# A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing

Supervisor: Carina Newen

Hemalatha Sekar.

# Uncertainty Types

Two types of uncertainty in scientific computing are described

Aleatory uncertainty

Epistemic uncertainty

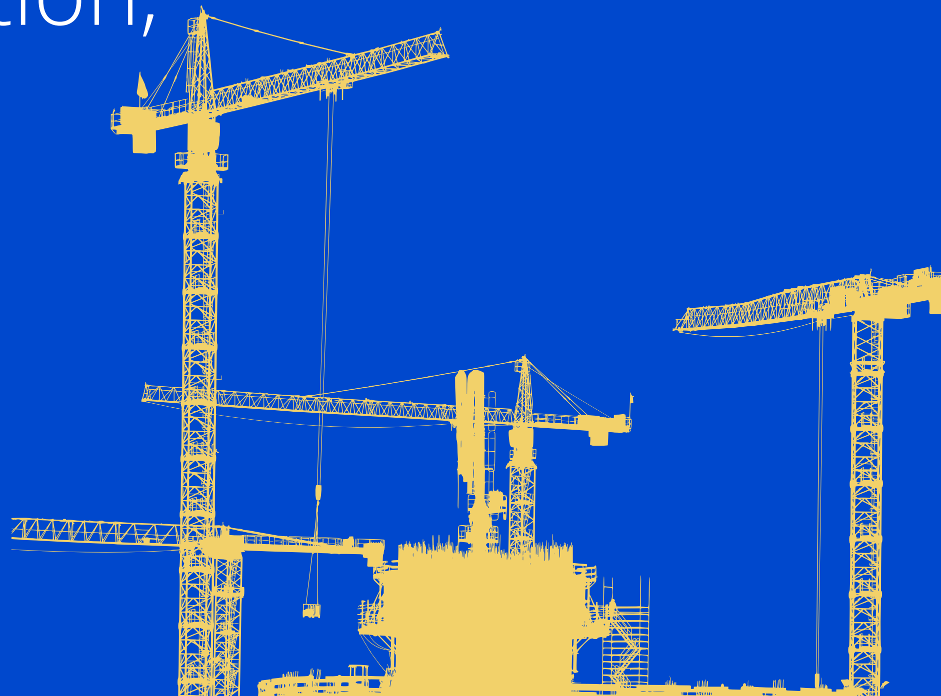
# Uncertainty Types

## Aleatory uncertainty

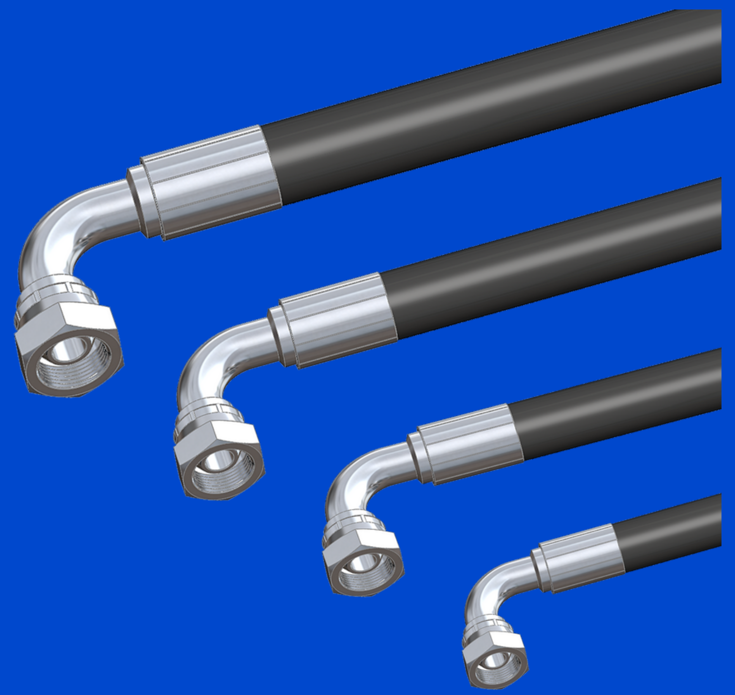
- Representative of randomness that differ for each iteration for the same experiment.
- Also known as irreducible uncertainty.
- Characterized either by PDF or CDF
- Uncertainity could be changed only if there is a change in manufacturing or quality control process.

## Epistemic uncertainty

- Lack of knowledge during the phase of analysis.
- Also known as reducible uncertainty.
- Characterized by interval.
- Reduced through conducting experiments, Improved numerical approximation, experts opinion etc.

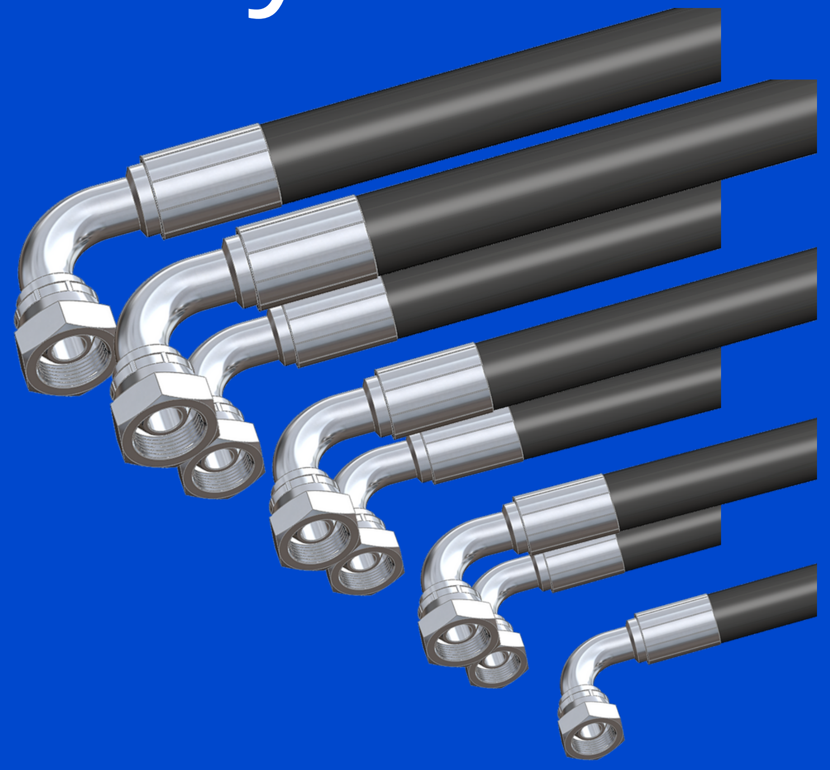


# Aleatoric + Epsidermic uncertainty



- Length of the part random variable -Aleatoric
- Not accuracy because of few samples from a population - Epsidermic

