Evaluating Imprecise Probabilities in Fusion Plasma Surrogates Using Conformal Prediction

Ander Gray ANDER.GRAY@UKAEA.UK
Vignesh Gopakumar VIGNESH.GOPAKUMAR@UKAEA.UK
William Hornsby WILLIAM.HORNSBY@UKAEA.UK
James Buchanan JAMES.BUCHANAN@UKAEA.UK
Stanislas Pamela STANISLAS.PAMELA@UKAEA.UK
Culham Science Centre, UKAEA, United Kingdom

Nuclear fusion plasma modelling Due to the infeasibility of iterative experimental design, next generation nuclear fusion reactors are being designed with a heavy emphasis on modelling and simulation. Fusion plasmas are highly non-linear and turbulent systems, and often require many hours on high performance computers to evaluate. Reliable surrogate models are therefore a must for integrated modelling, design optimisation, and uncertainty quantification in plasma simulation. Due to the cost, we must build surrogates with as few data points as possible, and machine learning models may struggle to fit accurately to these highly non-linear responses. We therefore seek reliable error estimation for surrogate modelling.

Conformal prediction Conformal prediction [4] presents itself as a simple method to add confidence intervals to any existing machine learning algorithm, or pre-trained machine learning model. This gives an estimate for the quality of the surrogate model with respect to its training data, capturing the surrogate's *model* uncertainty. Under some weak assumptions, the method works irrespective of the selected machine learning model, data-set, and sample size. The confidence intervals produced by conformal prediction are *valid*, or well-calibrated, in the sense that the true model response is guaranteed to be within the interval with at least a pre-selected probability.

Imprecise probability Recently Cella and Martin [2] have proposed an imprecise probabilistic interpretation of conformal prediction, through possibility theory, and describe how probability measures can be bounded in conformalised regression and classification problems. This also allows for the uncertainty estimates from conformal prediction to be further propagated [1], and makes a wider range of uncertainty quantification tasks accessible, for example reliability analysis.

A limitation of conformal prediction for surrogacy The confidence intervals produced by conformal prediction are valid only with respect to the distribution of the training data set. In the case of surrogate modelling, the input probability distribution may not be known at training time, i.e. the training distribution (usually a uniform grid) and the distribution we are actually interested in evaluating will not necessarily be the same.

This contribution If the relative density of the training and true input distribution are known, which is a fine assumption in surrogacy, then covariate shifting [3] may be used to produce valid confidence intervals with respect to an alternative distribution, without having to retrain the model. We use covariate shifting to evaluate input uncertainty, in combination with the surrogate's model uncertainty, in two plasma physics modelling scenarios. We further use covariate shifting evaluate an input parametric p-box to the conformalised surrogate. This is achieved by searching through parameters of the input p-box, and applying covariate shifting to each parametric distribution. For a given level of confidence, output p-box bounds are obtained by computing upper and lower cdf envelopes from random prediction sets.

References

[1] Michael Scott Balch. Mathematical foundations for a theory of confidence structures. *International Journal of Approximate Reasoning*, 53(7):1003–1019, 2012. doi:10.1016/j.ijar.2012.05.006.

[2] Leonardo Cella and Ryan Martin. Valid inferential models for prediction in supervised learning problems. In *International Symposium on Imprecise Probability: Theories and Applications*, volume 147, pages 72–82. PMLR, 2021.

[3] Ryan J Tibshirani, Rina Foygel Barber, Emmanuel Candes, and Aaditya Ramdas. Conformal prediction under covariate shift. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

[4] Vladimir Vovk, Alexander Gammerman, and Glenn Shafer. *Algorithmic learning in a random world*. Springer, 1st edition, 2005. doi:10.1007/b106715.