SAND2021-4692C

Is Discrete-Direct model calibration and uncertainty propagation more Trustworthy for estimating Tail Probabilities and Percentiles than Bayesian approaches?

Vicente Romero

V&V, UQ, and Credibility Processes Dept

Sandia National Laboratories

Albuquerque, NM

*Sandia National Laboratories document SAND2021-yyyy C (unlimited release)
**Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

ASME V&V Symposium 2021
May 19-20, on-line

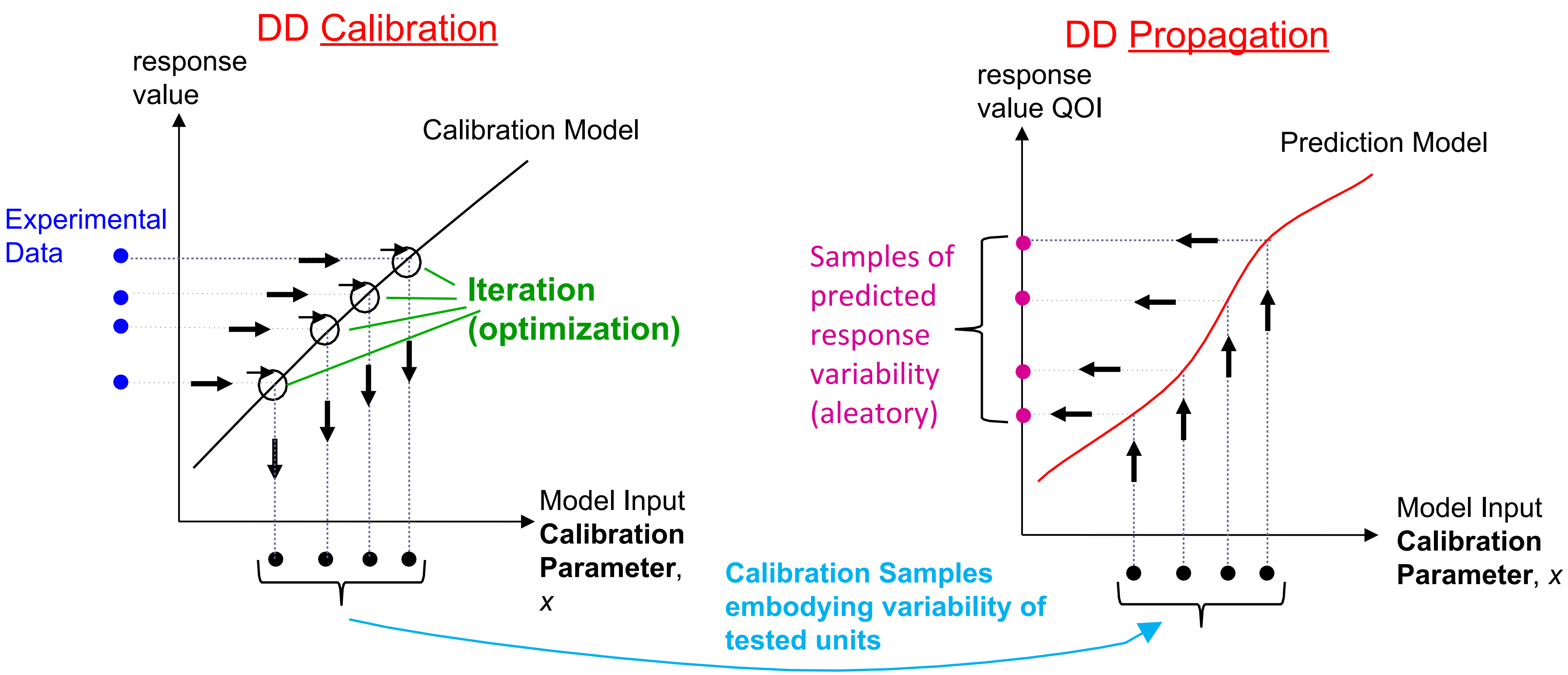
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Outline

- **Introduce the Discrete-Direct (DD) model calibration and uncertainty propagation approach:**
    - calibrate a material/phenomena/device model to each of multiple replicate experiments to incorporate stochastic variability effects into the model and propagate them to predictions