# Purely Aleatoric uncertainty



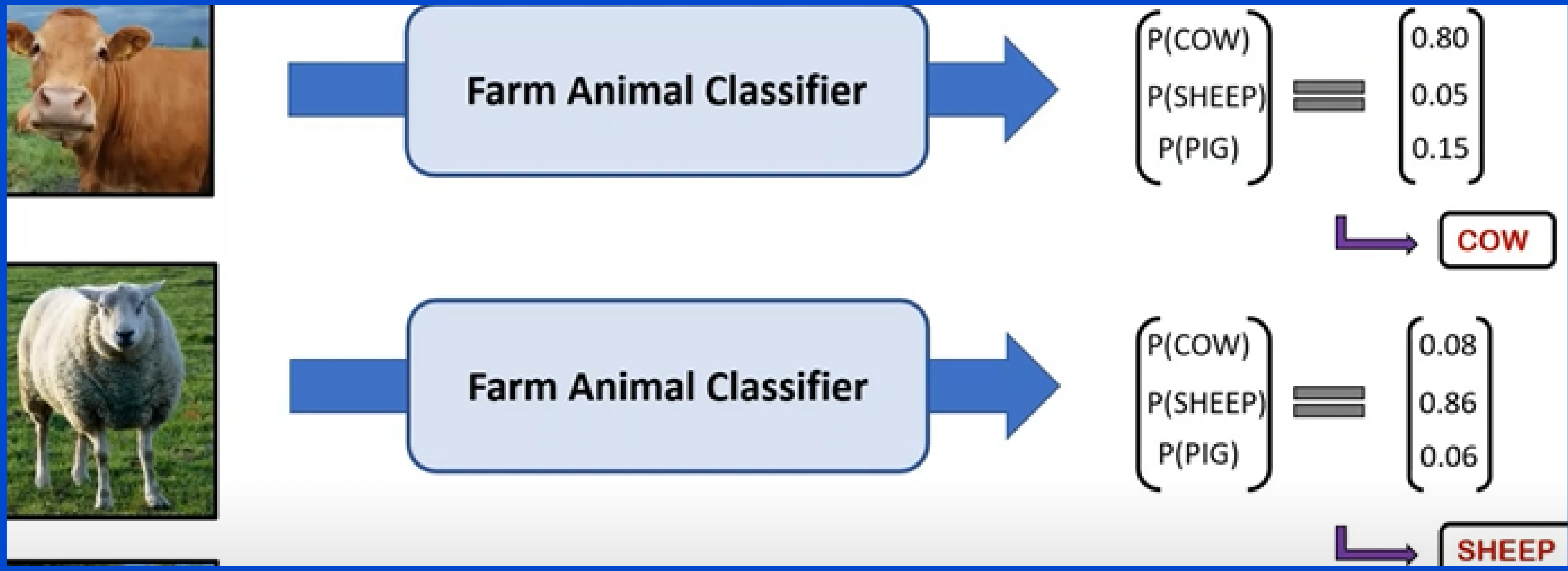
- With large number of samples, PDF is determined more accurately and precisely

# Sources of uncertainty

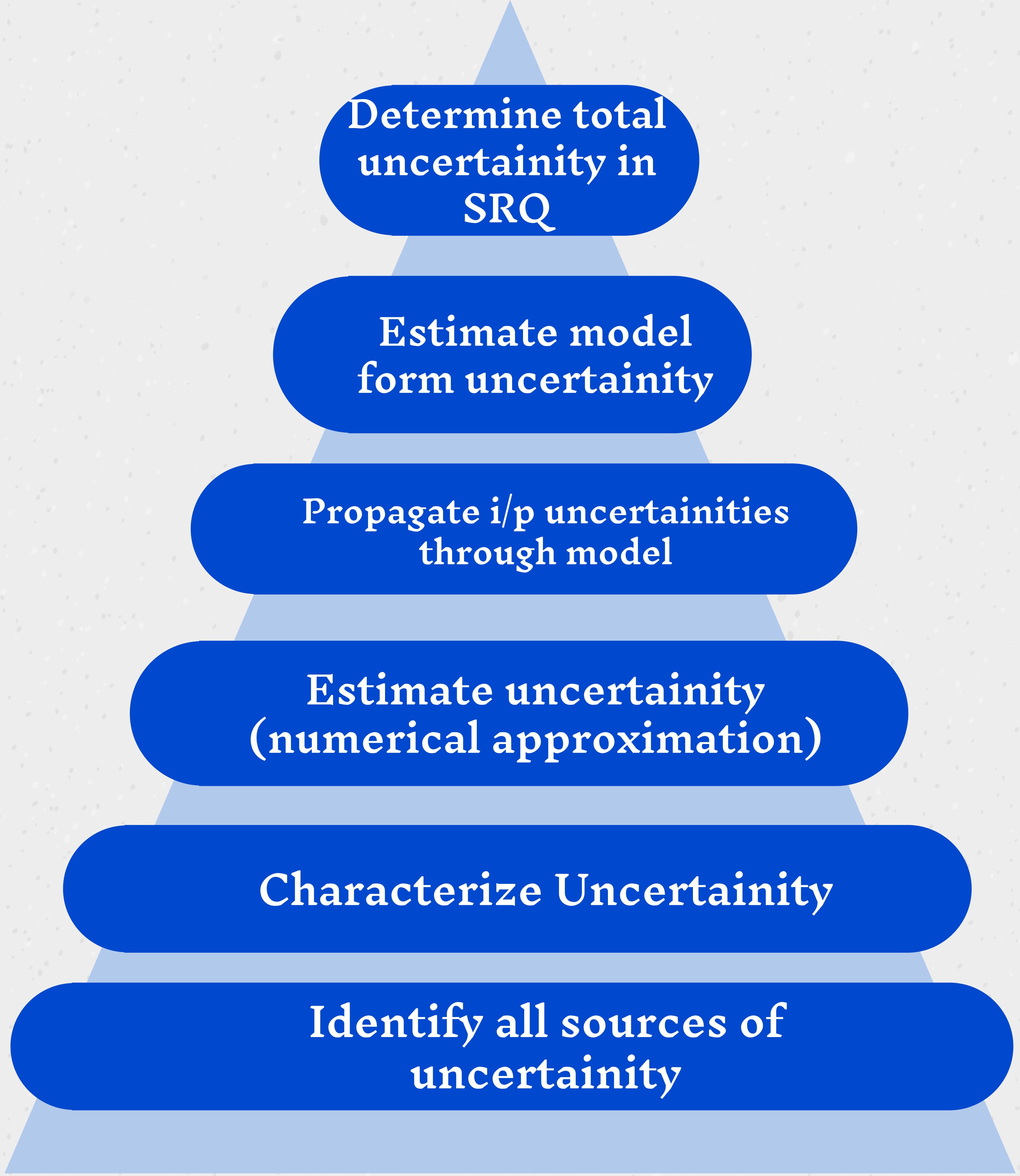
- Model Inputs
    - Parameters used in system
    - System surroundings
- Numerical approximations

