



Automation of the Uncertainty Quantification Process Based on Probability Boxes with DAKOTA

Robert A. Baurle and Andrew Norris
Hypersonic Airbreathing Propulsion Branch
NASA Langley Research Center





- **Motivation**
- **Uncertainty Quantification Framework**
- **Description of the DAKOTA software**
- **Details on how the UQ process is automated**
- **Application of the UQ framework**
 - Parameter sensitivity study for a scramjet combustor
- **Summary**



- **Scramjet engine design/analysis relies heavily on CFD.**
 - Trend is an ever increasing reliance on numerical data
 - Limitations (environment, scale, and duration) of ground test facilities
 - Advanced diagnostics hindered by harsh scramjet flow environment
 - Increased pressure to simulate more and test less
- **How good are these CFD simulations?**
 - Confidence (or lack there of) in the results is rarely communicated
 - Instead a deterministic simulation strategy is the current SoP
 - Typical goal is to maximize the coverage of the design space (e.g., increase resolution of aero-databases)
 - Can this culture continue to be justified?
- **A UQ-centric approach to CFD can provide “Risk Informed” decisions.**
 - Confidence in simulation data should accompany the raw data

Barriers Preventing Incorporation of UQ



- **Why is UQ not commonly a standard practice in the CFD workflow?**
 - 1) The CFD applicationist is often not well versed with UQ methods
 - UQ is its own complex scientific discipline, so standard “Joe” engineer may not know how best to incorporate it
 - 2) Workload of the CFD applicationist is inevitably increased
 - Preparation of “n” simulations instead of 1 deterministic one
 - 3) Increase computational cost
 - 1 deterministic simulation replaced with “n” simulations
- **Through automation, the first 2 items can be **eliminated** as barriers.**
- **The last barrier is addressed by the flexibility offered in the statistics package used.**
 - Goal is to keep “n” as small as possible (e.g., surrogate modeling)



- **Desired UQ framework attributes:**
 - Nonintrusive (no modifications required to CFD tool)
 - Computationally inexpensive (in a relative sense)
 - Retain the core features emphasized by the UQ community
- **Probability Box approach**
 - Aleatoric and Epistemic uncertainty sources are treated separately.
 - Aleatoric uncertainties are irreducible and inherently random, and are typically described by PDFs.
 - Epistemic uncertainties are systematic uncertainties that are a result of inadequate knowledge ... these sources of uncertainty can in principle be reduced if further knowledge is gained.
 - Reference: Roy and Oberkampf, "Computational Methods in Applied Mechanics and Engineering," Vol. 200:25, 2011, pp. 2131-2144.

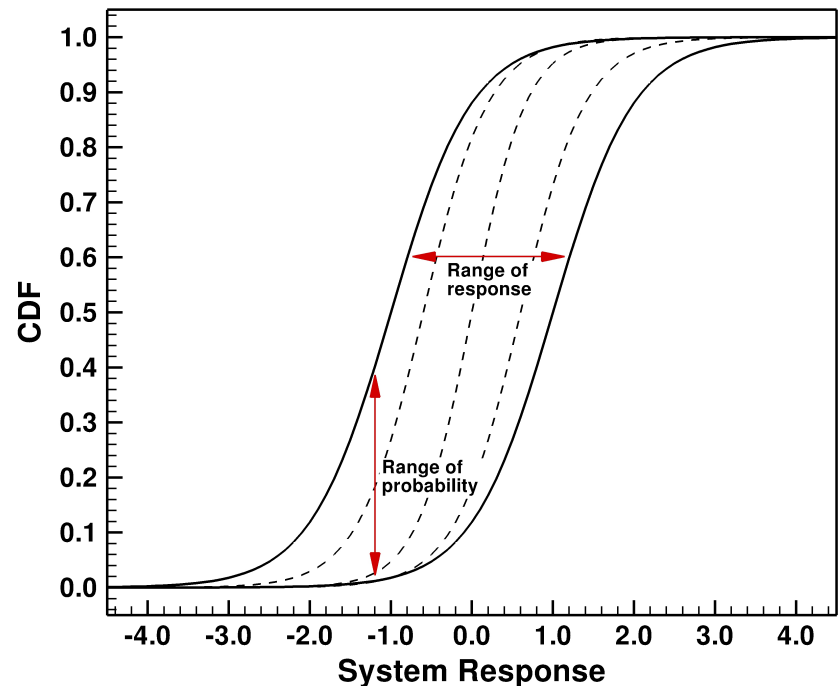
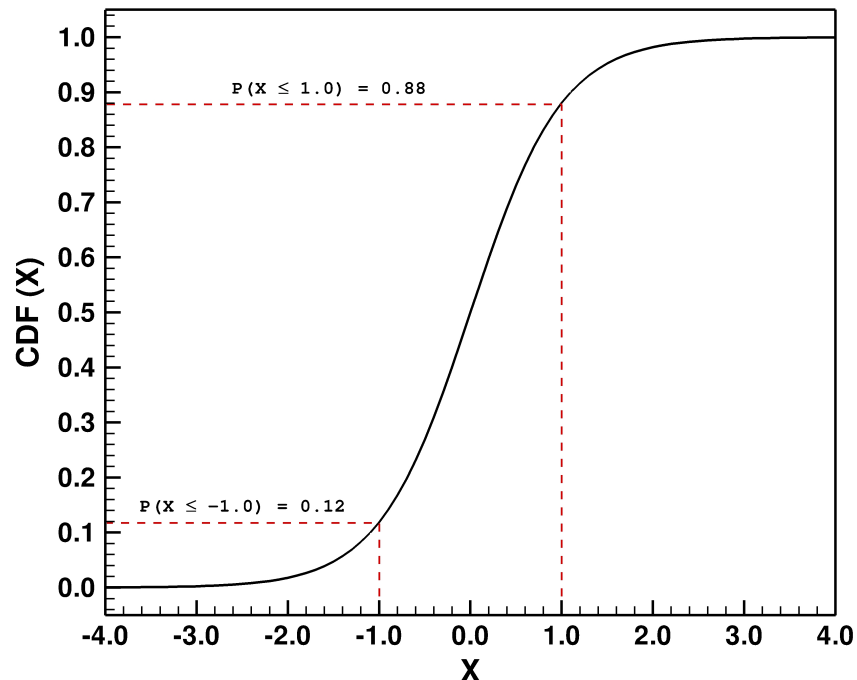
UQ via Probability Boxes



