



Validation Metrics for Applications with Mixed Uncertainty

Dr. Laura White
NASA Langley Research Center

DATAWorks 2026
4/23/2026



What is Validation?



“All models are wrong, but some are useful.”

-George Box, 1976¹

¹Box, G. E. P. (1976). Science and Statistics. *Journal of the American Statistical Association*, 71(356), 791-799



What is Validation?

“All models are wrong, but some are useful.”

-George Box, 1976¹

Oberkampff et al². address three aspects of validation:

- A comparison between computed system response quantities (SRQs) of interest and measured SRQs that provides a **quantifiable measurement of accuracy**

¹Box, G. E. P. (1976). Science and Statistics. *Journal of the American Statistical Association*, 71(356), 791-799

²Oberkampff, William L., and Timothy G. Trucano. "Verification and validation benchmarks." *Nuclear engineering and Design* 238.3 (2008): 716-743.4444



What is Validation?

"All models are wrong, but some are useful."

-George Box, 1976¹

Oberkampff et al. address three aspects of validation:

- A comparison between computed system response quantities (SRQs) of interest and measured SRQs that provides a **quantifiable measurement of accuracy**
- Uses **interpolation or extrapolation** of computer models to make predictions for conditions within the model's domain of use

¹Box, G. E. P. (1976). Science and Statistics. *Journal of the American Statistical Association*, 71(356), 791-799

²Oberkampff, William L., and Timothy G. Trucano. "Verification and validation benchmarks." *Nuclear engineering and Design* 238.3 (2008): 716-743.4444



What is Validation?

"All models are wrong, but some are useful."

-George Box, 1976¹

Oberkampff et al. address three aspects of validation:

- A comparison between computed system response quantities (SRQs) of interest and measured SRQs that provides a **quantifiable measurement of accuracy**
- Uses **interpolation or extrapolation** of computer models to make predictions for conditions within the model's domain of use
- Determination of whether the estimated accuracy of the computational model results satisfies the **accuracy requirements for the SRQs** of interest

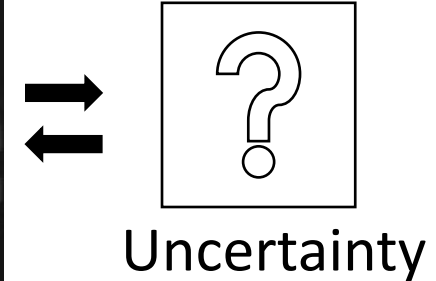
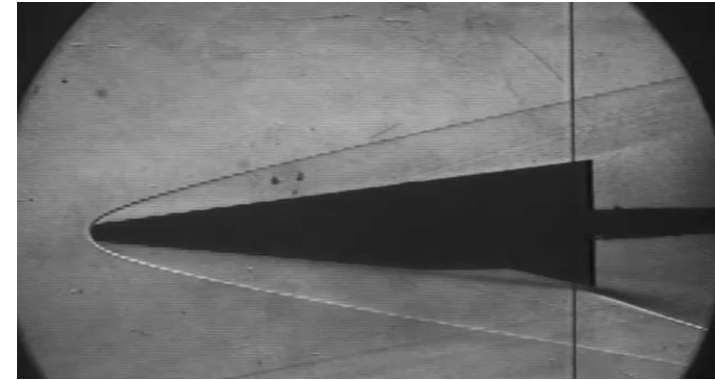
¹Box, G. E. P. (1976). Science and Statistics. *Journal of the American Statistical Association*, 71(356), 791-799

²Oberkampff, William L., and Timothy G. Trucano. "Verification and validation benchmarks." *Nuclear engineering and Design* 238.3 (2008): 716-743.4444

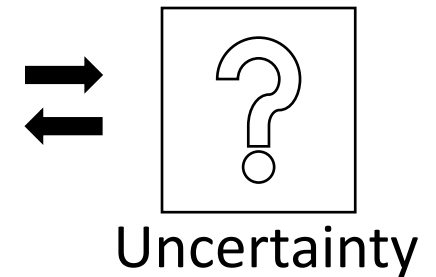
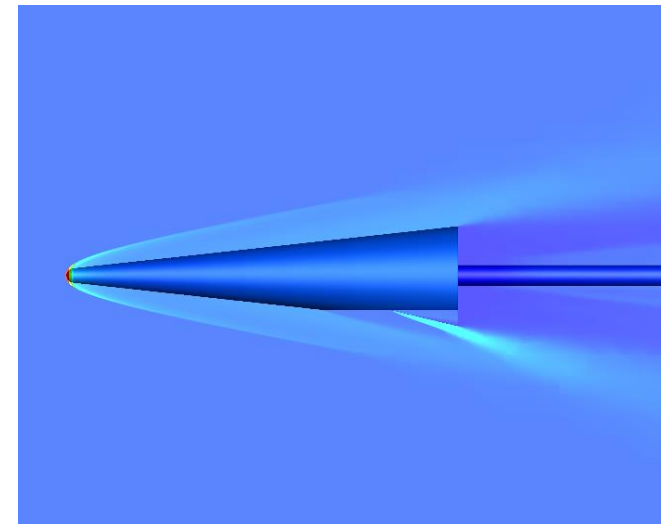
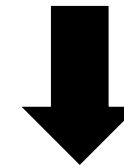


Model Validation – In this Context

- Model validation determines the degree of accuracy between the real world and the simulation
- The validation method captures differences between experimental and simulation data with aleatory and epistemic uncertainty included in the data



Validation





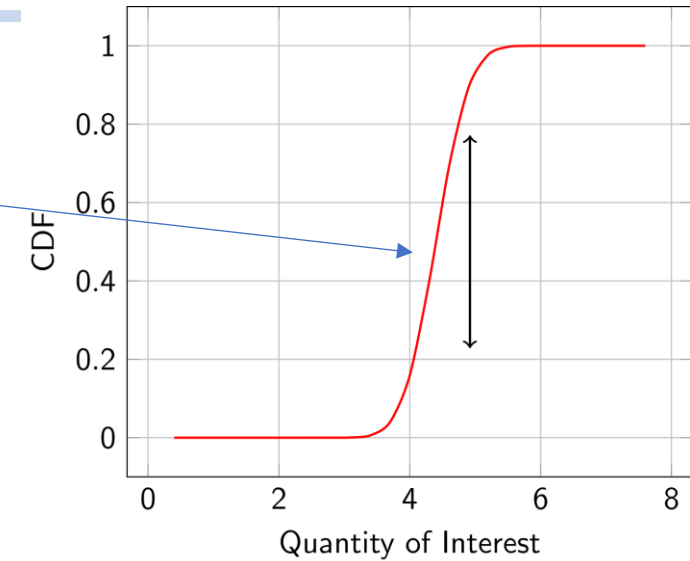
Introduction – Uncertainty Quantification

- Aleatory Uncertainty
 - aka stochastic uncertainty
 - Irreducible variability inherent in nature
 - Ex: flipping a coin, measurement error



Introduction – Uncertainty Quantification

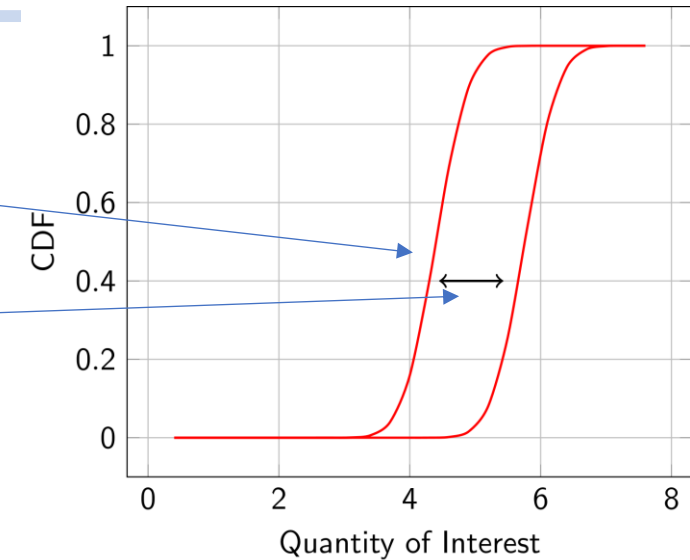
- Aleatory Uncertainty
 - aka stochastic uncertainty
 - Irreducible variability inherent in nature
 - Ex: flipping a coin, measurement error





Introduction – Uncertainty Quantification

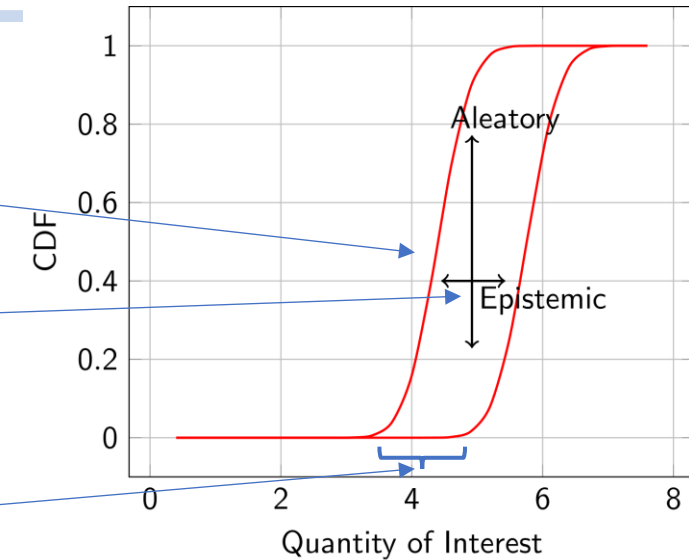
- Aleatory Uncertainty
 - aka stochastic uncertainty
 - Irreducible variability inherent in nature
 - Ex: flipping a coin, measurement error
- Epistemic Uncertainty
 - aka systematic uncertainty
 - Lack of knowledge
 - Ex: impact of radiation on humans, hum