  The iterative convergence error, discretization error, roundoff error and computer programming mistakes.

- Model form
    - Model validation.
    - Episdermic uncertainty.




Reference:https://www.youtube.com/watch?v=Osju20L6Z3I




# Uncertainty framework

The steps in Vertification, Validation and Uncertainity framework (Hypersonic nozzel flow)

# Hypersonic nozzel Flow

Arnold Engineering Development Complex crew members lower the NASA/Army Tiltrotor Test Rig into the 40-by 80-foot wind tunnel at Moffett Field in California. (Photo credit: U.S. Air Force)

- Replicates the air movement over aircrafts, vehicles and other objects.
- Engineers use it for further improvement in design, stability and cost effective etc.

Reference:https://boomsupersonic.com/flyby/post/what-is-wind-tunnel-testing

# Scenario

- Temperature < 80k ----> Condensation Occurs
- Decreases the flow quality with that high speed could damage the aircraft model.

# Hypothesis Stated

- To determine that the test section temperature should be greater than or equal to 80k with 95% confidence.

# Findings

- Test section static temperature of 85.3k is resulted through deterministic simulation which is 6% greater than the temperature specified.

Ref:https://imgur.com/gallery/qgxl5

# 1. Identify all sources of uncertainty

*Primary sources*

- Wind tunnel stagnation temperature
- Area downstream of the tunnel throat

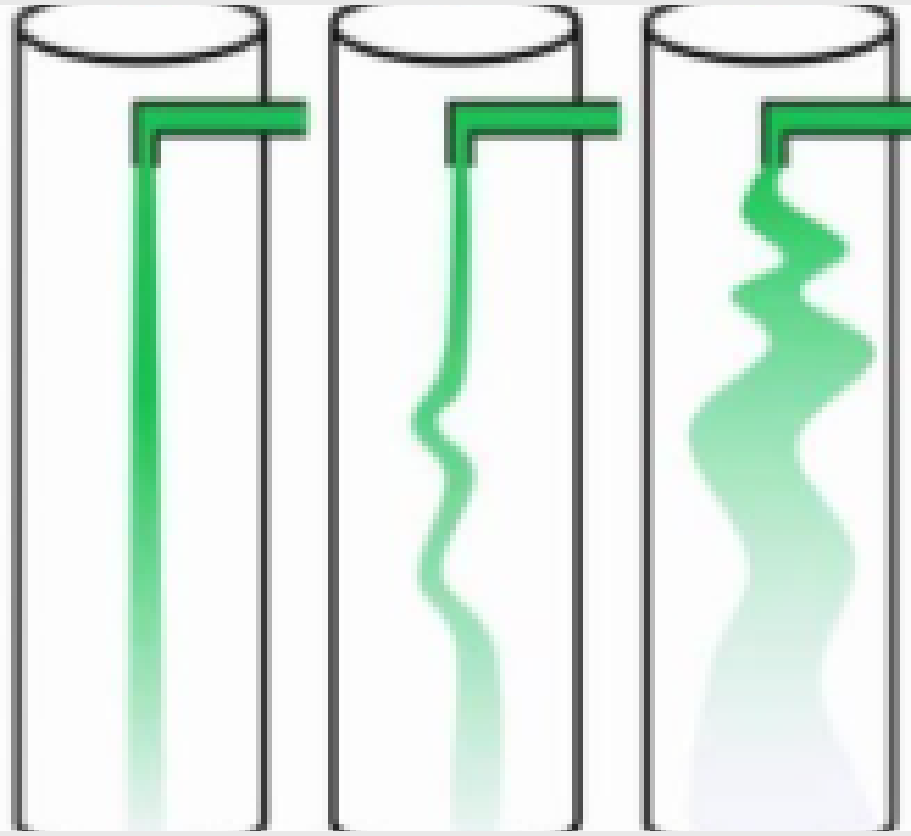
*Other sources*

- stagnation pressure
- Specific gas constant
- Ratio of specific heats
- Tunnel throat radius

NASA wind tunnel with the scale model of an airplane

Reference:https://en.wikipedia.org/wiki/Wind_tunnel

# 2. Characterize uncertainties



## *Wind tunnel stagnation temperature*

- It is an aleatory uncertainty
- Through run-to-run experiments, variations are normally distributed with mean stagnation temperature of 1200k with 3.33% coefficient of variation and 40k of standard deviation.

## Area downstream of the tunnel throat

- The wind tunnel **side-wall** boundary layer is not measured.
- The **state** of the boundary layer (laminar, transitional, or turbulent)is not known.
- Separate boundary layer simulations are performed(i.e fully laminar and turbulent)
  - Laminar boundary layer - 0.13m
  - Turbulent boundary layer - 0.14m

Reference:https://www.youtube.com/watch?v=5vGQFp_0C-AI

# 3. Estimate uncertainty due to numerical approximation

## Code Verification

- Removing bugs in the code.
- verification - the exact solution.

## Round-off and iterative error

- Simulations are advanced to achieve a steady state.
- Inserting the current solution of the discrete equations and evaluating the non-zero remainder.
- Iterative residuals are converged 12 orders of magnitude from their initial levels.

# Discretization error

- Estimated by running simulations on three systematically-refined meshes 128, 256, and 512 cells, the test section static temperature was found to be 85.307, 85.824, and 85.954 K, respectively.