Steps in the UQ process:

- (1) Identify and characterize each uncertainty
- (2) Prioritize the sources of uncertainty
- (3) Propagate the uncertainties through the “computational model” to extract the QoI

```
Do n = 1, Ne  
  Do m = 1, Na  
    evaluate QoI  
  End Do  
End Do
```





- **DAKOTA is a package developed out of SNL designed to provide robust software to enable advanced parametric analyses for the purposes of**
 - design exploration
 - model calibration
 - risk analysis
 - quantification of margins and uncertainty
 - optimization
- **The toolkit provides a flexible, extensible interface to perform any combination of the tasks above given data obtained from scientific simulation codes**
- **The software is publically available (GNU LGPL) and can be downloaded from: <https://dakota.sandia.gov/>**



- **Automated process to incorporate UQ in the CFD workflow:**
 - 1) Setup and simulate a baseline CFD case
 - Provides an opportunity to make sure that algorithmic parameters are set in a robust and efficient manner
 - Allows the additional simulations required for UQ to use the results from this baseline simulations as a “warm” start
 - 2) Create a small input file (5 required inputs) to govern the high-level UQ execution process
 - 3) Choose a DAKOTA template for your specific UQ problem
 - Pure aleatoric template
 - Pure epistemic template
 - Mixed epistemic/aleatoric template
 - 4) Execute “uq_setup” to setup the UQ simulation directory structure and VULCAN-CFD input files for each simulation required by your UQ problem
 - 5) Perform the CFD simulations
 - 6) Execute “uq_final” to process the CFD outputs (responses) of interest and generate the P-box data

UQ Driver Step



- **Example “uq_input.dat” file to control the UQ process:**

```
#  
# Input file for coupling VULCAN-CFD with Dakota  
#  
VULCAN_INPUT          rc22.inp  
CASE_NAME              RC22  
CASE_PATH              Cases  
DAKOTA_TEMPLATE        dakota_UQ  
POST_CODE              uq_vulout  
GCI_REFINE             2.0
```

DAKOTA Template Setup (aleatoric example)



```
environment
  tabular_data
    tabular_data_file = 'dakota_table.dat'
    top_method_pointer = 'UQ_ALEATORIC'

method
  id_method = 'UQ_ALEATORIC'
  sampling
    sample_type lhs
    samples = 1000
    seed = 17
    distribution cumulative
    model_pointer = 'SURROGATE'
  output verbose

model
  id_model = 'SURROGATE'
  surrogate global
    dace_method_pointer = 'DACE'
    reuse_points all
    polynomial quadratic
  ‡ gaussian_process surpack
  ‡ neural_network
    metrics = 'root_mean_squared' 'rsquared'

method
  id_method = 'DACE'
  sampling
    sample_type lhs
  ‡ samples = 50
    seed = 17
    model_pointer = 'DACE_MODEL'
  output verbose

model
  id_model = 'DACE_MODEL'
  single
```

```
variables
  ‡ normal_uncertain = 1
  ‡ means = 5637.5571
  ‡ std_deviations = 169.12
  ‡ descriptors = 'STATIC_PRESSURE'
  ‡ uniform_uncertain = 2
  ‡ lower_bounds = 277.6 0.49
  ‡ upper_bounds = 299.8 0.51
  ‡ descriptors = 'STATIC_TEMPERATURE' 'MACH_NO.'

interface
  analysis_drivers = 'dakota_script'
  fork
    file_tag

responses
  ‡ response_functions = 2
  ‡ descriptors = 'TOTAL_PRESSURE_LOSS' 'F<sub>x</sub>_[N]'
  no_gradients
  no_hessians
```

- Any VULCAN-CFD input parameter can be specified as an uncertain variable
- Available VULCAN-CFD output properties:
 - Any property from the force and moment history file (vulcan.ifam_his_1.tec file)
 - Any property written to the screen by the VULCAN-CFD postprocessor
 - Any property output by the propulsion system performance extraction utility (perf_ext)

