensional 2-shock platform for code validation at the National Ignition Facility," *Physics of Plasmas*, vol. 23, no. 4, p. 042708, apr 2016. [Online]. Available: <https://doi.org/10.1063/1.4947223http://aip.scitation.org/doi/10.1063/1.4947223>
- [9] S. A. MacLaren, L. P. Masse, C. E. Czajka, S. F. Khan, G. A. Kyrala, T. Ma, J. E. Ralph, J. D. Salmonson, B. Bachmann, L. R. Benedetti, S. D. Bhandarkar, P. A. Bradley, R. Hatarik, H. W. Herrmann, D. A. Mariscal, M. Millot, P. K. Patel, J. E. Pino, M. Ratledge, N. G. Rice, R. E. Tipton, R. Tommasini, and C. B. Yeaman, "A near one-dimensional indirectly driven implosion at convergence ratio 30," *Physics of Plasmas*, vol. 25, no. 5, p. 056311, may 2018. [Online]. Available: <https://doi.org/10.1063/1.5017976http://aip.scitation.org/doi/10.1063/1.5017976>
- [10] J. T. Larsen and S. M. Lane, "HYADESA plasma hydrodynamics code for dense plasma studies," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 51, no. 1-2, pp. 179–186, jan 1994. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0022407394900787http://linkinghub.elsevier.com/retrieve/pii/0022407394900787>
- [11] K. A. Keilty, E. P. Liang, T. Ditmire, B. A. Remington, K. Shigemori, and A. M. Rubenchik, "Modeling of Lasergenerated Radiative Blast Waves," *The Astrophysical Journal*, vol. 538, no. 2, pp. 645–652, aug 2000. [Online]. Available: <http://stacks.iop.org/0004-637X/538/i=2/a=645>
- [12] D. Leibbrandt, R. Drake, and J. Stone, "Zeus-2D Simulations of Laser-Driven Radiative Shock Experiments," in *High Energy Density Laboratory Astrophysics*. Berlin/Heidelberg: Springer-Verlag, 2005, pp. 273–276. [Online]. Available: http://link.springer.com/10.1007/1-4020-4162-4_39
- [13] K. L. Sequoia, M. S. Tillack, and H. A. Scott, "A comparison of Hyades and Cretin for modeling laser absorption in underdense plasmas," 2006. [Online]. Available: https://cer.ucsd.edu/{_}files/publications/UCSD-CER-06-09.pdf
- [14] R. H. H. Scott, F. Perez, J. J. Santos, C. P. Ridgers, J. R. Davies, K. L. Lancaster, S. D. Baton, P. Nicolai, R. M. G. M. Trines, A. R. Bell, S. Hulin, M. Tzoufras, S. J. Rose, and P. A. Norreys, "A study of fast electron energy transport in relativistically intense laser-plasma interactions with large density scalelengths," *Physics of Plasmas*, vol. 19, no. 5, p. 053104, may 2012. [Online]. Available: <http://aip.scitation.org/doi/10.1063/1.4714615>
- [15] S. X. Hu, B. Militzer, V. N. Goncharov, and S. Skupsky, "FPEOS: A First-Principles Equation of State Table of Deuterium for Inertial Confinement Fusion Applications," *Physical Review B*, vol. 84, Issue 22, id. 224109, vol. 84, no. 22, sep 2011. [Online]. Available: <http://arxiv.org/abs/1110.0001http://dx.doi.org/10.1103/PhysRevB.84.224109>
- [16] R. M. More, K. H. Warren, D. A. Young, and G. B. Zimmerman, "A new quotidian equation of state (QEOS) for hot dense matter," *Physics of Fluids*, vol. 31, no. 10, p. 3059, 1988. [Online]. Available: <https://aip.scitation.org/doi/10.1063/1.866963>
- [17] M. D. Rosen, "The physics issues that determine inertial confinement fusion target gain and driver requirements: A tutorial," *Physics of Plasmas*, vol. 6, no. 5, pp. 1690–1699, may 1999. [Online]. Available: <http://aip.scitation.org/doi/10.1063/1.873427>
- [18] A. Hamza, "Targets for Inertial Confinement Fusion," *Encyclopedia of Materials: Science and Technology*, pp. 1–11, jan 2005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B0080431526020428{#_}bib25