## Order of convergence

$$\hat{p} = \frac{\ln\left(\frac{T_{coarse} - T_{med}}{T_{med} - T_{fine}}\right)}{\ln(r)},$$

To = 1200 K
rts = 0.14m
r = 2

Coarse temp - 85.954k
med temp - 85.824
fine temp - 85.307

^p = 1.99

## Richardson extrapolation:

Uses two fines grids to obtain an estimate of the value

$$\overline{T} = T_{fine} + \frac{T_{fine} - T_{med}}{r^p - 1} = 85.998 \text{ K}.$$

## Roache's Grid Convergence Index

uncertainty estimate due to discretization on the coarse mesh of 128 cells

$$U_{NUM} = U_{DE} \cong 1.25\left|T_{coarse} - \overline{T}\right| = 0.86 \text{ K}.$$

Ref:C.J. Roy, W.L. Oberkampf / Comput. Methods Appl. Mech. Engrg. 200 (2011) 2131–2144

# 4. Propagate input uncertainties through the model

**Aleatoric uncertainty**

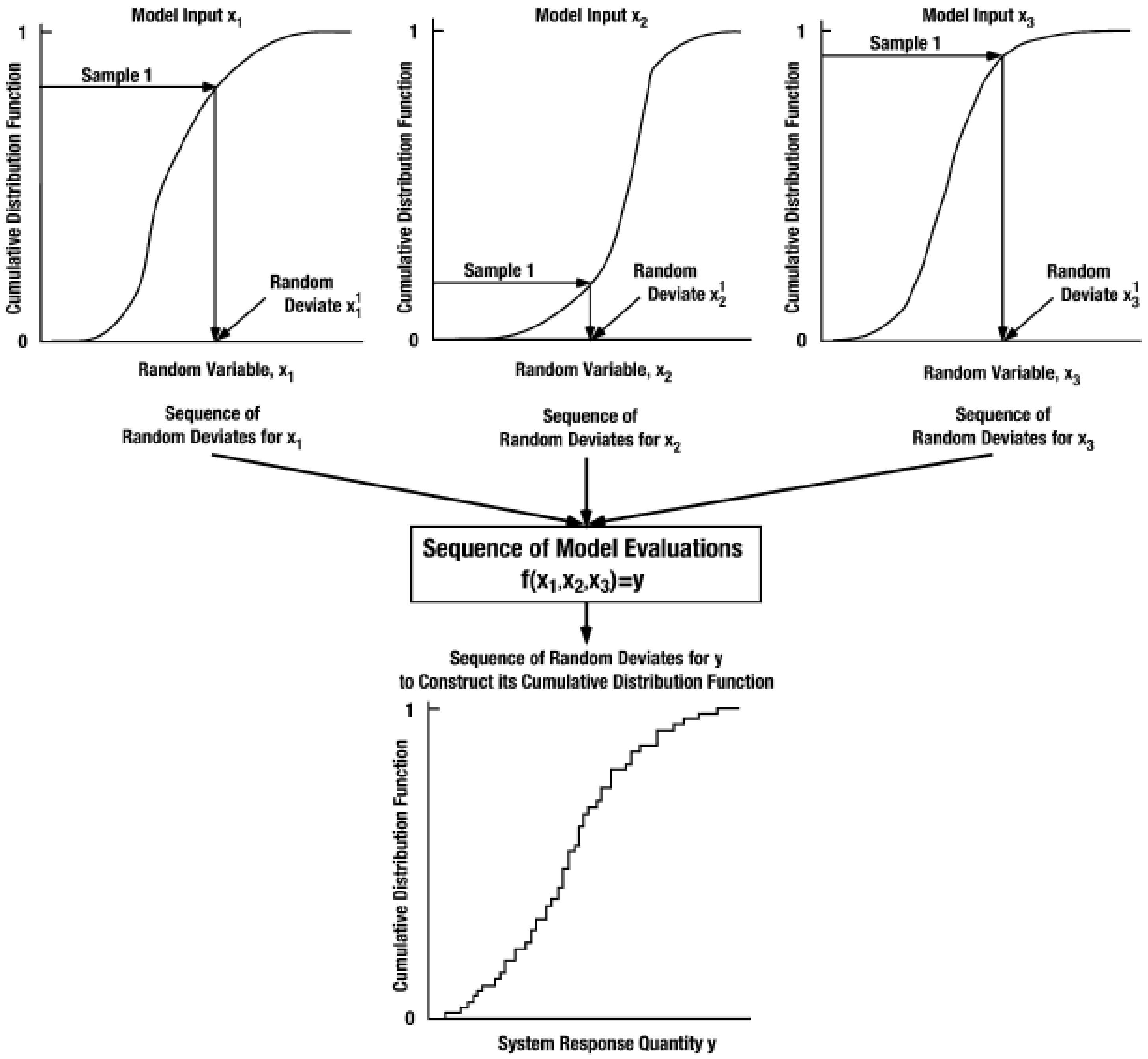
I/o samples (x1,x2,x3)

↓

Compute SRQ(y)