DAKOTA Template Setup (epistemic example)

```
environment
  tabular_data
    tabular_data_file = 'dakota_table.dat'
    top_method_pointer = 'UQ_EPISTEMIC'

method
  id_method = 'UQ_EPISTEMIC'
  sampling
    sample_type lhs backfill
    samples = 1000
    seed = 17
    model_pointer = 'SURROGATE'
    output verbose

model
  id_model = 'SURROGATE'
  surrogate global
    dace_method_pointer = 'DACE'
    reuse_points all
    polynomial quadratic
  ‡ gaussian_process surfpack
  ‡ neural_network
  metrics = 'root_mean_squared' 'rsquared'

method
  id_method = 'DACE'
  sampling
    sample_type lhs backfill
  ‡ samples = 50
    seed = 17
    model_pointer = 'DACE_MODEL'
    output verbose

model
  id_model = 'DACE_MODEL'
  single

variables
  ‡ continuous_interval_uncertain = 2
  ‡ lower_bounds = 0.7 0.7
  ‡ upper_bounds = 0.9 0.9
  ‡ descriptors = 'TURB._SCHMIDT_NO.' 'TURB._PRANDTL_NO.'
  ‡ discrete_uncertain_set
  ‡ string = 1
  ‡ elements_per_variable = 2
  ‡ elements = 'WILCOX-1998' 'WILCOX-2006'
  ‡ descriptors = 'TURBULENCE_MODEL'

interface
  analysis_drivers = 'dakota_script'
  fork
    file_tag

responses
  ‡ response_functions = 2
  ‡ descriptors = 'COMBUSTION_EFFICIENCY' 'Qdot_[W]'
  no_gradients
  no_hessians
```

- **Note that categorical variables are allowed**
 - Categorical variables can be strings, reals, or integers
- **Substantial computational cost savings can be realized if real-valued uncertain variables can be cast as a categorical variable**
 - Possible if the outputs of interest (responses) vary monotonically with the uncertain variable

Prepare CFD Simulations for UQ



- **Execute “uq_setup” which performs the following tasks:**
 - 1) Extracts the following information from the UQ driver file (uq_input.dat)
 - The directory (CASE_NAME) that contains the baseline VULCAN-CFD input file and all files to perform the simulation (grid file, database files, restart files, etc.)
 - The directory (CASE_PATH) that will house all of the specific CFD simulation directories required by the UQ problem
 - The name of the DAKOTA input template created for the UQ
 - 2) Creates the actual DAKOTA input file (given the template) needed by DAKOTA to generate the random samples of each UQ variable
 - 3) Executes DAKOTA and extracts the random sample sets that were generated and creates VULCAN-CFD input files from the baseline input file
 - 4) Each of these VULCAN-CFD input files (and all supporting files from the baseline simulation directory) are copied to independent (sequentially numbered) directories inside the “CASE_PATH” folder.
- **At this point, all of the VULCAN-CFD simulations are setup and ready for execution**

Perform the CFD Simulations



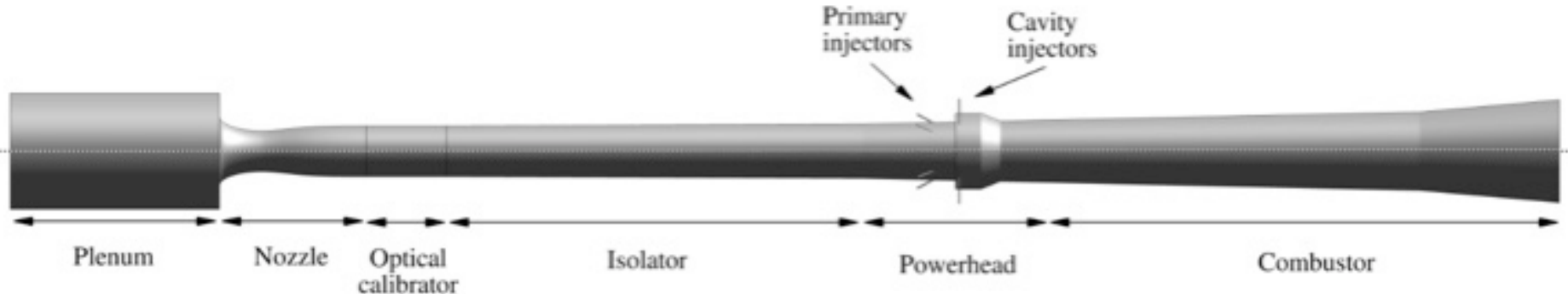
- **This step proved difficult to automate in a “hands-off” fashion due to the following unknowns:**
 - 1) Specific HPC batch processing environment used (PBS, SLURM, ...)
 - 2) Wall clock time required to complete each simulation
 - Are enough processors available to allow all simulations be completed within 1 batch submittal?
 - Are enough processors available for even 1 simulation be completed within 1 batch submittal?
- **Based on this observation, the decision was made to loosely couple VULCAN-CFD with DAKOTA in the UQ process**
 - This offers the highest level of flexibility for the user based on the computational resources available to them
 - The CFD applicationist is likely to be very knowledgeable with maximizing throughput through the HPC machines they have access to