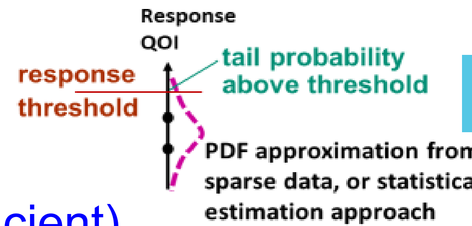
- **Compare DD, Bayesian, and other calibration-prediction approaches for stochastic calibration-propagation problems under conditions of sparse replicate test data:**
    - cost
    - complexity
    - trustworthiness

“Discrete Direct” (DD) Model Calibration and Uncertainty Propagation



- Propagate the <u>discrete values</u> of the calibration parameters
- Straightforwardly extends to problems with <u>multiple</u> calibration parameters
- *N* runs of model to propagate *N* param. values or sets from *N* calibration experiments
- Simple to <u>update</u> w/new experiments/data that may become available (w/out Bayes’ rule & machinery)

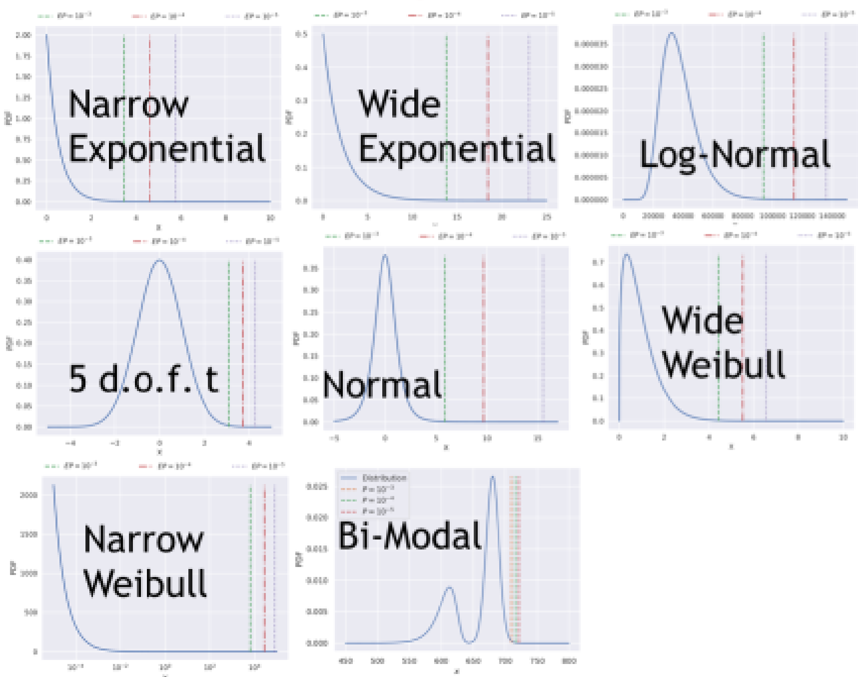
Testing and Optimization of Sparse-Sample 1-D Tail-Probability Estimation Methods



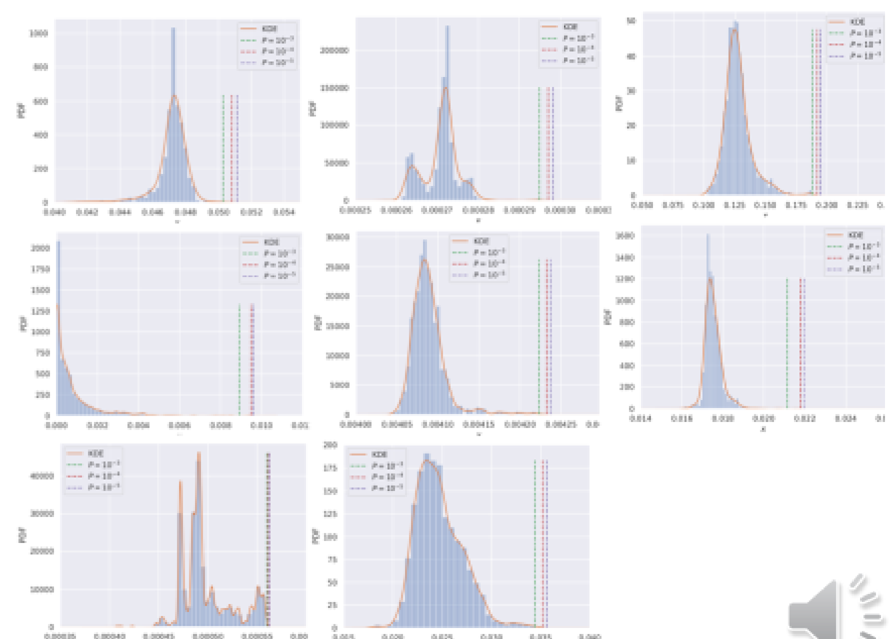
Objective: be conservative but not overly conservative (efficient)