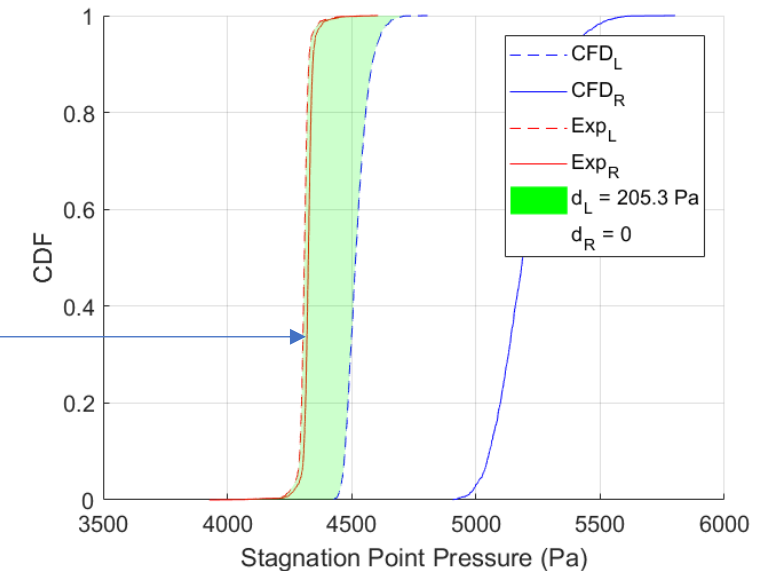
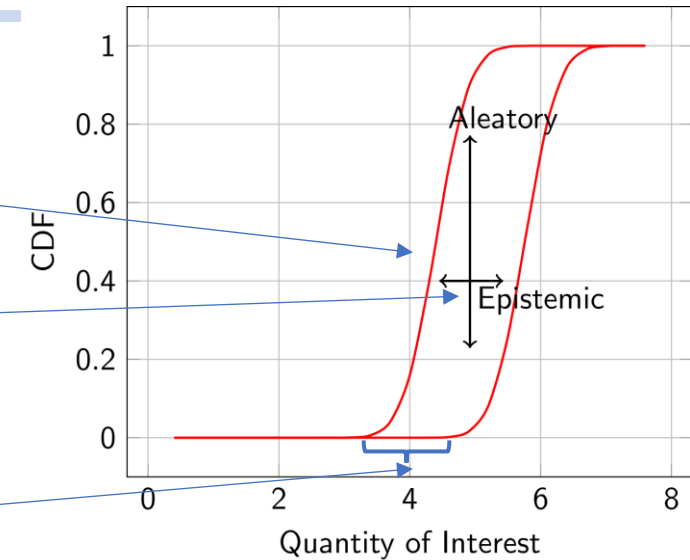
Introduction – Uncertainty Quantification

- Aleatory Uncertainty
 - aka stochastic uncertainty
 - Irreducible variability inherent in nature
 - Ex: flipping a coin, measurement error
- Epistemic Uncertainty
 - aka systematic uncertainty
 - Lack of knowledge
 - Ex: impact of radiation on humans, hum
- Mixed Uncertainty
 - A combination of aleatory and epistemic uncertainty exists corresponding to different inputs into the system
 - The end form is a probability box with **multiple possible distributions** that could represent the true quantity of interest distribution



Introduction – Uncertainty Quantification

- Aleatory Uncertainty
 - aka stochastic uncertainty
 - Irreducible variability inherent in nature
 - Ex: flipping a coin, measurement error
- Epistemic Uncertainty
 - aka systematic uncertainty
 - Lack of knowledge
 - Ex: impact of radiation on humans, hum
- Mixed Uncertainty
 - A combination of aleatory and epistemic uncertainty exists corresponding to different inputs into the system
 - The end form is a probability box with **multiple possible distributions** that could represent the true quantity of interest distribution
- Model-form Uncertainty
 - Modeling and quantification of the uncertainties about the ability of a computational model to truly represent the physics underlying the problem to be solved





Quantify Model-Form Uncertainty by Validation Metric

Defn: A mathematical operator that measures the difference between the SRQ obtained from a simulation result and one obtained from experimental measurements [Oberkampff, Roy, 2010¹]

Metric Options:

- Hypothesis Testing
- Comparison of Means
- Area Metrics

¹Oberkampff, William L., and Christopher J. Roy. *Verification and validation in scientific computing*. Cambridge university press, 2010.



Quantify Model-Form Uncertainty by Validation Metric

Defn: A mathematical operator that measures the difference between the SRQ obtained from a simulation result and one obtained from experimental measurements [Oberkampff, Roy, 2010¹]

Metric Options:

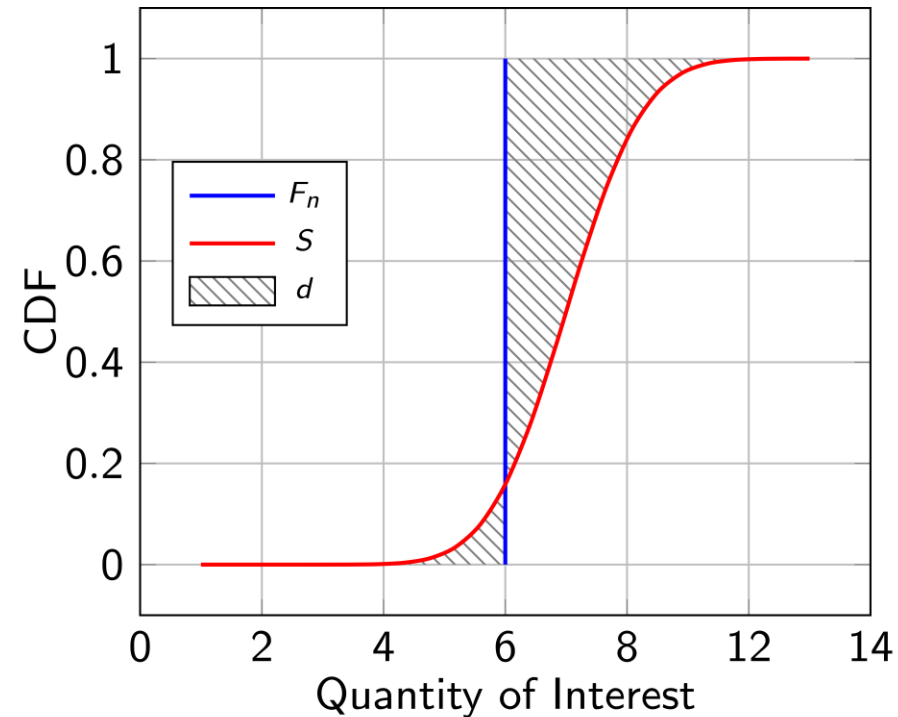
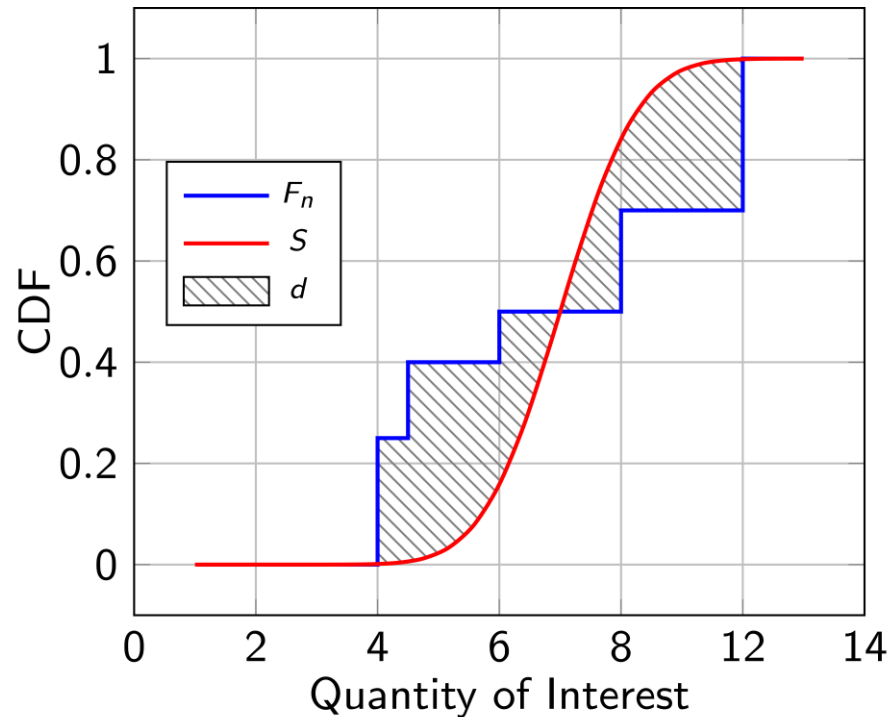
- Hypothesis Testing
- Comparison of Means
- **Area Metrics**

¹Oberkampff, William L., and Christopher J. Roy. *Verification and validation in scientific computing*. Cambridge university press, 2010.

Area Metric Operator – Aleatory Uncertainty

- Defined as the area between the prediction distribution S and the data distribution F_n which measures the mismatch between them.

$$d(S, F_n) = \int_{-\infty}^{\infty} |S(x) - F_n(x)| dx$$

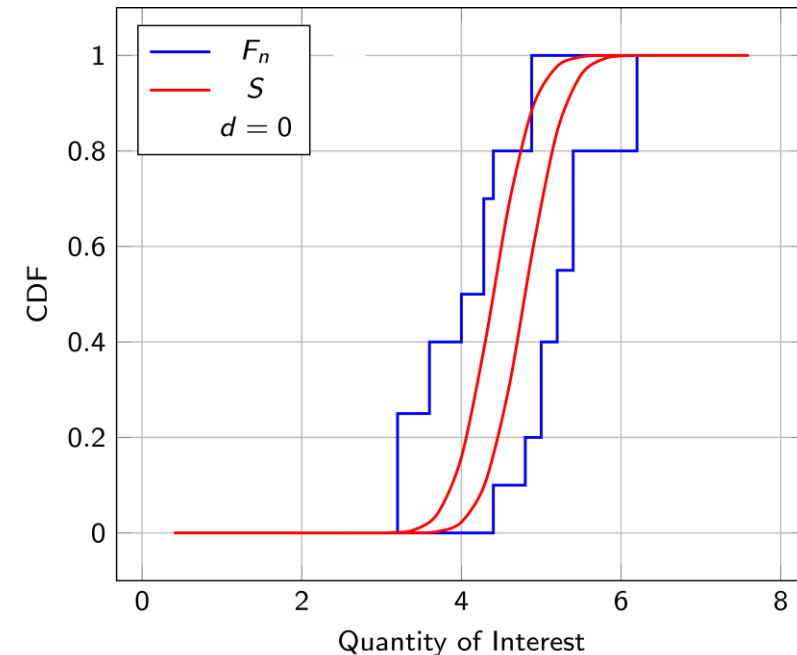
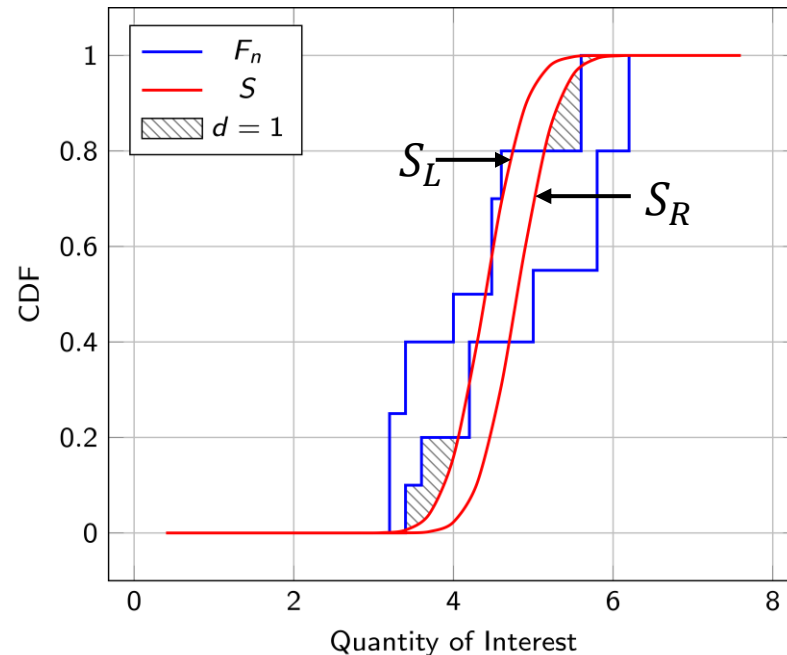