Process CFD Outputs to Quantify Uncertainty



- **Execute “uq_final” which performs the following tasks:**
 - 1) Extracts the following information from the UQ driver file (uq_input.dat)
 - The directory (CASE_PATH) that houses all of the specific CFD simulation directories required by the UQ problem
 - The name of the DAKOTA input template created for the UQ
 - The name of the post-processing code used to extract the desired responses
 - The GCI refinement factor for estimating numerical error bounds [OPTIONAL]
 - 2) Executes Tecplot (if available) in batch mode and outputs convergence information for all of the CFD simulations to an image file.
 - L2 norm of the residual errors and the mass flow rate error are output
 - 3) Executes the post-processing code/utility to extract the desired CFD output (responses) from each CFD simulation
 - 4) Creates the actual DAKOTA input file (given the template) needed by DAKOTA to perform the P-Box UQ process
 - 5) Executes DAKOTA and extracts the CDFs used to build the P-Box
- **At this point, the UQ data (in the form of a P-box) have been output to files for convenient post-processing.**

- A scramjet combustor simulation (RC22 round direct-connect geometry) was chosen for demonstration purposes



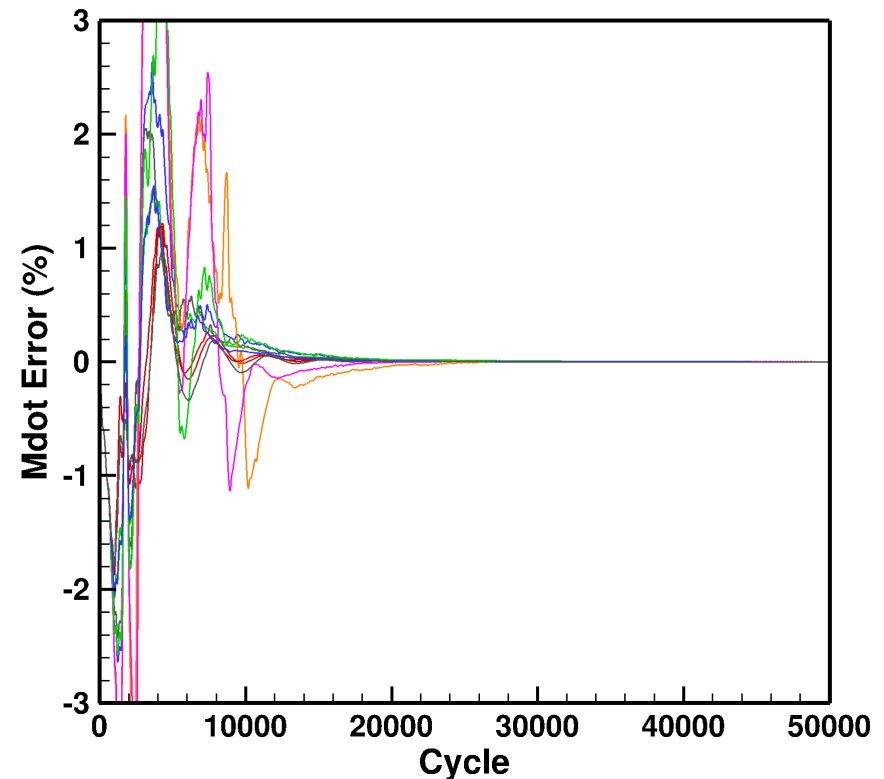
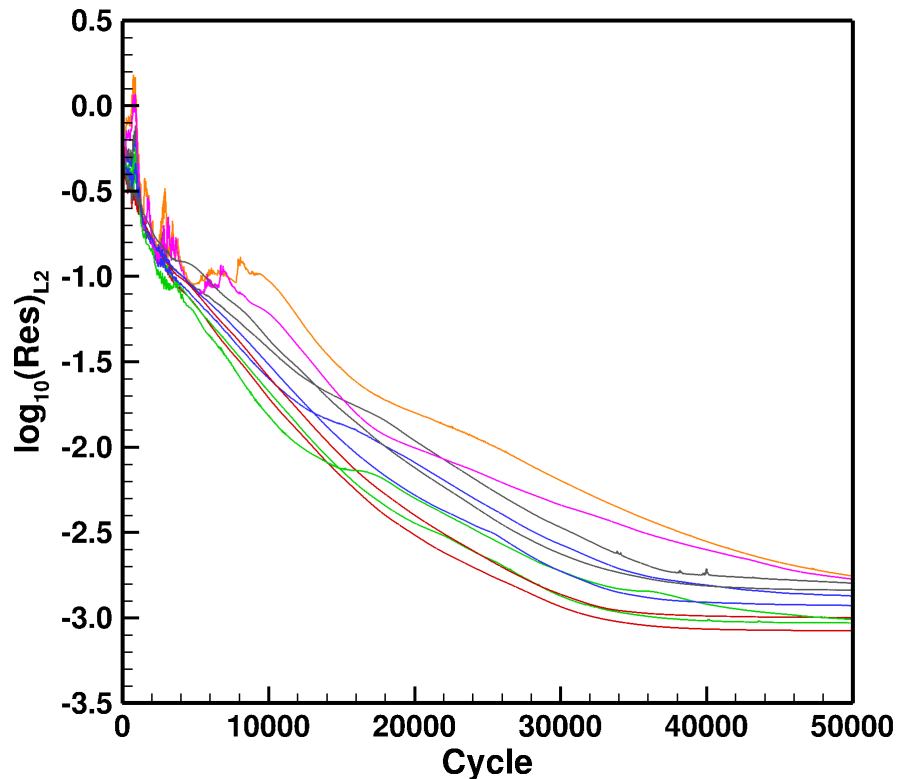
- A Mach 2.2 operating point was chosen with a global ER = 0.5
 - 92% through primary injector ports, 8% through cavity injection ports
- The Menter-BSL model was chosen as the turbulence model with a 6-species, 3-step ethylene finite rate model chosen for chemistry
- For this demonstration, a parameter sensitivity study is performed to gauge the importance of several aspects of the simulation
 - Uncertainty in Sc_t value (epistemic variable)
 - Uncertainty in heat resistance coating thickness (aleatoric variable)
- Two grid levels were used to assess the numerical errors

