SAND2015-4729PE

Exceptional service in the national interest

Sandia National Laboratories

An Overview of QMU at Sandia

Lisa Miller

Integrated Stockpile Evaluation Data Analyst (2957)

These slides were adapted from training material developed by Sandia's Statistics and Surveillance Assessment Department (org. 00415). For more information, contact:

Justin Newcomer, PhD
Principal Statistician
Department of Statistics and Surveillance Assessment
Sandia National Laboratories
Albuquerque, NM
505-844-5853
jtnewco@sandia.gov

U.S. DEPARTMENT OF ENERGY
NNSA National Nuclear Security Administration

Prepared by Sandia National Laboratories Albuquerque, New Mexico 87185 and Livermore, California 94550
Approved for public release; further dissemination unlimited.

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Key Events Change the Landscape of NW Certification

- June 6, 1989 – FBI Raid on Rocky Flats Plant
  - Nuclear operations shut down
- 1991 – W88 Completes Phase 5
- 1991-1992 – Cancelled phase 3 activities of several weapon programs
  - First time that there were no active NW development programs
- September, 23 1992 – Last underground nuclear weapon test
- 1993 – Beryllium and pit support functions transferred to LANL
- 1994 – National Defense Authorization Act (Public Law 103-160)
  - Multifaceted program to increase the understanding of the enduring stockpile
  - Predict, detect, and evaluate potential problems of the aging stockpile
  - Refurbish and re-manufacture weapons and components, as required
  - Maintain the science and engineering institutions
- 1994-1995 – Science Based Stockpile Stewardship is Initiated
- 1995 – JASON study voices concerns about NW assessment and certification
  - First explicit use of the term margin in this context
- 2000-2003 – LANL & LLNL begin to formalize their methodology for NW certification
  - B. T. Goodwin, R. J. Juzaitis (2003), "National Certification Methodology for the Nuclear Weapons Stockpile", UCRL-TR-223486 (published in 2006)

Quantification of Margins and Uncertainties (QMU) will be the framework used for future NW assessment and certification!