↓

Map the probability in CDF

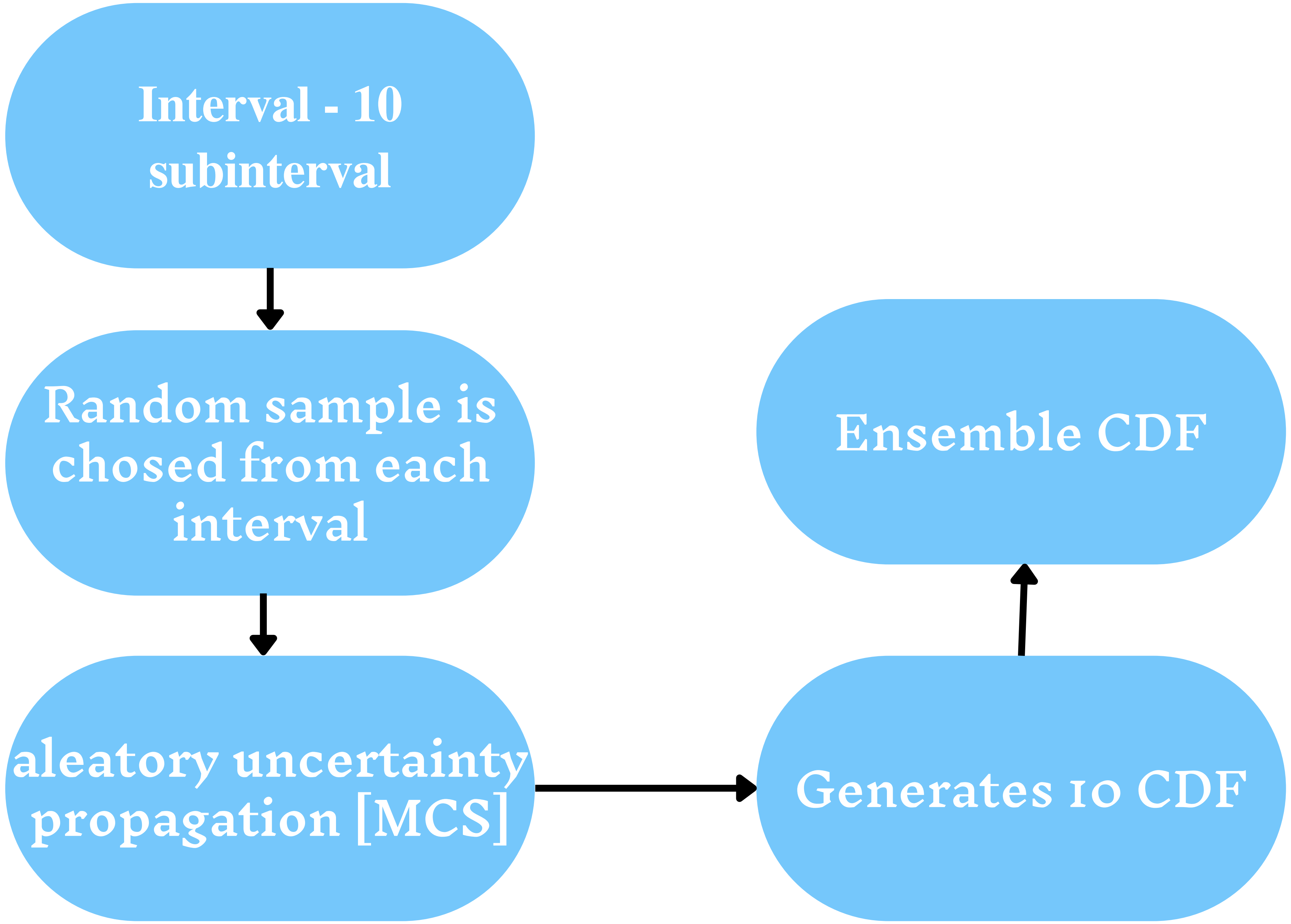


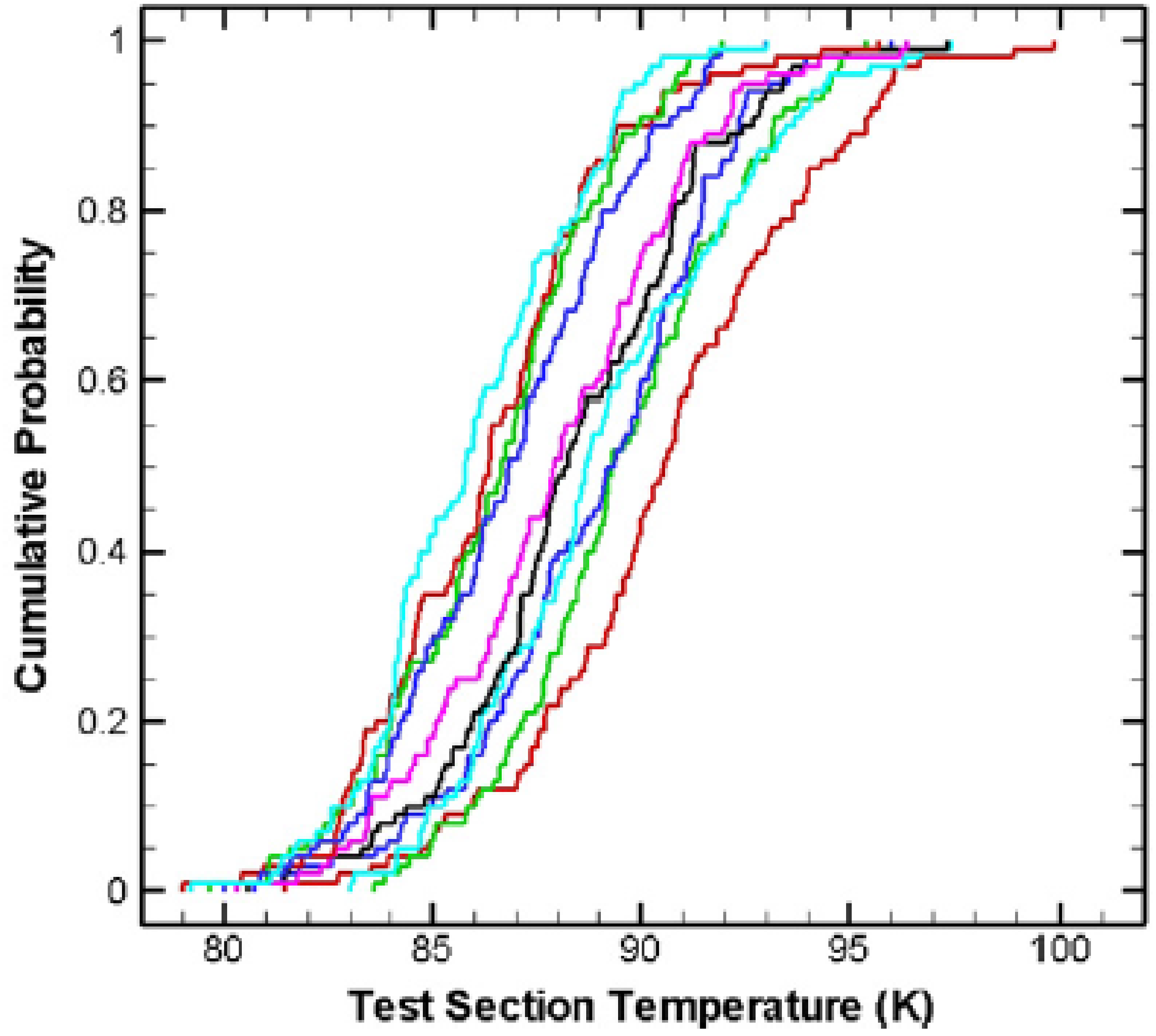

C.J. Roy, W.L. Oberkampf / Comput. Methods Appl. Mech. Engrg. 200 (2011) 2131–2144