- **UQ initiated from a baseline setup for this RC22 simulation**
- **Included GCI option in the UQ driver input file**
- **Chose a mixed epistemic/aleatoric DAKOTA template**
 - 1 epistemic variable (continuous interval uncertain variable)
 - Turbulent Schmidt number ($0.5 < Sc_t < 0.7$)
 - 1 aleatoric variable (uniform uncertain variable)
 - Powerhead heat resistance coating thickness variability ($0.01'' < h < 0.02''$)
 - 10 Samples chosen to build surrogates for each grid level (minimum requirement = 6)
 - 4 responses selected
 - Total pressure recovery
 - Combustion efficiency
 - Total heat load
 - Powerhead heat load
- **Automated UQ framework utilized to extract uncertainties**
 - UQ setup utility executed to setup the simulations directory structure and input files
 - Performed the CFD simulations
 - UQ finalize utility executed to assess simulation convergence and extract P-box data

DAKOTA/VULCAN-CFD Demonstration



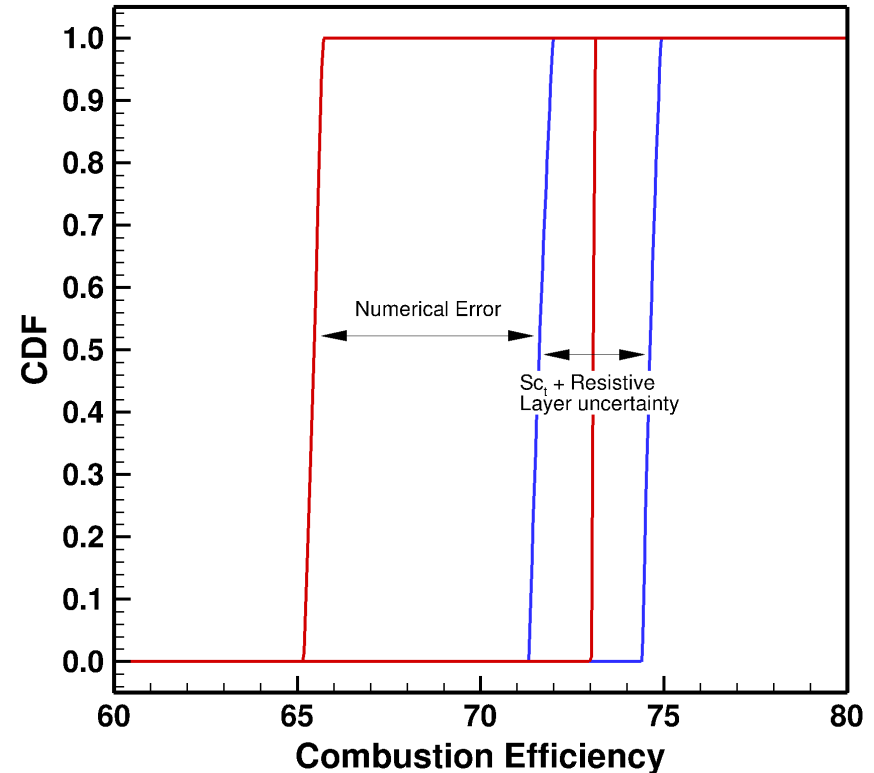
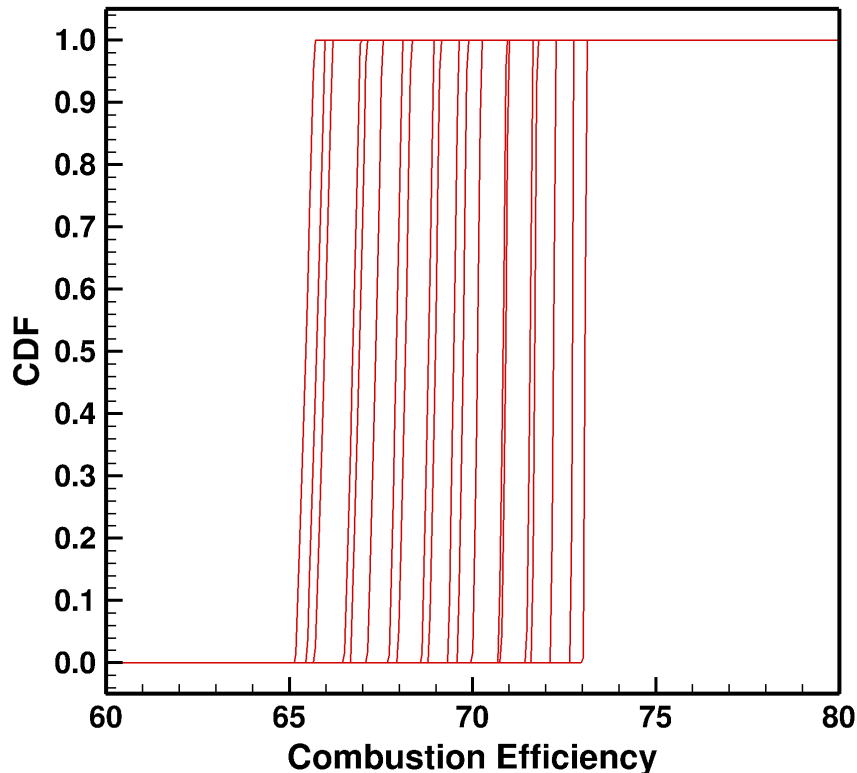
- **Image files directly output from UQ finalize utility to assess convergence**
 - Tecplot Macro files left in place for easy processing of other iterative convergence measures