Area Metric Operator – Mixed Uncertainty

- Combining epistemic and aleatory uncertainties results in a p-box, and Ferson et al. (2009)¹ proposed a novel area metric.
- This metric evaluates the difference by calculating the area between the prediction p-box and the data p-box

$$d(S, F_n) = \int_{-\infty}^{\infty} \Delta([S_R(x), S_L(x)], [F_{nR}(x), F_{nL}(x)]) dx$$

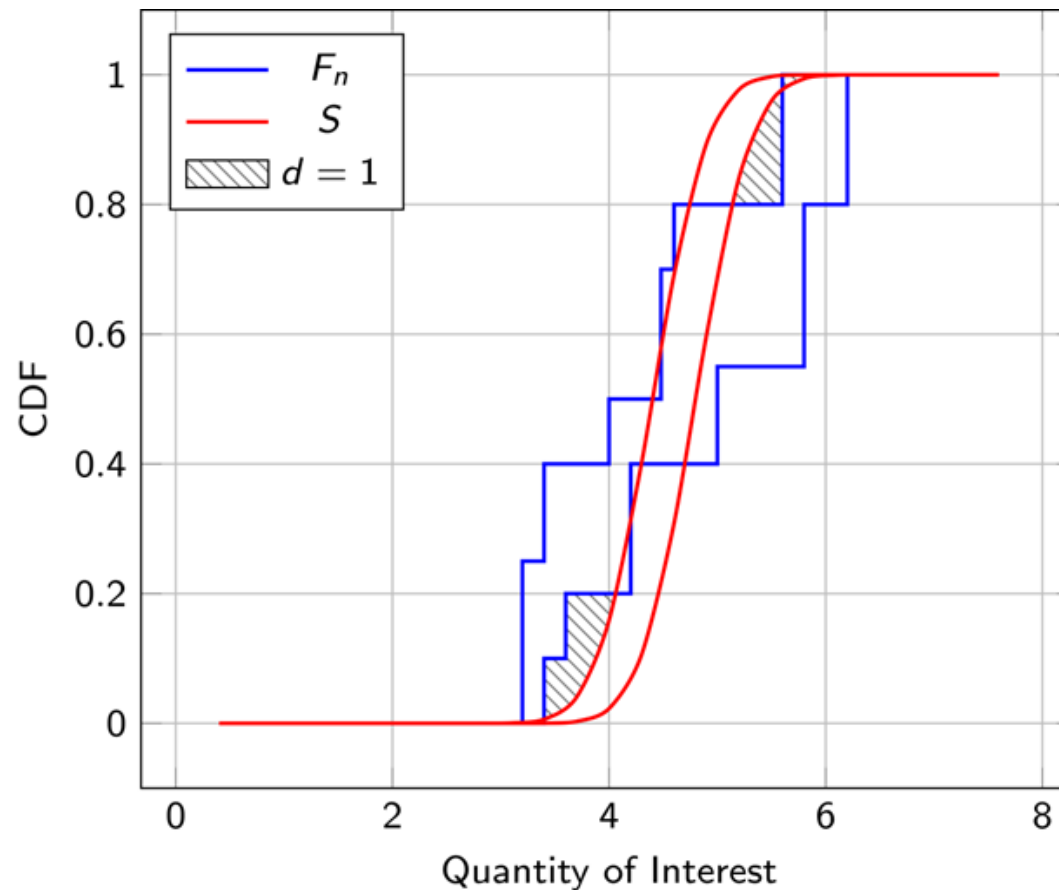
Here, $\Delta(A, B) = \min_{a \in A, b \in B} |a - b|$, R and L represent left and right bounds of the distributions



¹Ferson, Scott, and William L. Oberkampf. "Validation of imprecise probability models." *International Journal of Reliability and Safety* 3.1-3 (2009): 3-22.

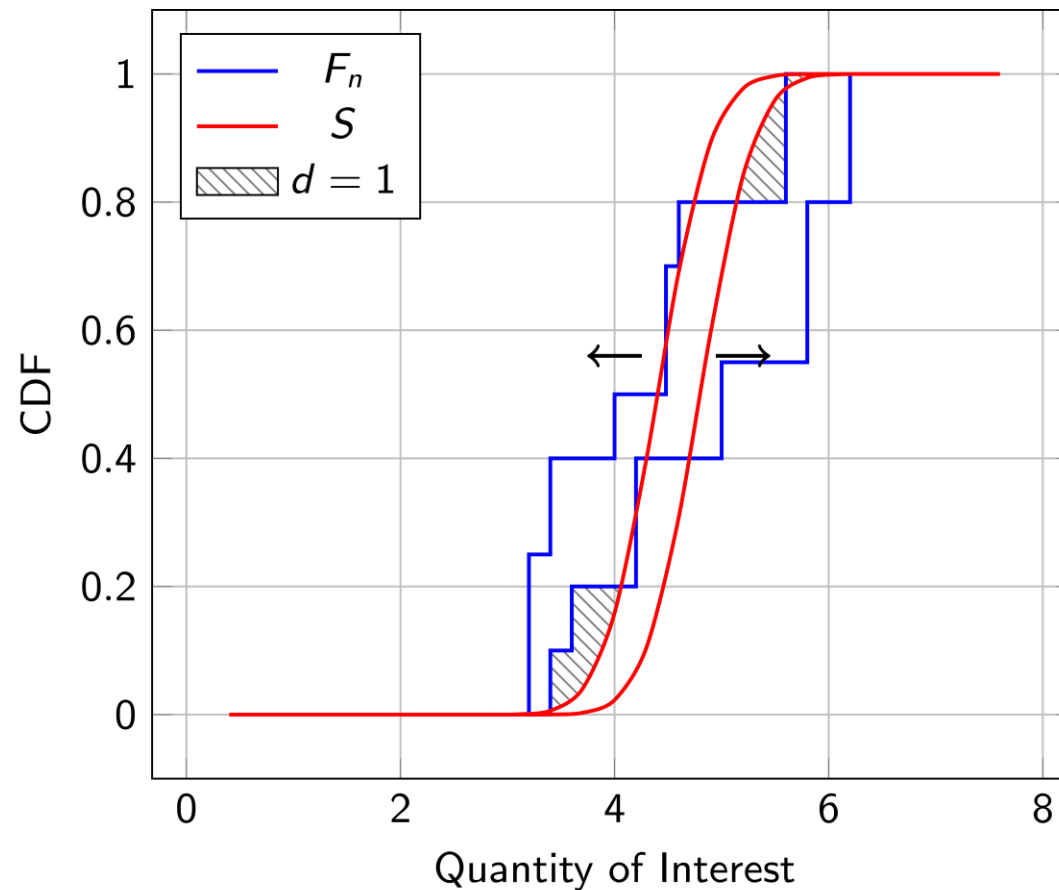
Validation Metric Operator Shortcomings

- d can be used to quantify the model form uncertainty and is appended to the p-box



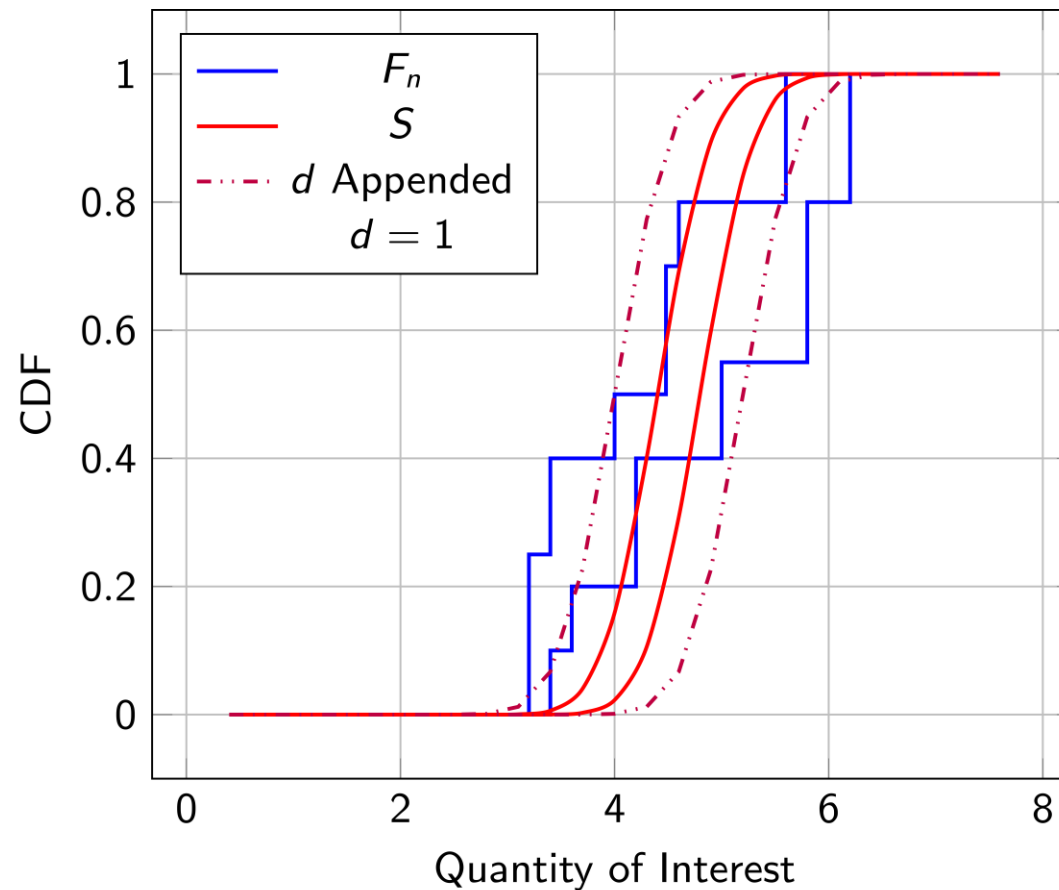
Validation Metric Operator Shortcomings

- d can be used to quantify the model form uncertainty and is appended to the p-box