- Large numerical study with > 3e+08 performance tests
- ~20 established and newly developed methods: variants, combinations, hybrids
- tail probability magnitudes 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}
- # samples N = 2, 3, 4,…,20
- 16 diverse distribution shapes below
- 10K random sampling trials for each combinations of the above factors studied

8 analytical PDFs



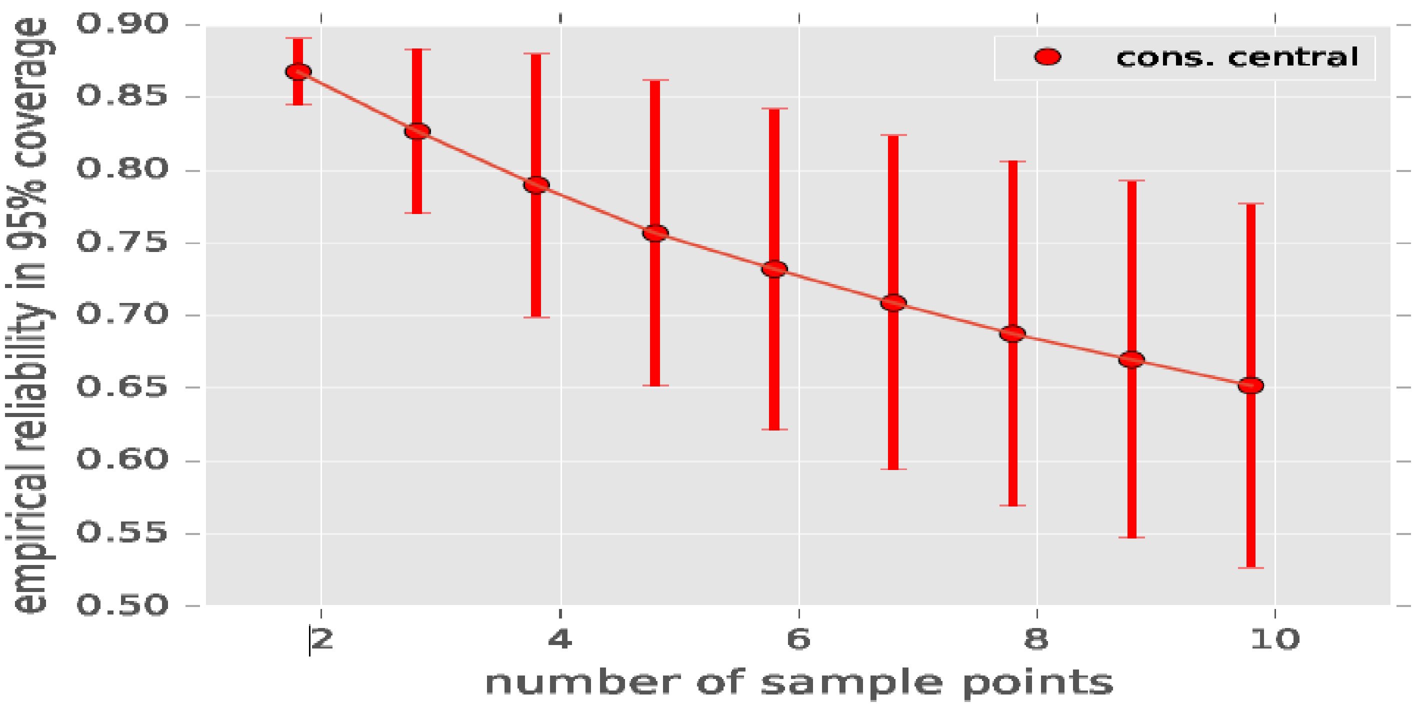
8 empirical PDFs



Robustness of 95/90 TIs for Bounding Central 95% of Non-Normal Distributions

Sandia National Laboratories

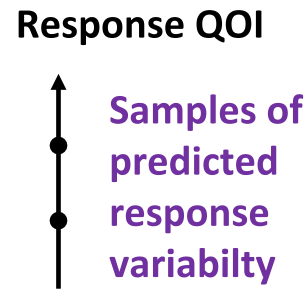
Empirical success rates on 144 Non-Normal distributions (bars capture 90% of the 144 cases)



- **Can get higher reliabilities for N≥3 by using Statistical Jackknifing and averaging resulting TIs**

DD Simplicity, Cost, and Trustworthiness Advantages: maps Multi-D UQ ➦ 1-D UQ

- <u>Example</u>:
  - a model with 3 calibration parameters
  - N=2 replicate experiments



- **<u>DD</u>: 2 calibrations ➔ 2 parameter sets ➔ 2 runs of prediction model ➔ 2 values of response**
  - get a 1-D UQ problem with 2 samples of response
  - empirical confidence levels on bounding estimates of response statistics

- **Some <u>other</u> calibration-propagation approaches would try to infer a 3-D Joint PDF of variability of the calibration parameters from the 2 calibration parameter sets (data points) in the 3-D space, then propagate the JPDF**
  - get a highly questionable JPDF and predicted PDF of response, uncertainty would be difficult and expensive to reliably estimate
  - JPDF propagation requires high expense or a surrogate model (added complexity, uncertainty, and more runs of the prediction model than DD)

Bayesian Approaches:

Additional Difficulties and Uncertainties

- **Bayesian approaches would typically propose PDF model forms for the marginal distributions of the calibration-parameter aleatory variations**
  - ***hyper-parameters* of the proposed aleatory PDFs are assigned prior distributions to describe their uncertainties**
  - **<u>little to no prior knowledge</u> would generally be available concerning hyper-parameter values, especially for non-Normal distributions which have non-intuitive parameters**
  - **so, broad uncertainties would most prudently be assigned**