A Sandia QMU Timeline

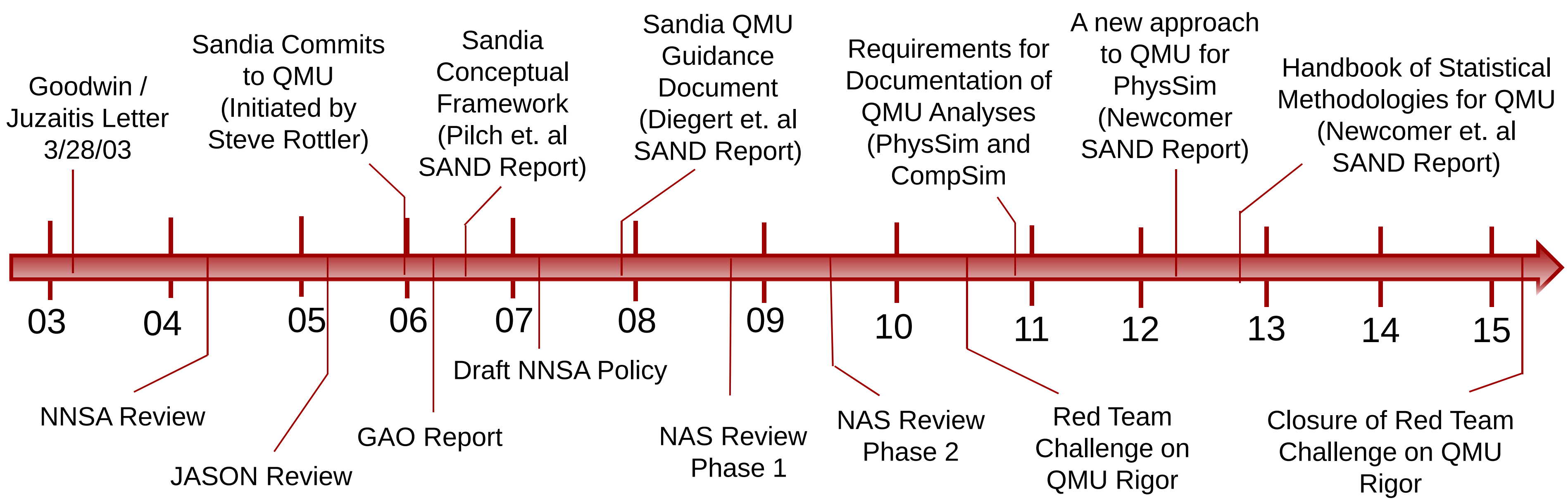
Sandia National Laboratories

- QMU is a new physics package certification methodology
- Sandia has its own certification methodology
- QMU is a physics lab only issue

- QMU applies to the whole lifecycle of the whole weapon
- Sandia is an equal partner in QMU
- Culture change, transition period
- Limited technical guidance

- Review what is working and what is not working
- Engineering analysis and reviews
- Increase formalism, transparency, and rigor of analyses and reporting
- Understanding of requirements

- Interface margins
- Environmental margins
- Larger role of CompSim
- Full system QMU



- Both QMU theory and implementation **are still being developed**
- Common definitions need to be developed
- Additional coordination to implement QMU is needed among the 3 NW labs