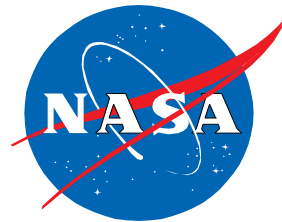
DAKOTA/VULCAN-CFD Demonstration



- P-box image depicting the uncertainty in combustion efficiency



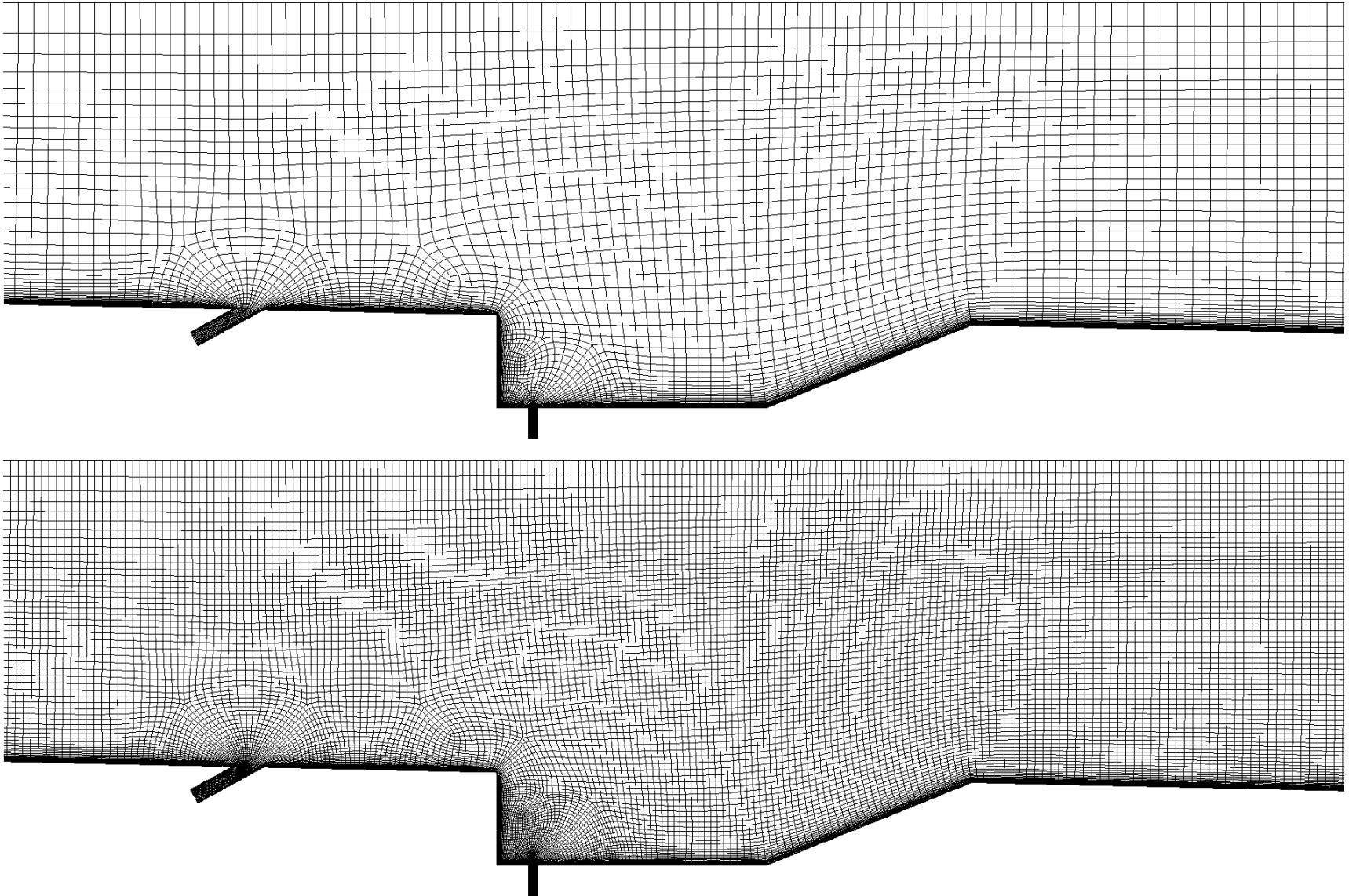
- **The SNL DAKOTA package has been coupled with VULCAN-CFD to provide a fully automated framework for UQ.**
- **The automation process removes the CFD applicationist from details required to integrate UQ into the CFD workflow.**
 - Minimizes user error with the simulation bookkeeping
 - Minimizes the UQ knowledge required by the CFD practitioner
- **The uncertainty framework used by this utility is in the form of a P-box that treats aleatoric uncertainties separately from epistemic uncertainties.**
- **The framework was demonstrated in a realistic environment using simulations of a scramjet combustor configuration.**
- **Although not emphasized, the automation strategy can in principle be applied to other CFD tools as well.**