Validation Metric Operator Shortcomings

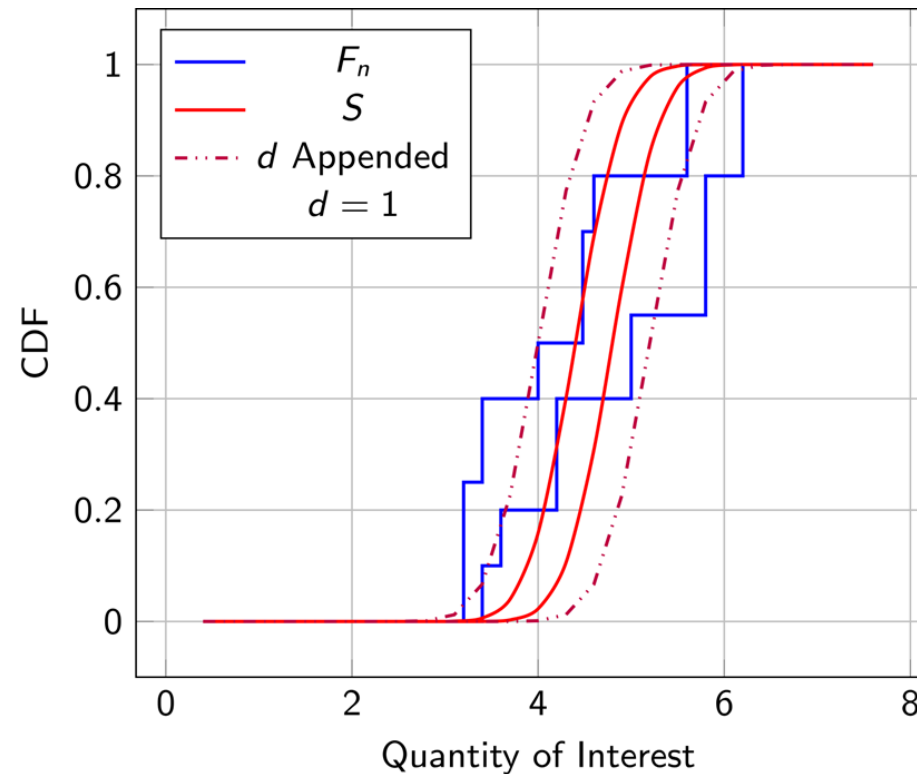
- d can be used to quantify the model form uncertainty and is appended to the p-box





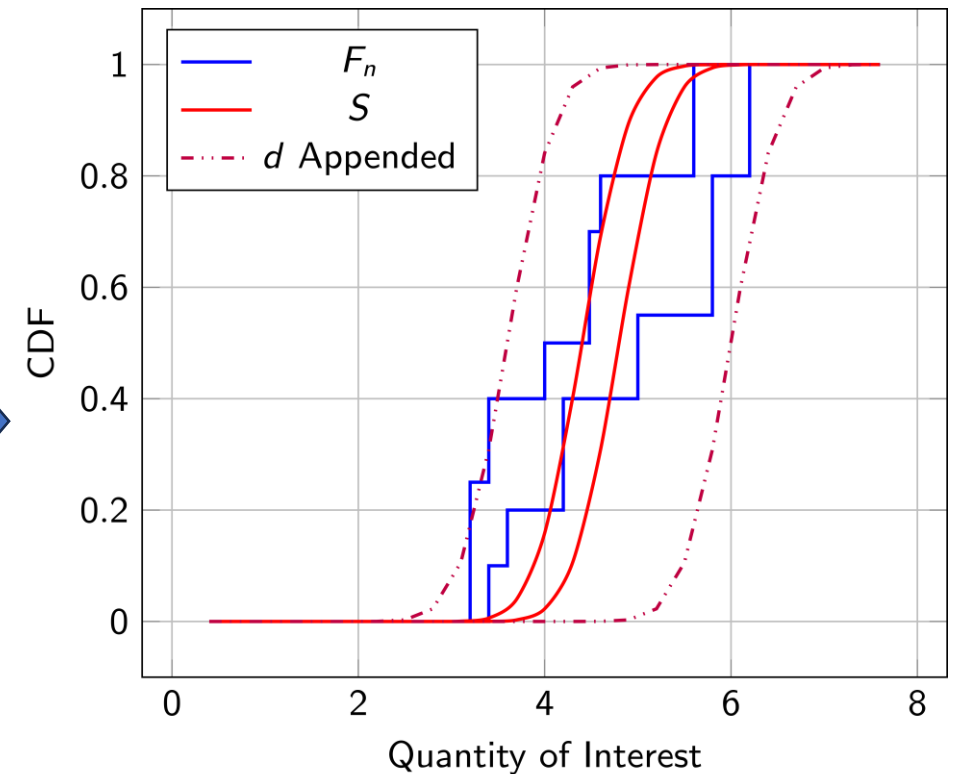
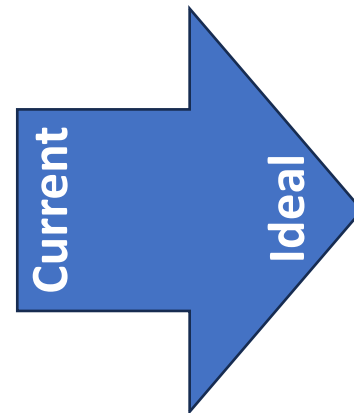
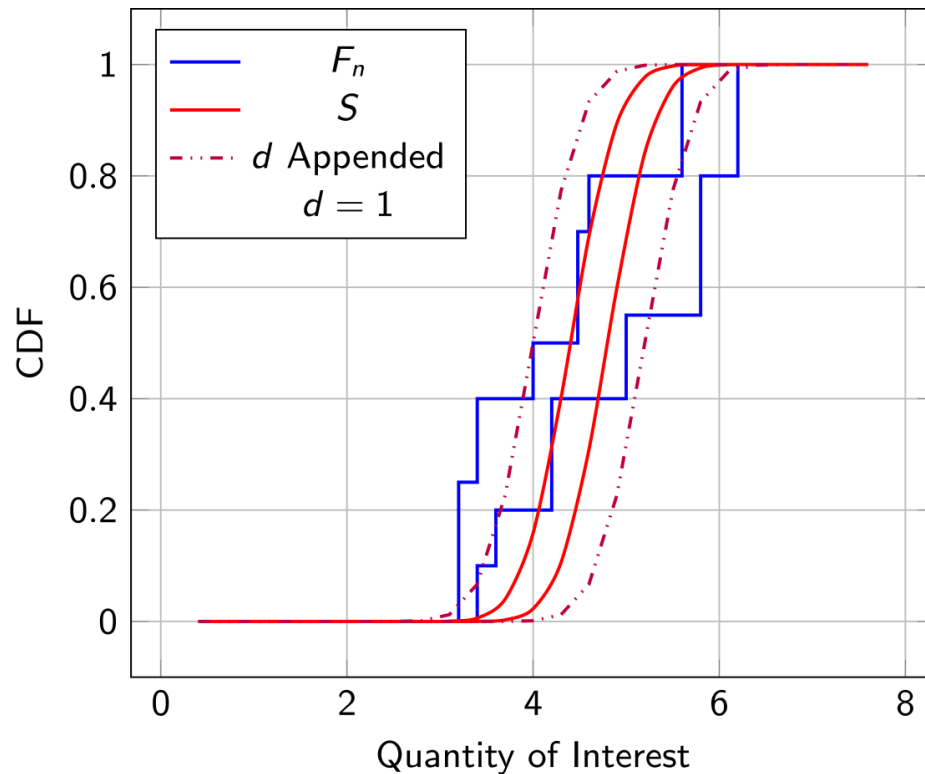
Validation Metric Operator Shortcomings

- d can be used to quantify the model form uncertainty and is appended to the p-box
- However, for mixed uncertainty this allows for over-compensating for the model form uncertainty in some areas while not capturing enough of it in other areas.



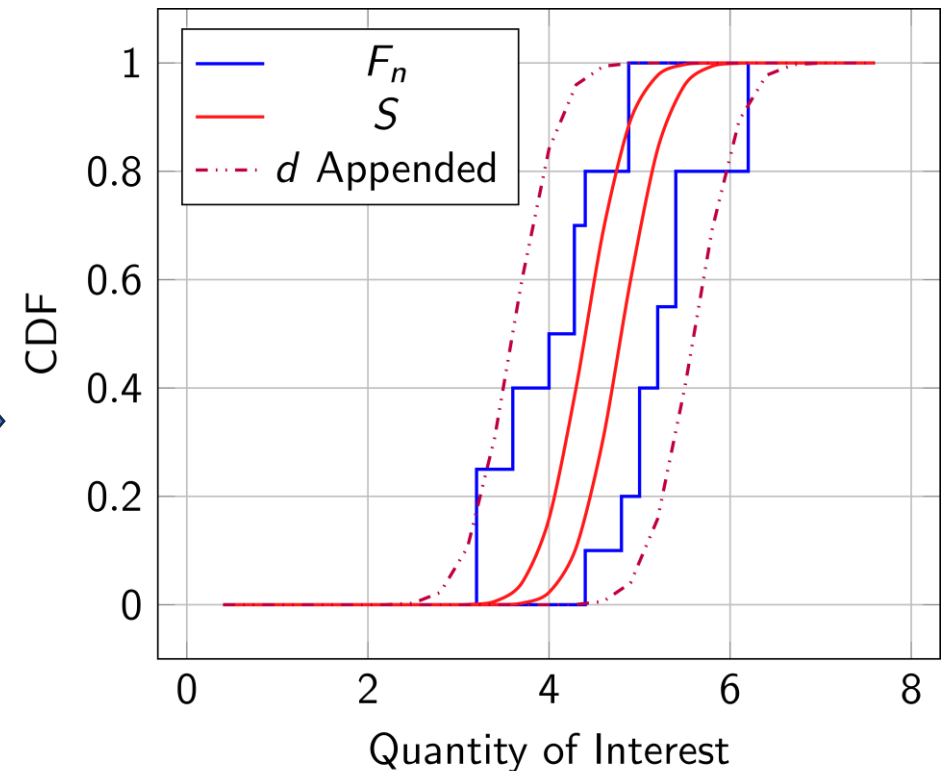
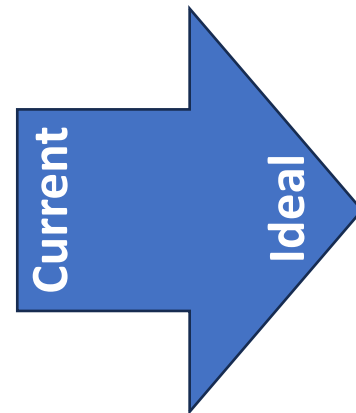
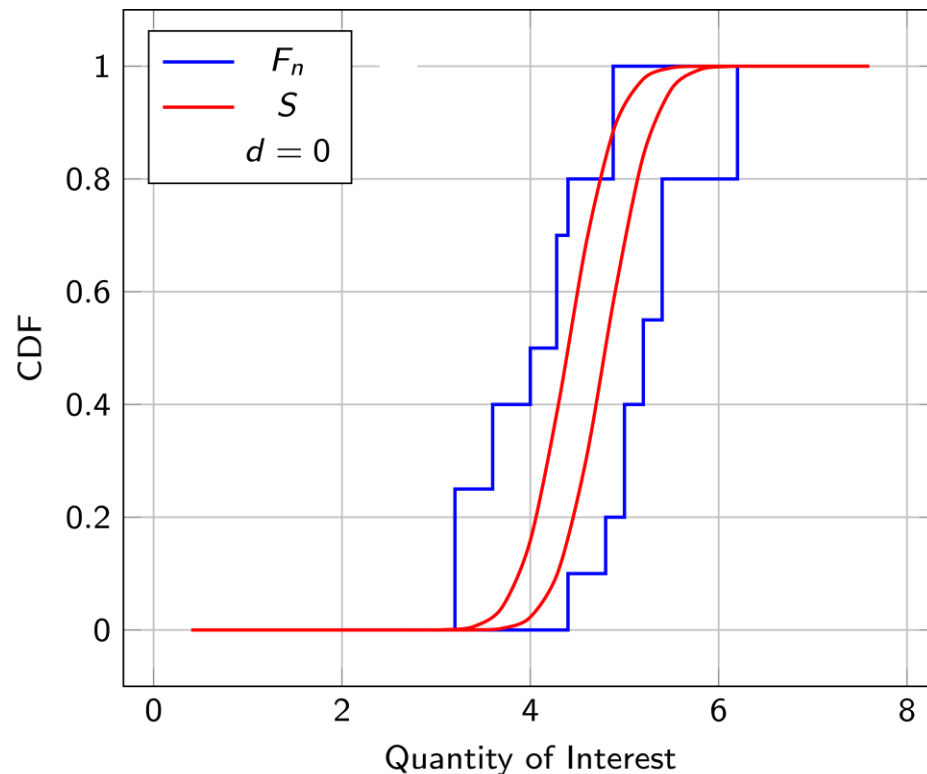
Validation Metric Operator Shortcomings