What is QMU and Why Do We Do It

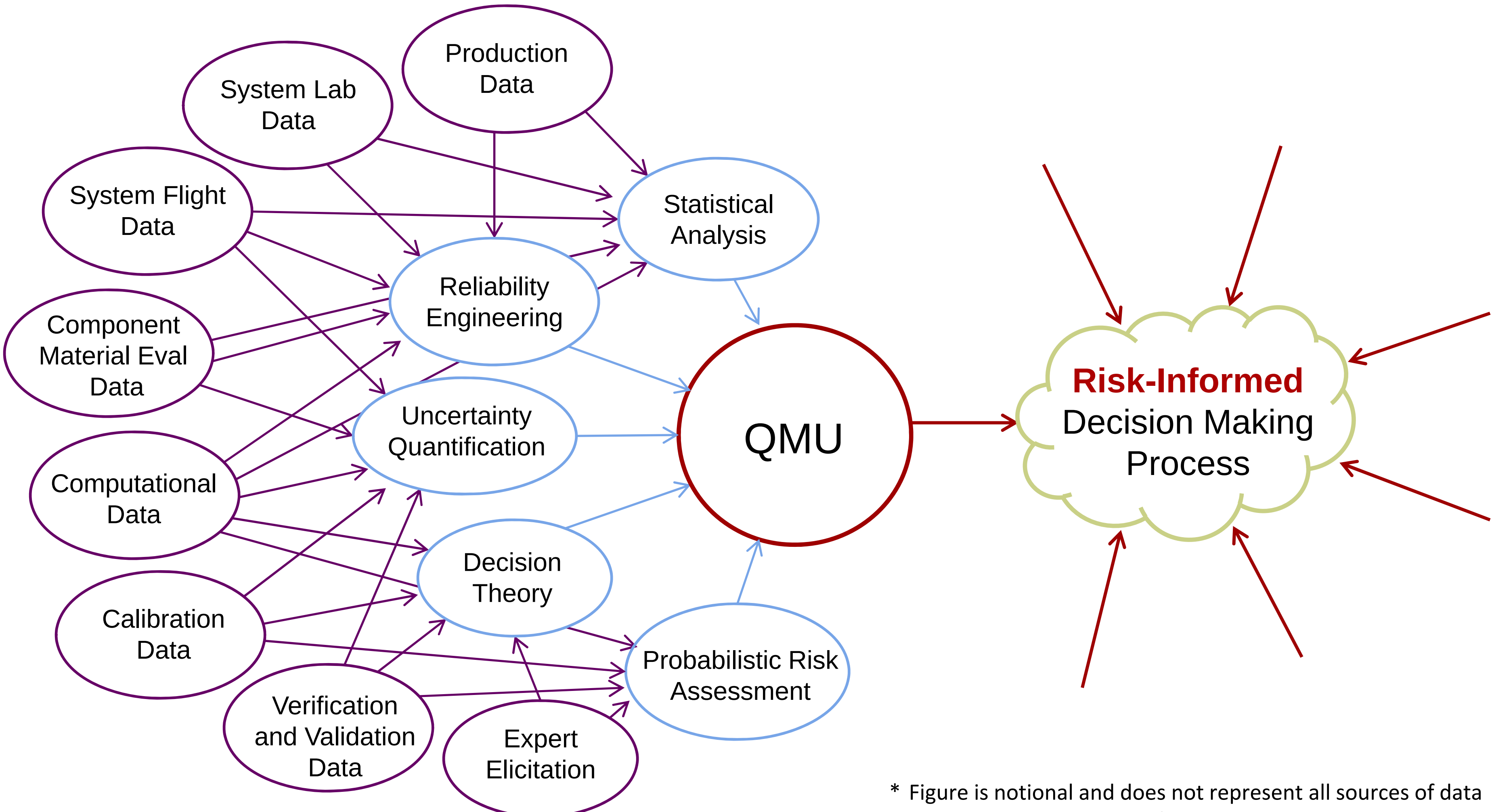
Sandia National Laboratories

- Quantification of Margins and Uncertainties (QMU) is a framework with applications that include the following:
  - **As a measure of sensitivity to change**: A characteristic with a large margin relative to its uncertainty can tolerate more change before jeopardizing performance compared to one with a small margin relative to its uncertainty.
  - **As a means of detecting a trend**: This can be used to predict when a performance characteristic will no longer have sufficient margin.
  - **As a means of determining a performance impact**: In some cases we can estimate the proportion of units that would fail to achieve their expected outcome (i.e. quantify the impact due to margin insufficiency).
    - Requires a credible estimate of a parameter's distribution through QMU **and** a credible understanding of the pass/fail limit relative to meeting performance requirements
  - **As a means of identifying sample size**: Performance characteristics with smaller Margin/Uncertainty (M/U) should drive a higher sampling rate than those with a larger M/U, since they are more sensitive to change.

QMU Supports a Risk-Informed Decision Making Process

The Complexity of QMU at Sandia

- QMU is a conceptual framework that ***is evolving*** over time that outlines a process for ***communicating*** the ***confidence*** in the stockpile.



* Figure is notional and does not represent all sources of data and the complete range of methodologies that could be used.