## Monte Carlo sampling

Ref:C.J. Roy, W.L. Oberkampf / Comput. Methods Appl. Mech. Engrg. 200 (2011) 2131–2144

# Epistemic uncertainty




Ensembled CDF

**Latin hypercube sampling**

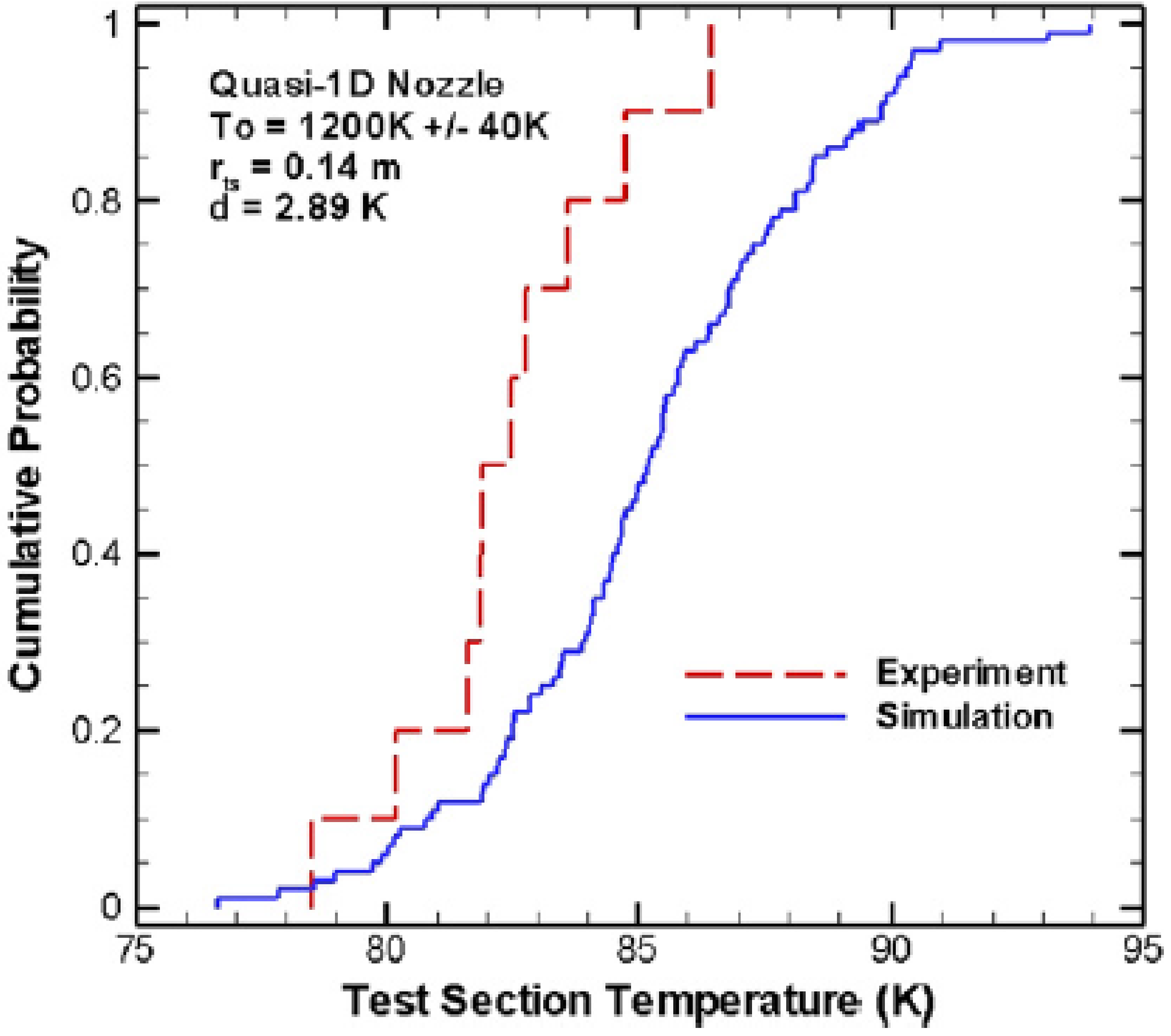
widest extent used to construct P- box

Ref: C.J. Roy, W.L. Oberkampf / Comput. Methods Appl. Mech. Engrg. 200 (2011)

# 5. Estimate model form uncertainty

- Consider an example, for stagnation pressure of 20 MPa, the area validation metric is unknown. Provided three random validation experiment outcomes as sample for stagnation pressure  7MPa,10MPa, 12MPa.