- d can be used to quantify the model form uncertainty and is appended to the p-box
- However, for mixed uncertainty this allows for over-compensating for the model form uncertainty in some areas while not capturing enough of it in other areas.



Validation Metric Operator Shortcomings

- d can be used to quantify the model form uncertainty and is appended to the p-box
- However, for mixed uncertainty this allows for over-compensating for the model form uncertainty in some areas while not capturing enough of it in other areas.

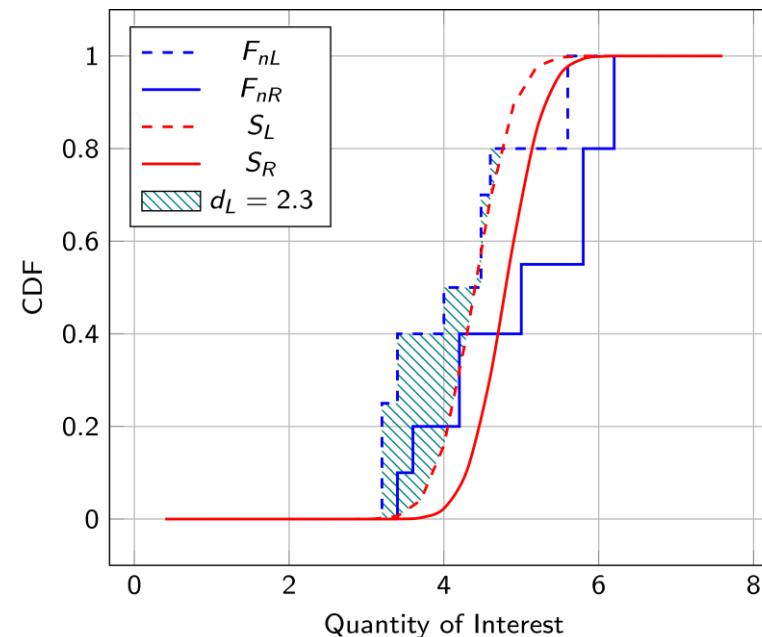




Alternative Area Validation Metric – Mixed Uncertainty

- We proposed a new area validation metric in 2020¹
- This metric handles the disagreement between the left and right bounds of prediction distribution and data distribution.
- Define a left (L) and right (R) metric as follows:

$$d_i(S_i, F_{ni}) = \int_{-\infty}^{\infty} g_i(x) dx \quad \text{where} \quad g_L(x) = \begin{cases} 0 & x \text{ s.t. } S_L(x) \geq F_{nL}(x) \\ F_{nL}(x) - S_L(x) & x \text{ s.t. } S_L(x) < F_{nL}(x) \end{cases} \quad \text{for } i \in \{L, R\}$$

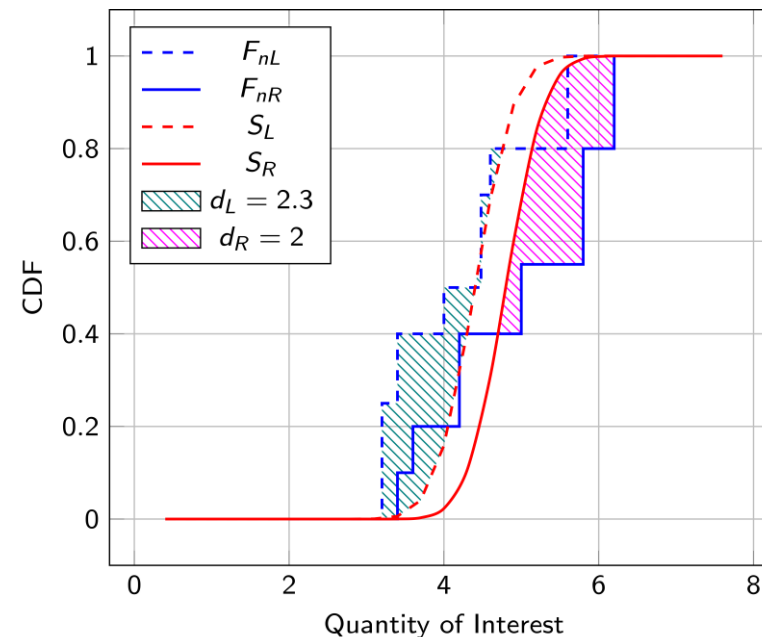


¹Brune, Andrew J., Thomas K. West IV, and Laura M. White. "Calibration probe uncertainty and validation for the hypersonic material environmental test system." *Journal of Thermophysics and Heat Transfer* 34.2 (2020): 404-420.

Alternative Area Validation Metric – Mixed Uncertainty

- We proposed a new area validation metric in 2020¹
- This metric handles the disagreement between the left and right bounds of prediction distribution and data distribution.
- Define a left (L) and right (R) metric as follows:

$$d_i(S_i, F_{ni}) = \int_{-\infty}^{\infty} g_i(x) dx \quad \text{where} \quad g_R(x) = \begin{cases} 0 & x \text{ s.t. } S_R(x) \geq F_{nR}(x) \\ S_R(x) - F_{nR}(x) & x \text{ s.t. } S_R(x) < F_{nR}(x) \end{cases} \quad \text{for } i \in \{L, R\}$$

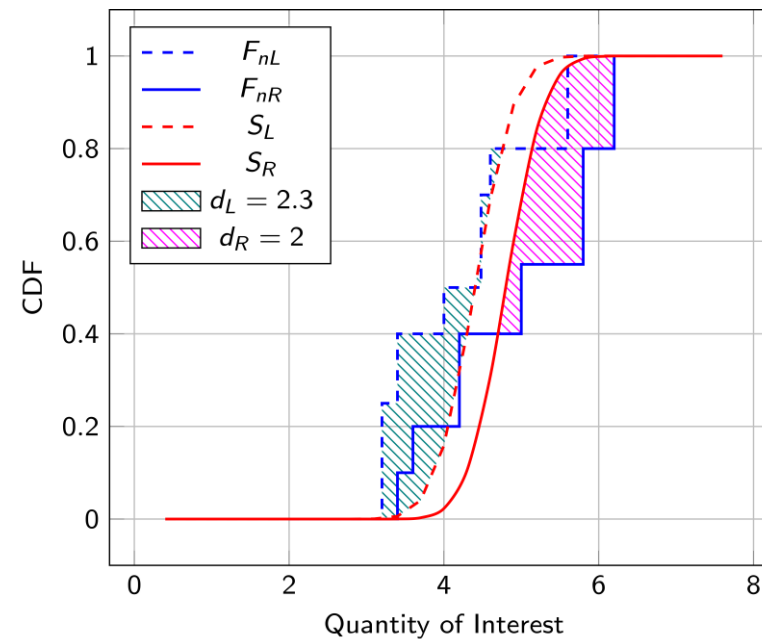


¹Brune, Andrew J., Thomas K. West IV, and Laura M. White. "Calibration probe uncertainty and validation for the hypersonic material environmental test system." *Journal of Thermophysics and Heat Transfer* 34.2 (2020): 404-420.