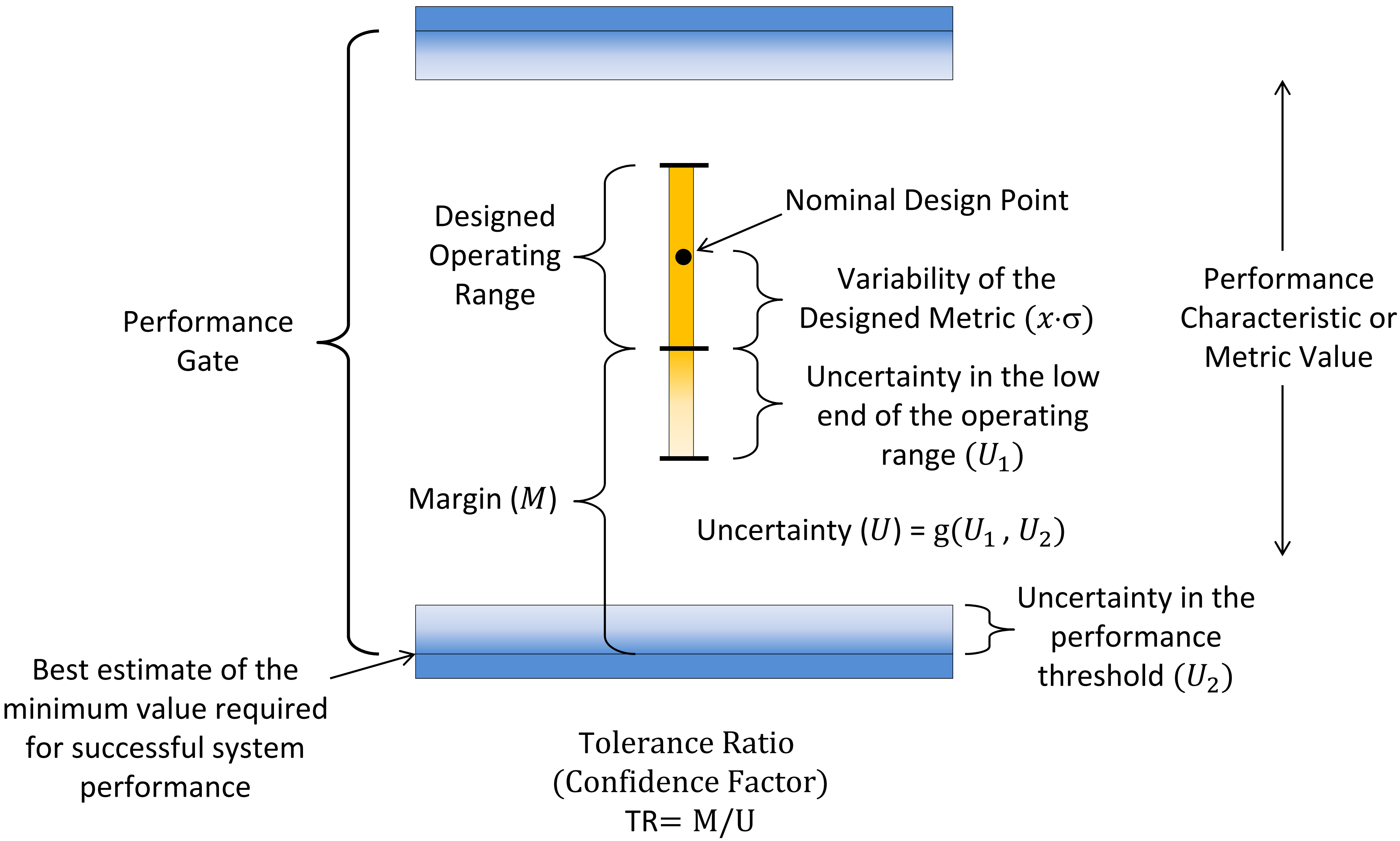
Key Elements of QMU

1. Specification of **Performance characteristics and their thresholds**
   - *Performance* is the ability of a bomb, a warhead, or a component to provide the proper function (e.g., timing, output, response to different environments) when exposed to the sequence of design environments and inputs.

