

DeepUQ

AUTHOR 1 BECKY NEVIN, RNEVIN@FNAL.GOV



ABSTRACT

This project aims to create a framework for calibrating uncertainty expectations in various ML and statistical models. It builds on DeepBench’s pendulum module, controlling error injection, calculating an analytic expectation for the error impact on the final confidence intervals, and comparing this expectation to that produced by various uncertainty-aware ML and statistical techniques. Here, we pursue the following modeling techniques: hierarchical and non-hierarchical Hamiltonian Monte Carlo sampling (**numpyro**), hierarchical and non-hierarchical simulation-based inference (**mackelab**), and deep ensembles. We inject aleatoric error on the pendulum parameters individually (L , θ_0 , and a_g) at a variety of levels (1%, 10%, and 50%). We compare this to expectations of aleatoric and epistemic error from the various ML techniques and explore the bias and confidence of the models compared to our analytic expectation.

1. INTRODUCTION

Cite other UQ techniques, mostly Caldeira & Nord.

2. METHODS

2.1. *Uncertainty definition and injection*

2.2. *Modeling techniques*

2.2.1. *HMC Sampling*

2.2.2. *SBI*

2.2.3. *DE*

3. ANALYSIS

3.1.

4. ACKNOWLEDGEMENTS

Make sure to cite [Harris et al. \(2020\)](#) all of your sources [Hunter \(2007\)](#).

You can also optionally provide contributions by person:

Becky Nevin—Author 1 contributed X Y and Z

Author 2—Author 2 contributed A B and C

If you work with the DeepSkies research group; please include the following text:

We acknowledge the Deep Skies Lab as a community of multi-domain experts and collaborators who’ve facilitated an environment of open discussion, idea-generation, and collaboration. This community was important for the development of this project.

REFERENCES

Harris, C. R., Millman, K. J., van der Walt, S. J., et al.
2020, Nature, 585, 357, doi: [10.1038/s41586-020-2649-2](https://doi.org/10.1038/s41586-020-2649-2)

Hunter, J. D. 2007, Computing in Science & Engineering, 9,
90, doi: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55)

APPENDIX

A. APPENDIX

A.1. *Table References*