Alternative Area Validation Metric – Mixed Uncertainty

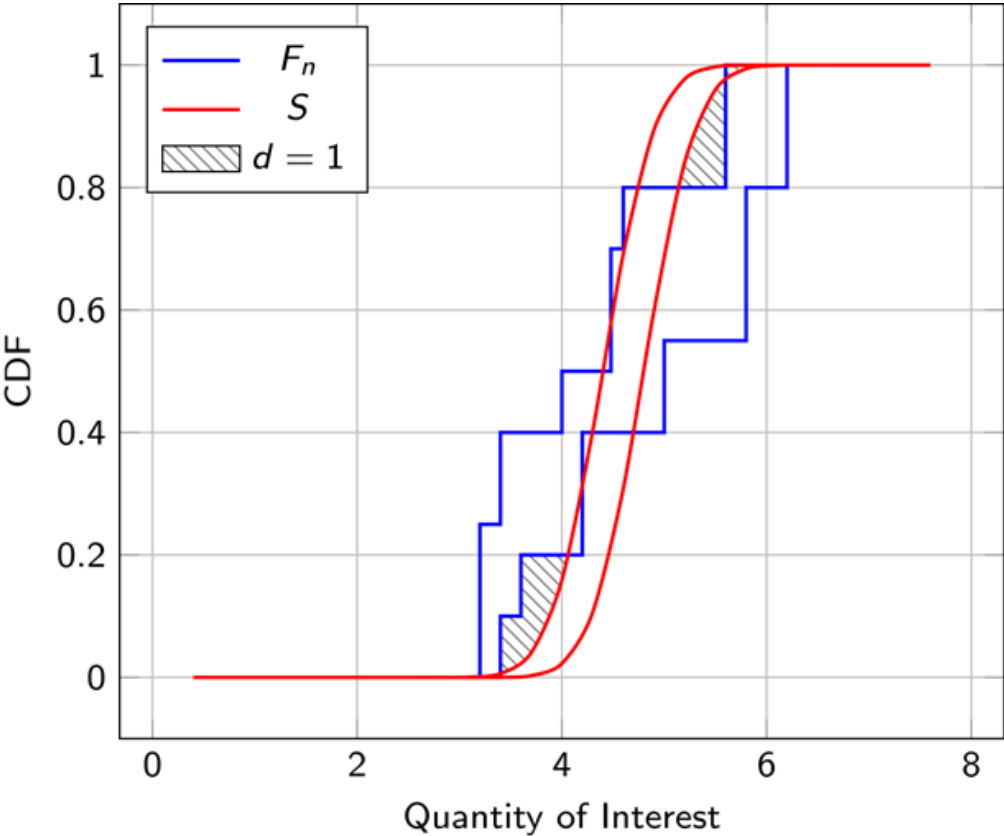
- When $d_L > 0$ there is an overprediction in the simulation data compared to experimental results
- When $d_R > 0$ there is an underprediction in the simulation data compared to experimental results
- When $d_R = d_L = 0$ then the CFD data already covers at least 95% of the possible experimental results



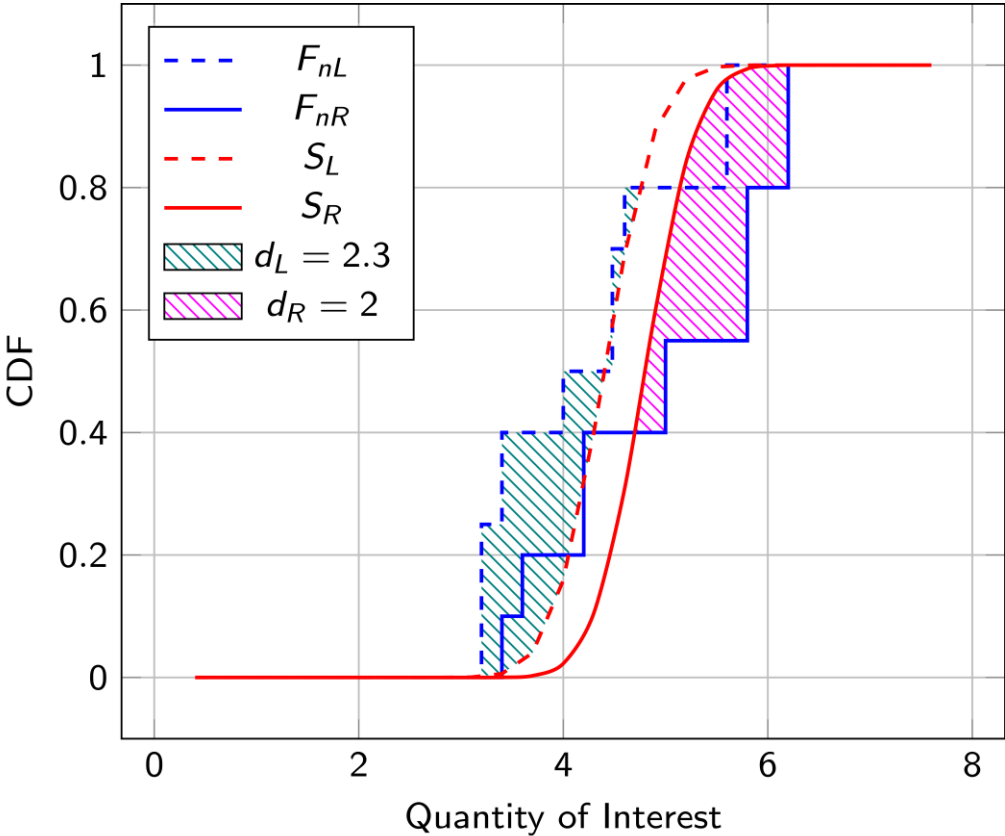
Comparison



Ferson et al. Metric



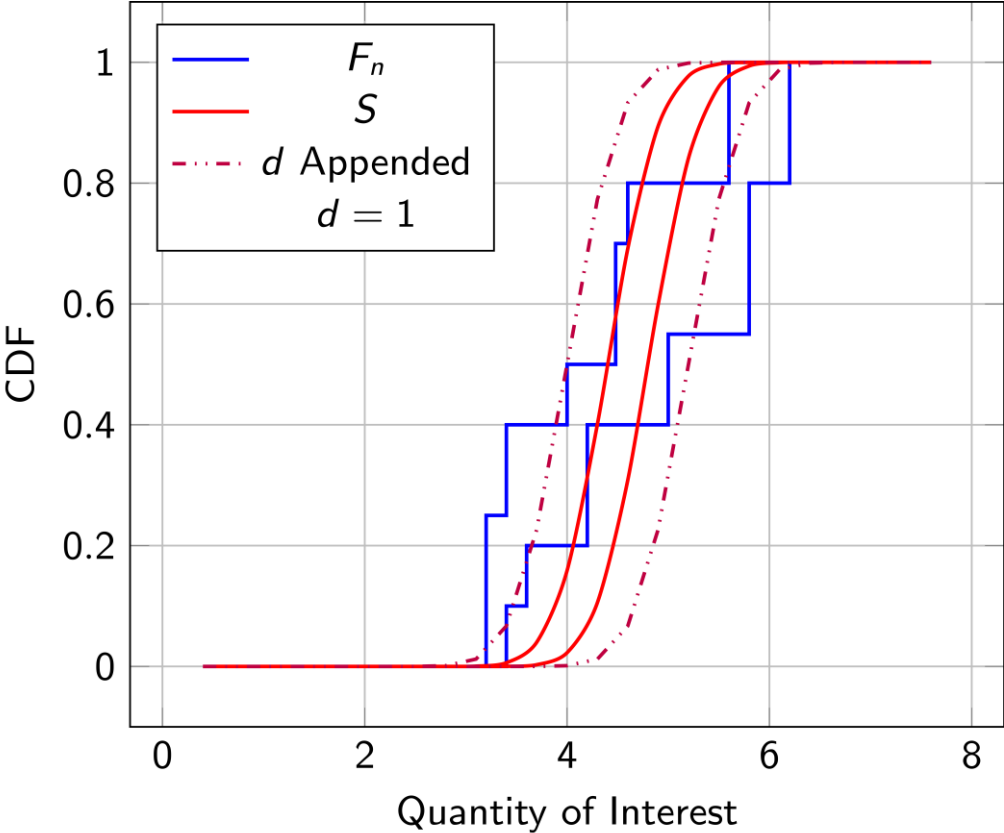
Brune et al. Metric



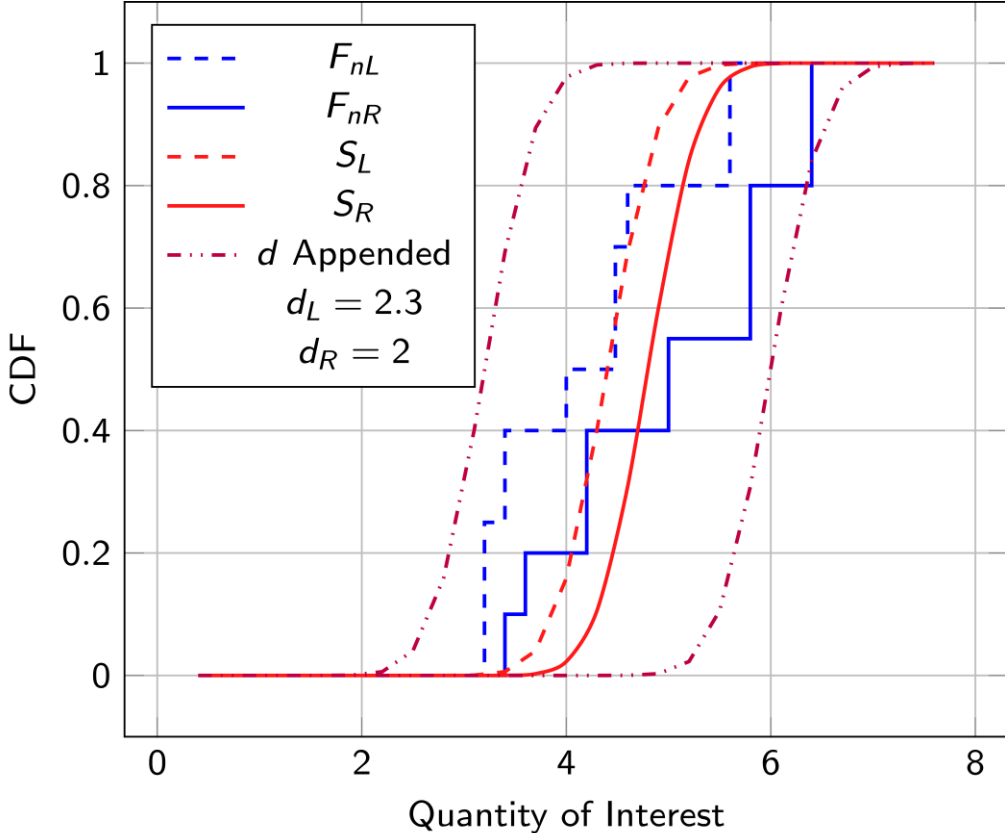
Comparison



Ferson et al. Metric



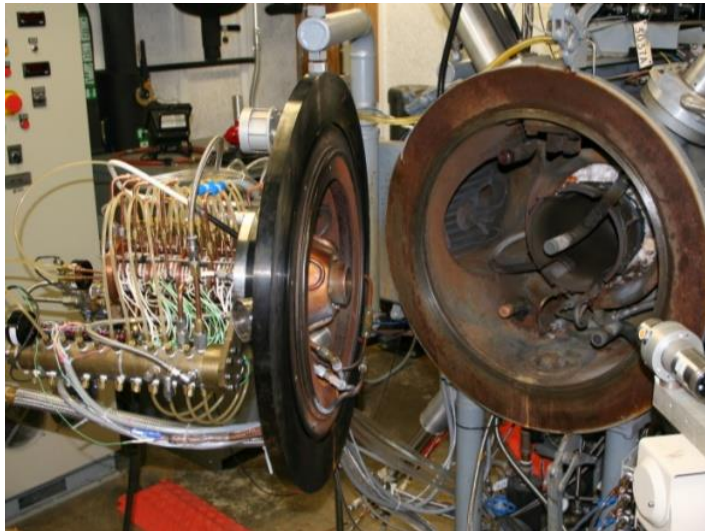
Brune et al. Metric



Applications



- Stagnation Point Calibration [Brune et al.¹]
 - Experimental data: from NASA Langley Hypersonic Materials Environmental Test System (HyMETS) facility
 - Simulation data: from Langley Aerothermodynamic Upwind Relaxation Algorithm (LAURA-5) to compute the flow-field
 - 47 mixed uncertainty parameters