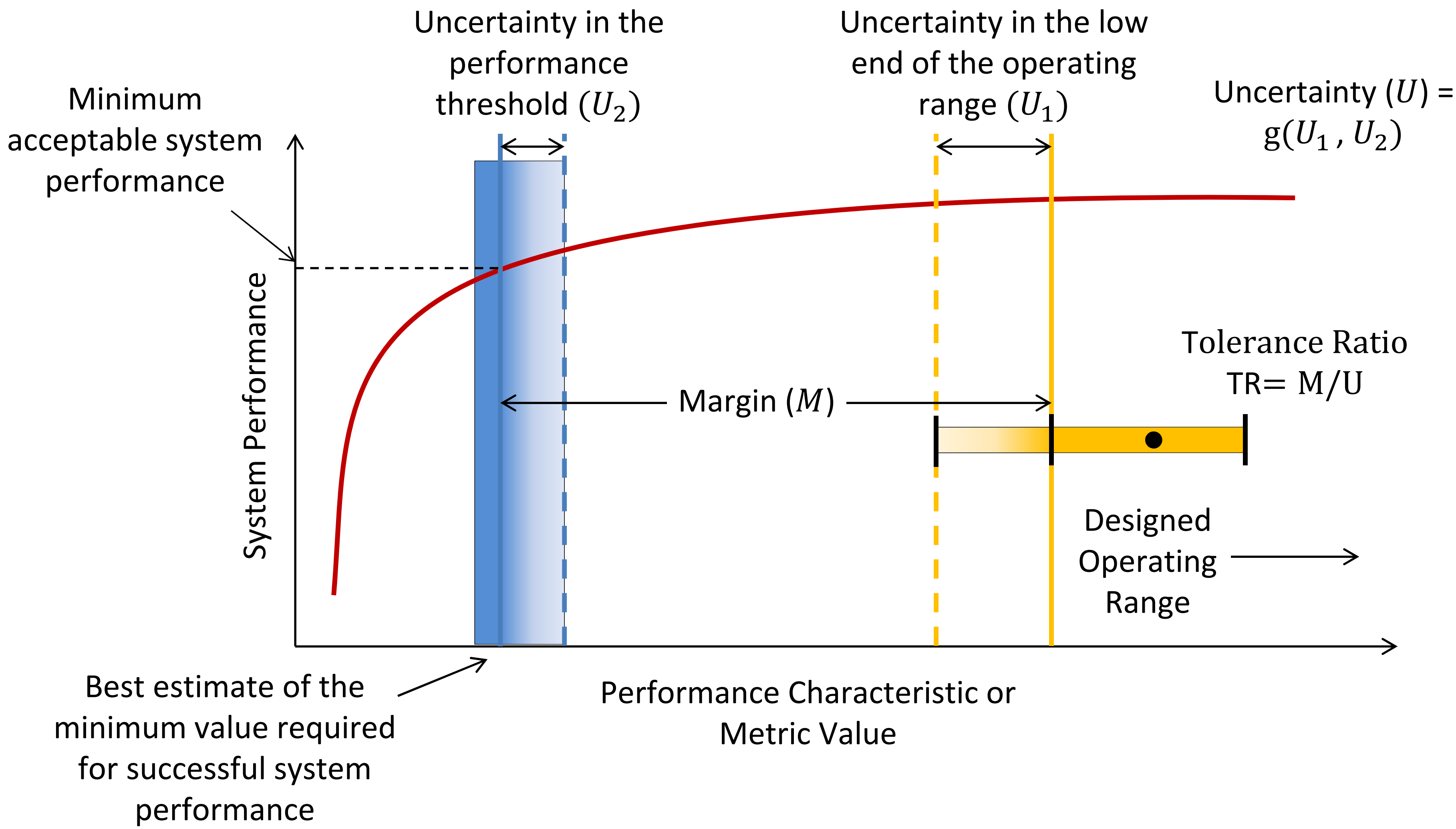
2. Identification and quantification of **Performance margins**
   - A *performance margin* is the difference between the required performance of a system and the demonstrated performance of a system, with a positive margin indicating that the expected performance exceeds the required performance.

3. Quantification of **uncertainty** in the performance thresholds and the performance margins as well as in the larger framework of the decisions being contemplated
   - Two types of uncertainty that we account for, quantify, and aggregate within QMU:
     1. ***Aleatory uncertainty* (variability)**
        – Variability in manufacturing processes, material composition, test conditions, and environmental factors, which lead to variability in component or system performance.
     2. ***Epistemic uncertainty* (lack of knowledge)**
        – Model form uncertainty, both known and unknown unknowns in scenarios, and limited or poor-quality physical test data.