Back-up Slides ...



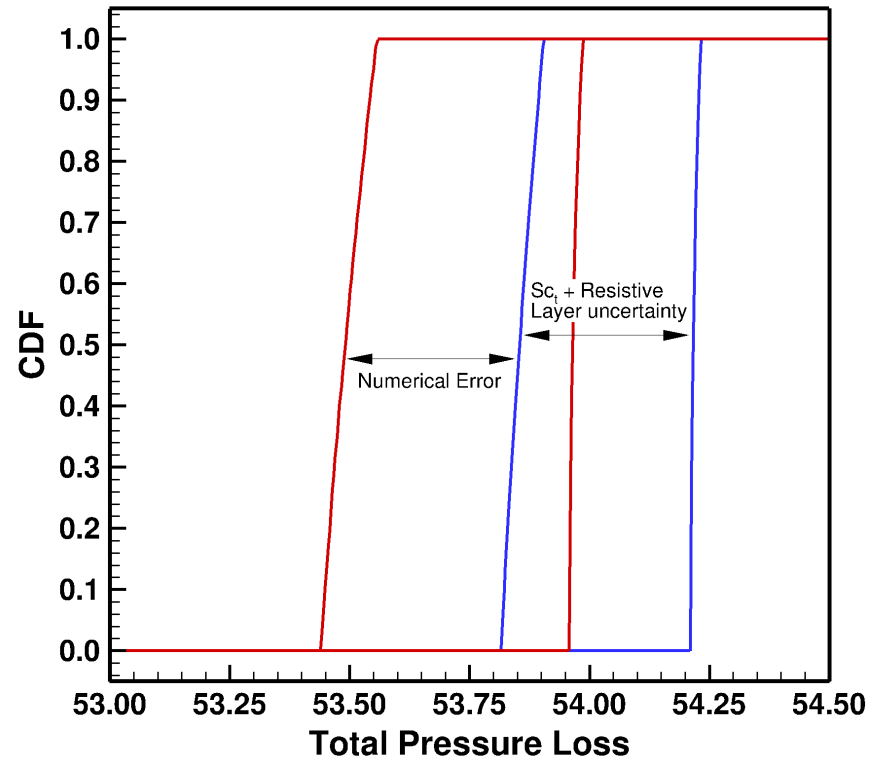
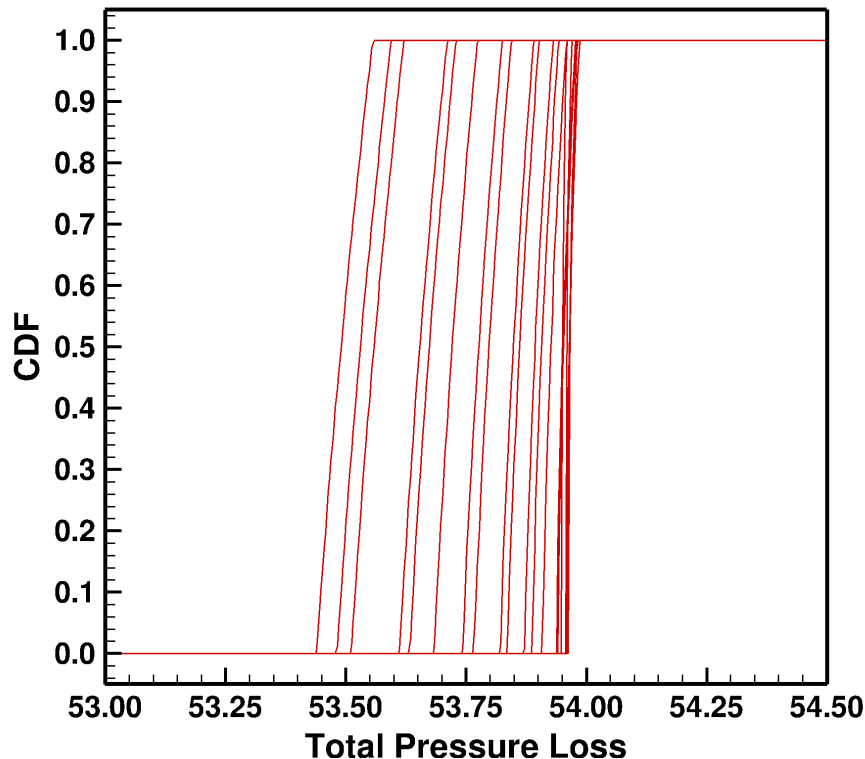
DAKOTA/VULCAN-CFD Demonstration