- [19] Z. Gomes, M. J. Jarvis, I. A. Almosallam, and S. J. Roberts, "Improving photometric redshift estimation using GPz: size information, post processing, and improved photometry," *Monthly Notices of the Royal Astronomical Society*, vol. 475, no. 1, pp. 331–342, mar 2018. [Online]. Available: <http://arxiv.org/abs/1712.02256http://dx.doi.org/10.1093/mnras/stx3187https://academic.oup.com/mnras/article/475/1/331/4718473>
- [20] K. J. Duncan, M. J. I. Brown, W. L. Williams, P. N. Best, V. Buat, D. Burgarella, M. J. Jarvis, K. Malek, S. J. Oliver, H. J. A. Rottgering, and D. J. B. Smith, "Photometric redshifts for the next generation of deep radio continuum surveys - I: Template fitting," *Monthly Notices of the Royal Astronomical Society, Volume 473, Issue 2, p.2655-2672*, vol. 473, pp. 2655–2672, sep 2017. [Online]. Available: <http://arxiv.org/abs/1709.09183http://dx.doi.org/10.1093/mnras/stx2536>
- [21] H. Peng and X. Bai, "Comparative evaluation of three machine learning algorithms on improving orbit prediction accuracy," *Astrodynamics*, pp. 1–19, aug 2019. [Online]. Available: <http://link.springer.com/10.1007/s42064-018-0055-4>
- [22] J. Nocedal, "Updating quasi-Newton matrices with limited storage," *Mathematics of Computation*, vol. 35, no. 151, pp. 773–773, sep 1980. [Online]. Available: <http://www.ams.org/jourcgi/jour-getitem?pii=S0025-5718-1980-0572855-7>
- [23] I. Almosallam, "Heteroscedastic Gaussian processes for uncertain and incomplete data," Ph.D. dissertation, University of Oxford, 2017. [Online]. Available: <https://ora.ox.ac.uk/objects/uuid:6a3b600d-5759-456a-b785-5f89cf4ede6d>
- [24] J. Nocedal, "Updating quasi-Newton matrices with limited storage," *Mathematics of Computation*, vol. 35, no. 151, pp. 773–773, sep 1980. [Online]. Available: <http://www.ams.org/jourcgi/jour-getitem?pii=S0025-5718-1980-0572855-7>
- [25] R. T. Clemen, "Combining forecasts: A review and annotated," *Tech. Rep.*, 1989. [Online]. Available: <https://pdfs.semanticscholar.org/7117/9279738b91df0520061b351cb3e0124a411c.pdf>
- [26] A. D. Cobb, S. J. Roberts, and Y. Gal, "Loss-Calibrated Approximate Inference in Bayesian Neural Networks," may 2018. [Online]. Available: <http://arxiv.org/abs/1805.03901>
- [27] C. M. Bishop, *Pattern recognition and machine learning*. Springer, 2006.
- [28] V. Gopalaswamy, R. Betti, J. P. Knauer, N. Luciani, D. Patel, K. M. Woo, A. Bose, I. V. Igumenshchev, E. M. Campbell, K. S. Anderson, K. A. Bauer, M. J. Bonino, D. Cao, A. R. Christopherson, G. W. Collins, T. J. B. Collins, J. R. Davies, J. A. Delettrez, D. H. Edgell, R. Epstein, C. J. Forrest, D. H. Froula, V. Y. Glebov, V. N. Goncharov, D. R. Harding, S. X. Hu, D. W. Jacobs-Perkins, R. T. Janezic, J. H. Kelly, O. M. Mannon, A. Maximov, F. J. Marshall, D. T. Michel, S. Miller, S. F. B. Morse, J. Palastro, J. Peebles, P. B. Radha, S. P. Regan, S. Sampat, T. C. Sangster, A. B. Sefkow, W. Seka, R. C. Shah, W. T. Shmyada, A. Shvydky, C. Stoeckl, A. A. Solodov, W. Theobald, J. D. Zuegel, M. G. Johnson, R. D. Petrasso, C. K. Li, and J. A. Frenje, "Tripled yield in direct-drive laser fusion through statistical modelling,"

Nature, vol. 565, no. 7741, pp. 581–586, jan 2019. [Online]. Available:
<http://www.nature.com/articles/s41586-019-0877-0>