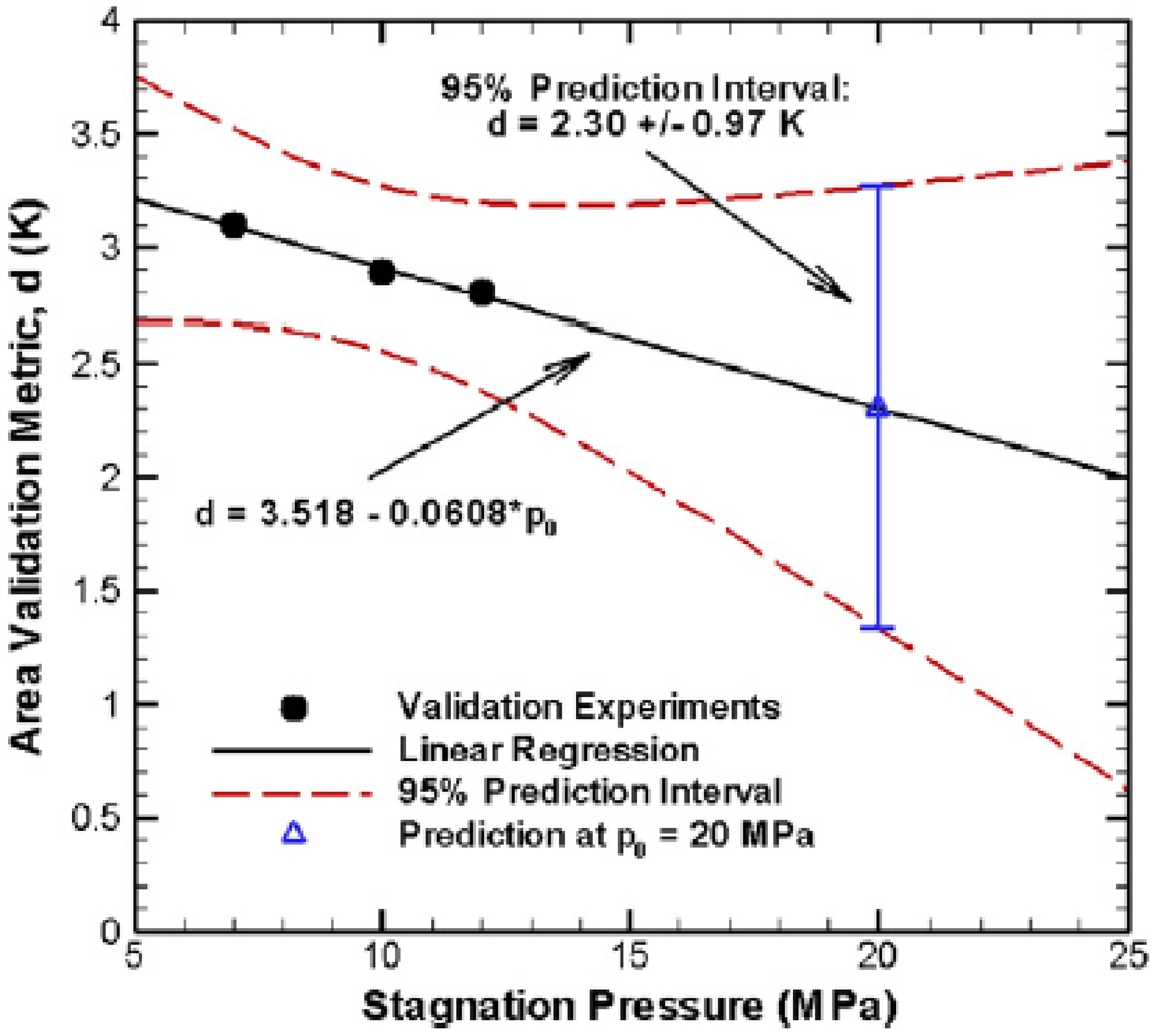
1. Ten synthetic measurements of the SRQ (test section static temperature) are chosen to be: SRQEXP = [78.5, 80.2, 81.6,81.8, 81.9,82.5, 82.7,83.6, 84.7,86.4] K
2. Propagating the input uncertainty (aleatory and epistemic) through the model to form CDF.
3. Retrieving the CDF formed from experimental observation.
4. Area between these two CDF is known to be the area validation metric d = 2.89K.




Computation of Area validation metric d

Ref:C.J. Roy, W.L. Oberkampf / Comput. Methods Appl. Mech. Engrg. 200 (2011) 2131–2144

5. Similarly,
7Mpa - 3.1k
10MPa - 2.89k
12Mpa - 2.8k are computed.
6. Compute Simple Linear Regression from the obtained value considering the stagnation temperature as an independent variable, and area validation metric as the dependent variable
y^ = 3.518 - 0.0608xk.
7. Compute prediction interval




$$\hat{y} \pm t_{\alpha/2,N-d}\, s \sqrt{1 + \frac{1}{N} + \frac{N(x-\bar{x})^2}{N\sum_N x_i^2 - \left(\sum_N x_i\right)^2}}.$$

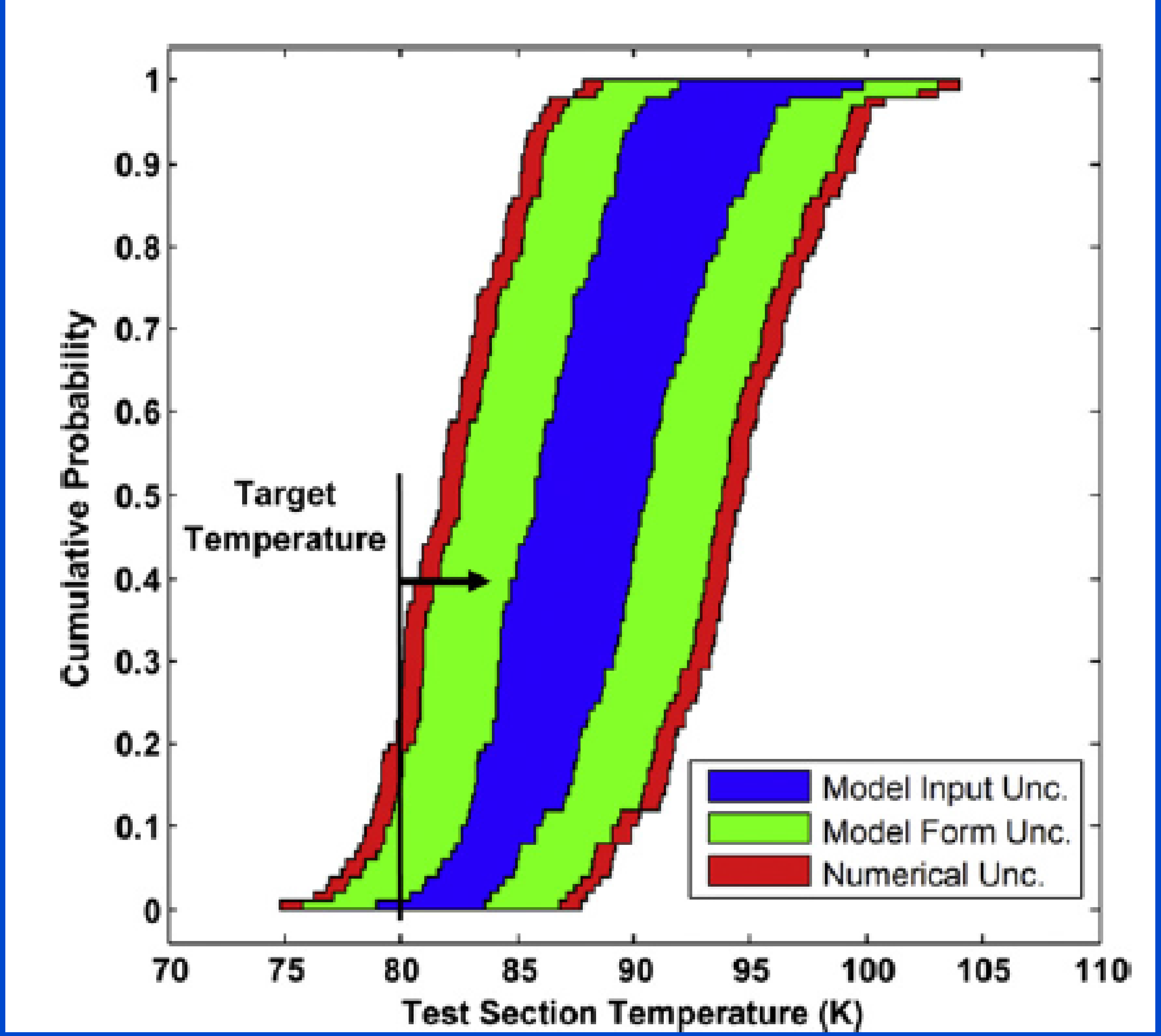
N - number of validation experiments[N = 3]
x - stagnation pressure [x=20MPa]
d - degrees of freedom [d=2]
s - sqrt.MSE [s=0.02433k]