DAKOTA/VULCAN-CFD Demonstration



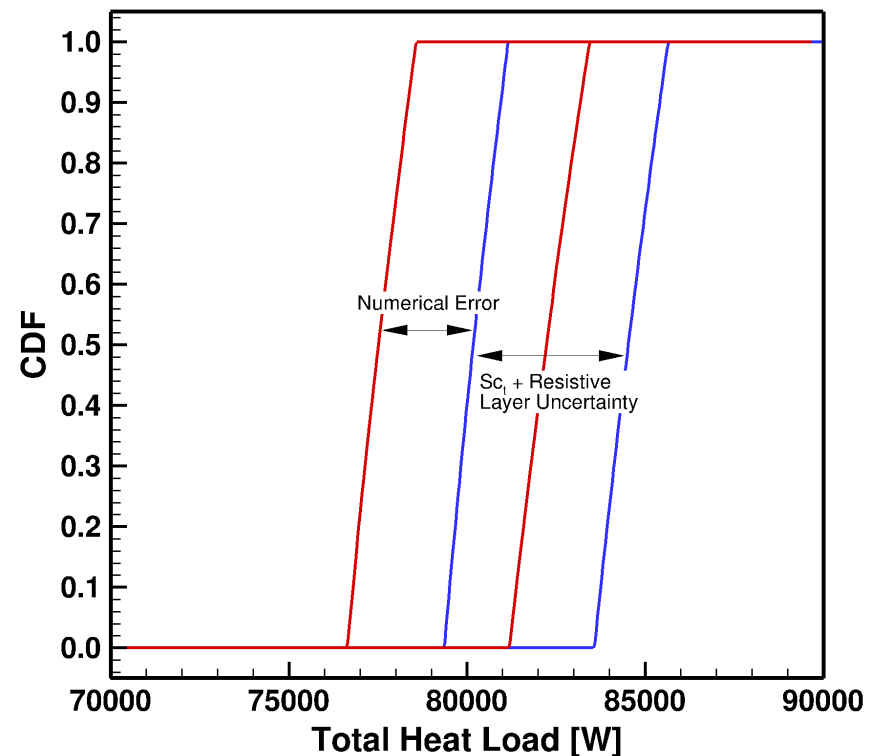
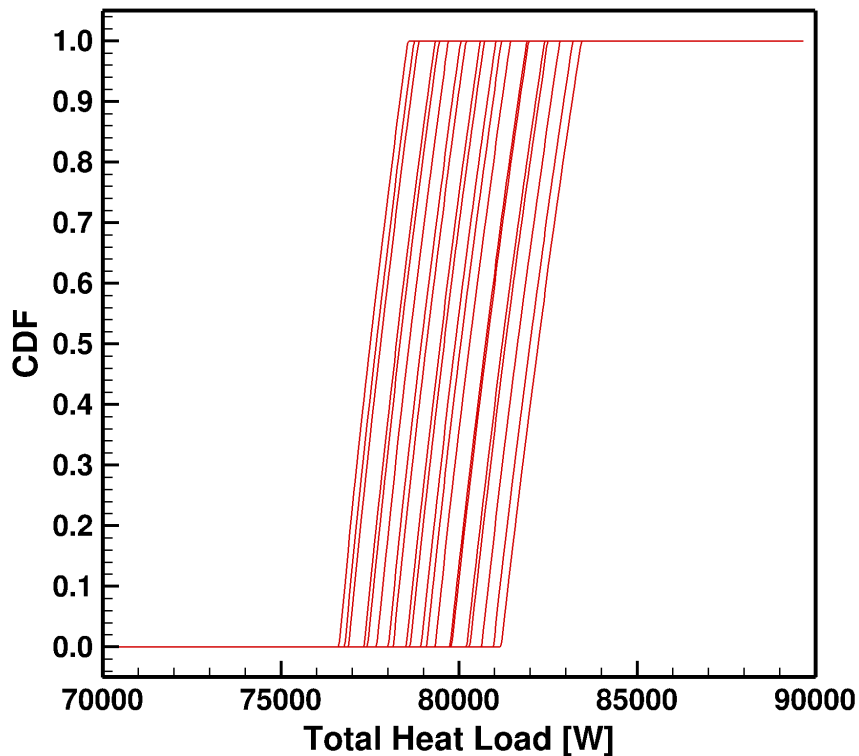
- P-box image depicting the uncertainty in total pressure recovery



DAKOTA/VULCAN-CFD Demonstration



- P-box image depicting the uncertainty in total heat load



DAKOTA/VULCAN-CFD Demonstration



- P-box image depicting the uncertainty in powerhead heat load