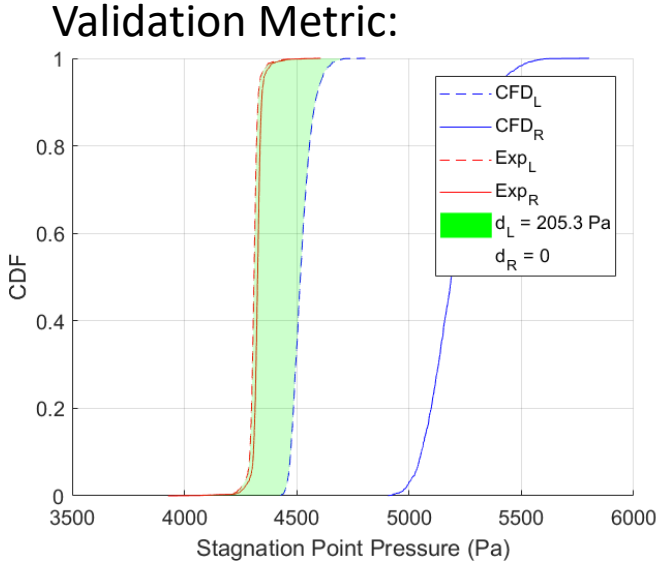
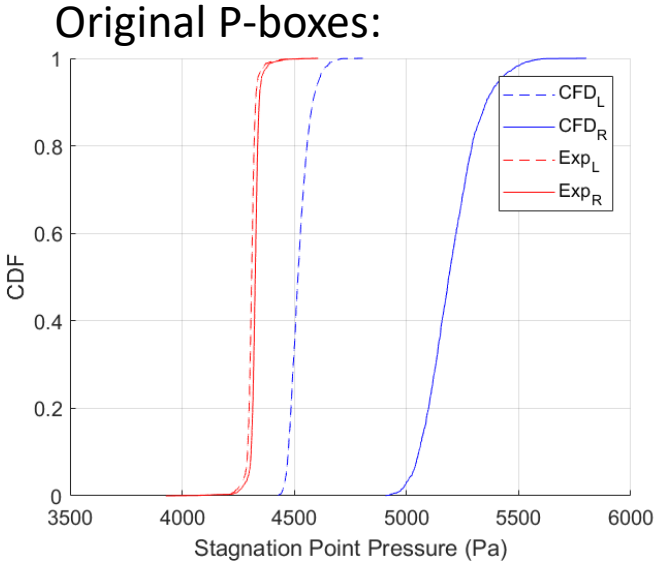


¹Brune, Andrew J., Thomas K. West IV, and Laura M. White. "Calibration probe uncertainty and validation for the hypersonic material environmental test system." *Journal of Thermophysics and Heat Transfer* 34.2 (2020): 404-420.

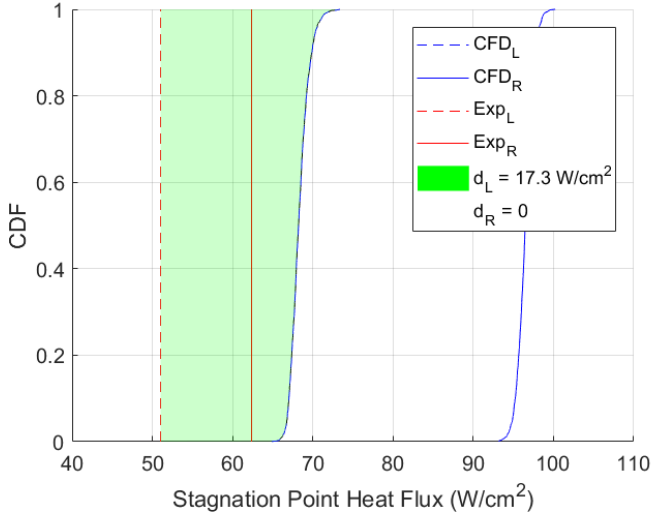
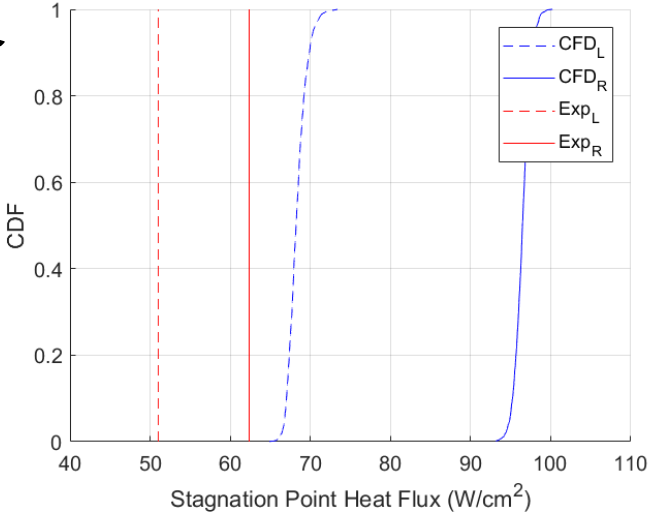
Validation Metric Results



Stagnation Point Pressure



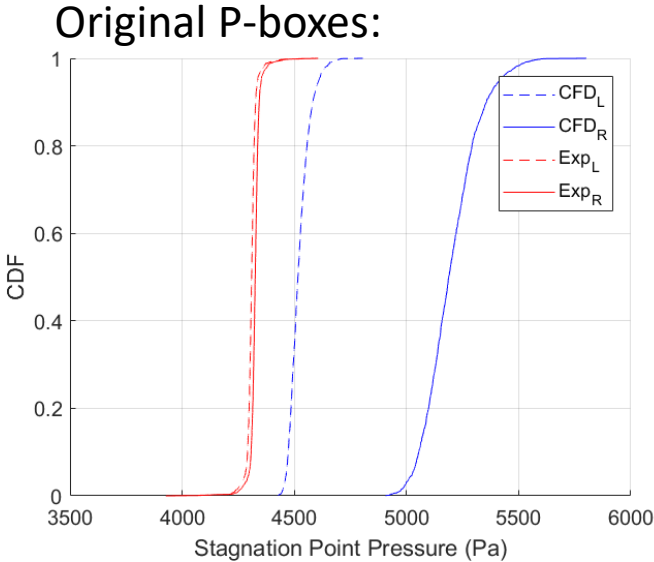
Stagnation Point Heat Flux



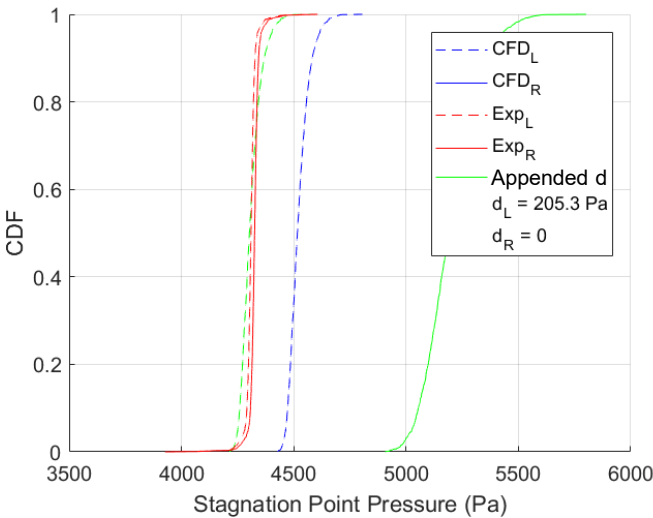
Validation Metric Results



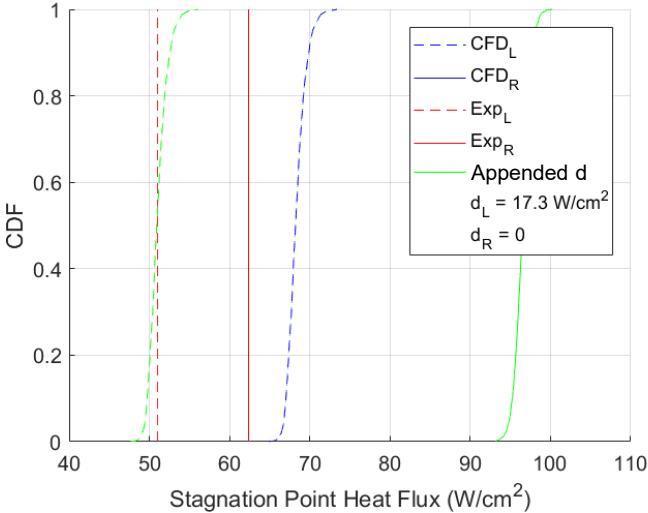
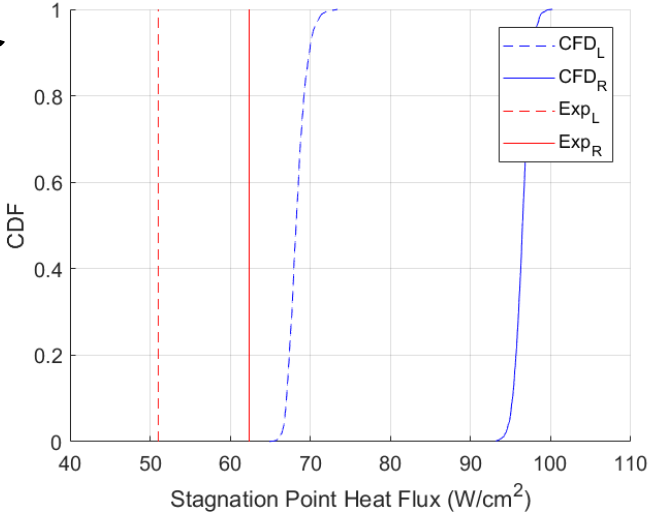
Stagnation Point Pressure



Validation Metric:



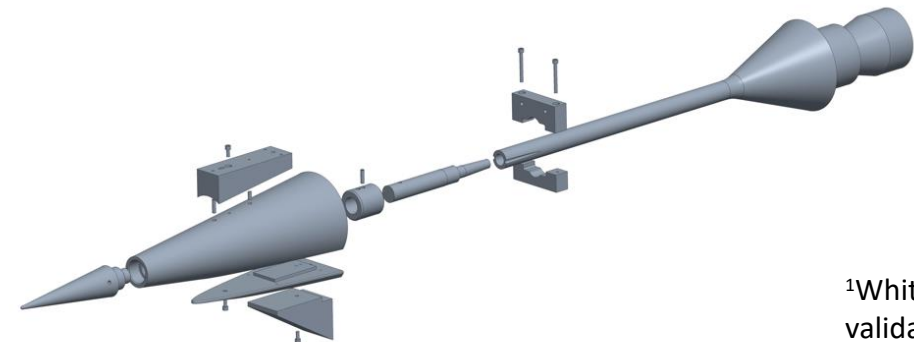
Stagnation Point Heat Flux



Applications



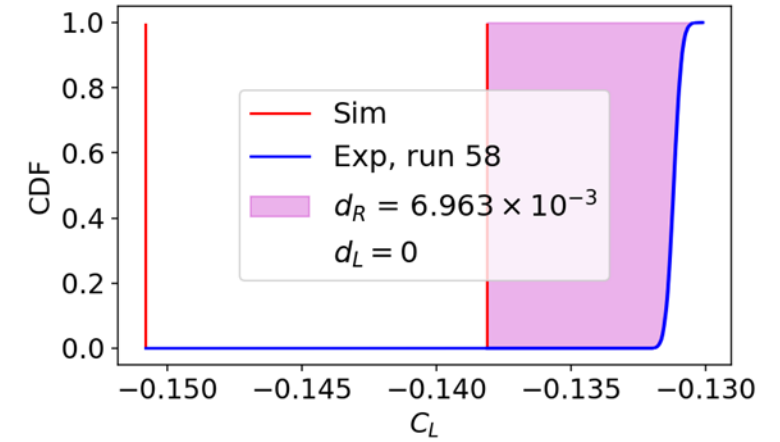
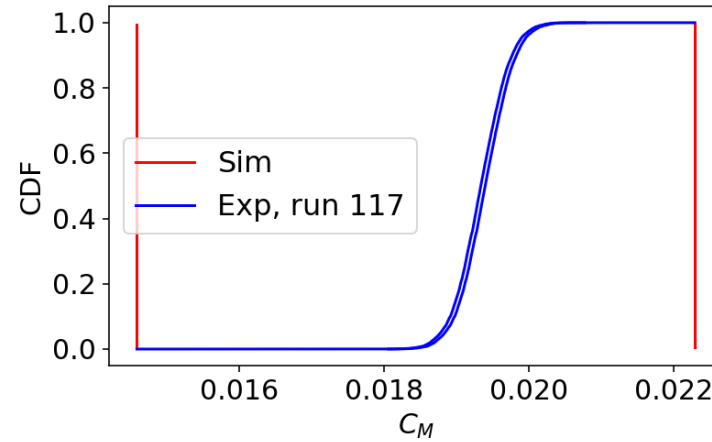
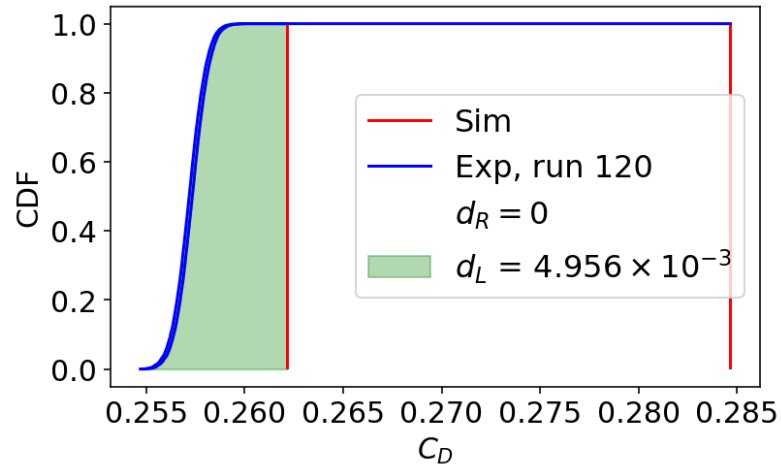
- Multi-configurable Cone-Slice-Flap Hypersonic Wind Tunnel Model¹
 - Experimental data: from the 20-inch Mach 6 Tunnel at NASA Langley Research Center
 - Simulation data: from NASA's Cart3D inviscid software and VULCAN-CFD Spalart-Allmaras (SA) and Menter Shear-Stress Transport (SST) models
 - Mixed and epistemic uncertainty



¹White, Laura, et al. "CFD validation study of a hypersonic cone-slice-flap configuration." *AIAA Scitech 2021 Forum*. 2021.

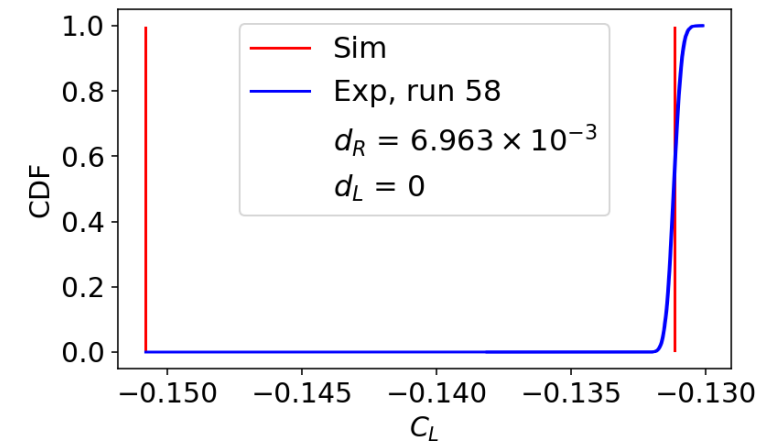
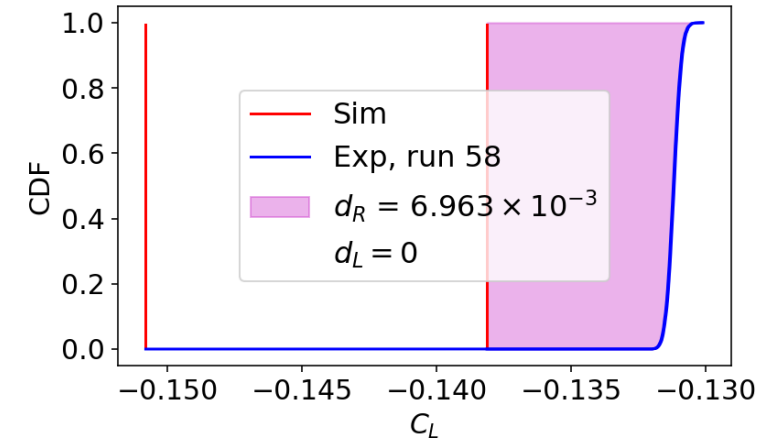
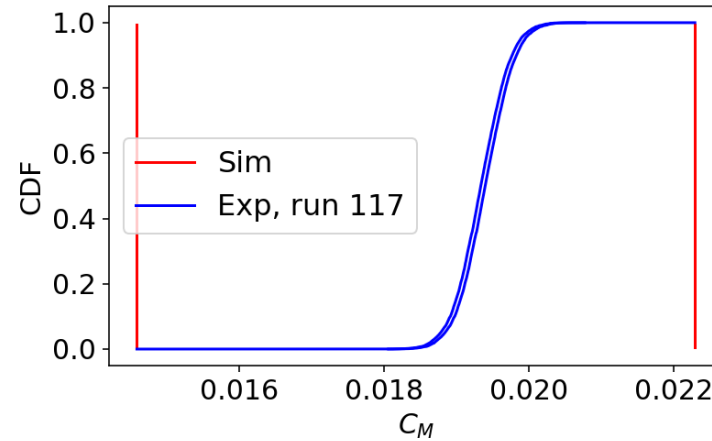
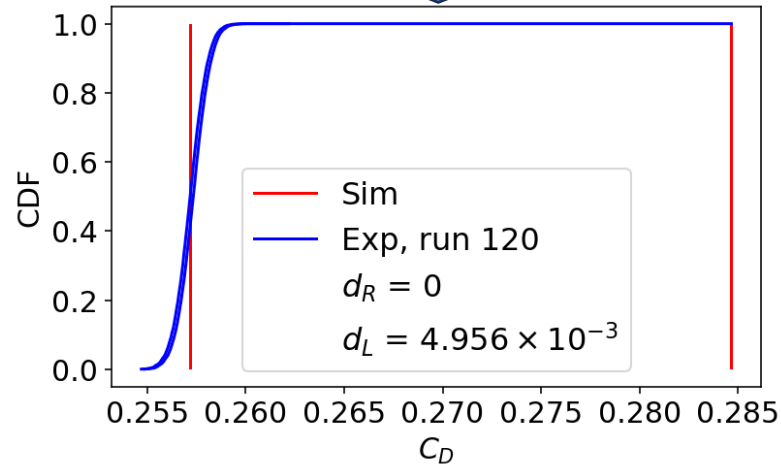
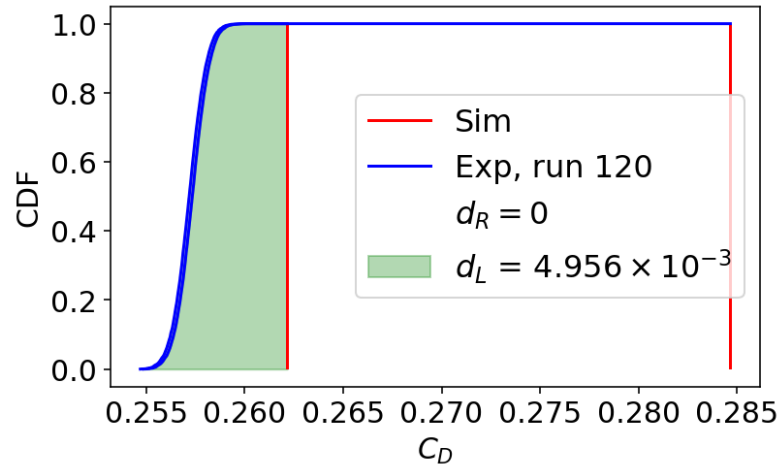


Validation Metric Results





Validation Metric Results





Conclusions

- Validation metrics quantify model-form uncertainty for systems with aleatory, epistemic, or mixed uncertainty.
- Standard area metrics can misrepresent disagreement in mixed-uncertainty cases.
- The alternative area metric more accurately captures left/right bound differences between simulation and experimental data.
- Applications demonstrate improved characterization of model-form uncertainty for hypersonic configurations and aerothermal heating.



Questions?