8. The resulting 95% prediction interval for the area validation metric at p = 20 MPa is d = 2.30 ± 0.97 K [d=3.27k]

Ref:C.J. Roy, W.L. Oberkampf / Comput. Methods Appl. Mech. Engrg. 200 (2011) 2131–2144

# 6. Determine total uncertainty in the SRQ

1. The p-box is determined by propagating aleatory and epistemic uncertainties model inputs through the model in condition (p = 20 MPa).
2. Append the area validation metric, i.e., d = 3.27 K, to the left and right sides of the p-box.
3. Uncertainty due to numerical approximation UNUM = 0.86 K is appended to the left and right sides of the p-box.
4. There is a 25% chance that the test static temperature would fall below 80k at 95% CI.




Nondeterministic prediction of uncertainty

Ref:C.J. Roy, W.L. Oberkampf / Comput. Methods Appl. Mech. Engrg. 200 (2011) 2131–2144

# 6. Conclusion

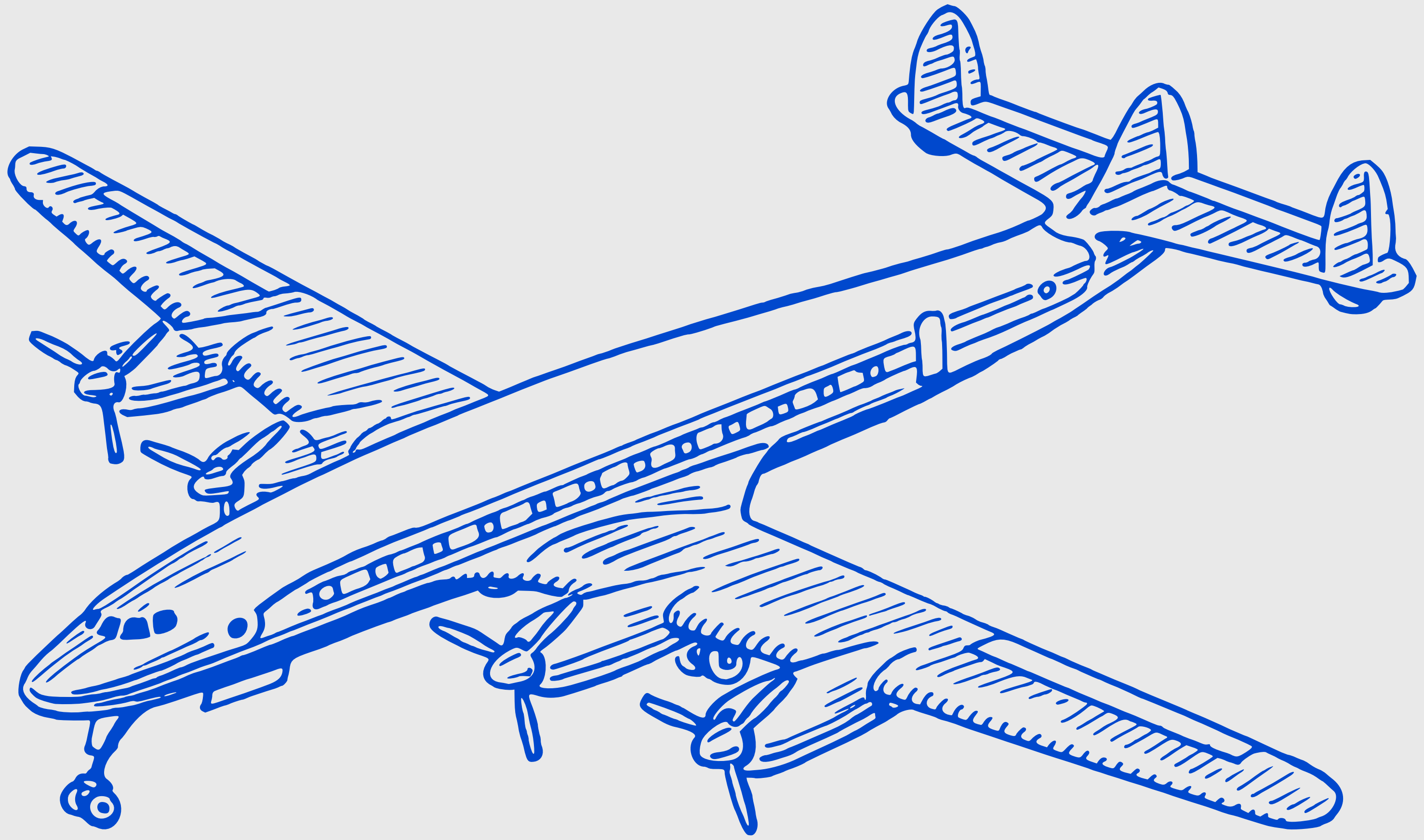
- This predicted uncertainty is precisely shown to the decision-makers to avoid putting customers or environments at risk from uncertainties.
- It separates the aleatory and epistermic uncertainty and focus on numerical solution error and model form uncertainty directly.

## When it can be used? :

When the decision-makers find the observations or system response quantities to be inaccurate.

## Where it can be used? :

Predictions of high consequences of the system (human lives, national security, safety measures)



# Thank you

*HEMALATHA SEKAR*
*hemalatha.sekar@tu-dortmund.de*