  - **The hyper-parameters are calibrated by Bayesian updating using experimental data to reduce their uncertainties**
  - **sparse experimental data → hyper-parameter uncertainties will remain relatively broad after Bayesian updating**
  - **The proposed aleatory PDF model forms themselves will likely be incorrect in the first place because sparse data does not support accurate selection of PDF model forms**

Additional Difficulties/Uncertainties with Bayesian Approaches (cont'd)

Sandia National Laboratories

- **The many formulation options in Bayesian approaches can lead to substantial <u>analyst-to-analyst and results variability</u>**
  - Use a hyper-parameter formulation or don't?
  - Use a discrepancy term or don't?
  - What type of parametric and/or non-parametric PDF model forms or formulations, and Priors, will be used?
  - What sampling approach (MCMC or other) will be used, and what surrogate modeling approach will be used to make it affordable?

- **A Bayesian-derived JPDF with linear dependency approximations will have substantial error/uncertainty that must be appropriately characterized and accounted for**
  - *How is this done? Has the effectiveness been characterized in realistic and representative test problems under random trials?*
  - *How well can I trust the results in real applications?*

Closing Remarks

- **The DD calibration-UQ approach for sparse calibration data is versatile and relatively simple, inexpensive, and reliable in terms of characterized confidence of giving conservative estimates of response statistics**
- **The methodology has been confirmed on several test problems**
  - 4 calibration variables, strain-rate dependent material plasticity model calibrated to 4 stress-strain curves, 16 Can-Crush output responses (ASME UQ-Risk journal paper)
  - 8 calibration variables, material plasticity model calibrated to 5 stress-strain curves, Pressure Vessel max load
  - radiation-damaged electronics (3 device models calibrated to time-dependent functional data curves, then used in circuit response tail-probability predictions)
- **A Sandia document with more detailed discussion of these issues will likely be available by the time you see this presentation (contact vjromer@sandia.gov)**