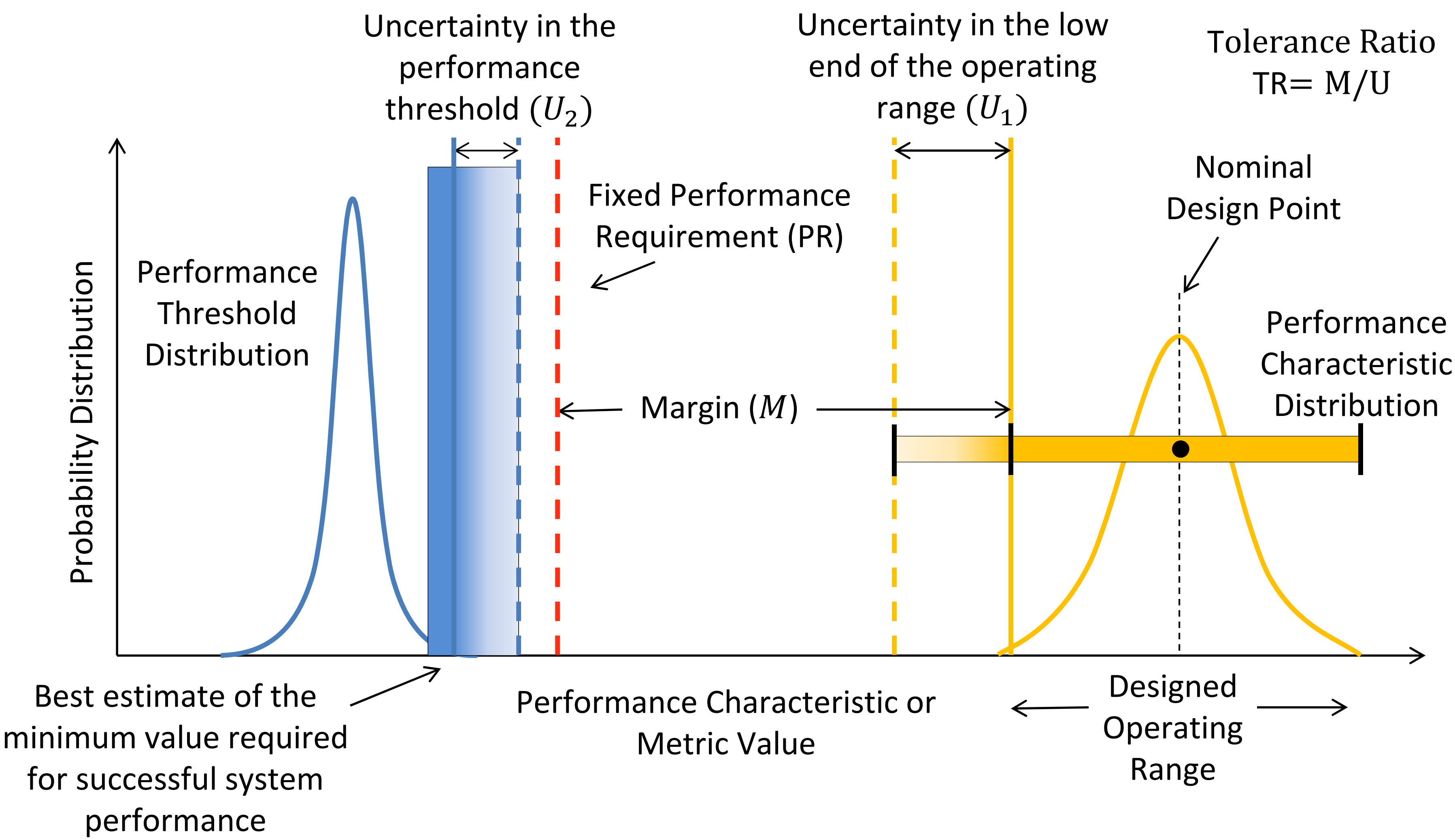
Conceptual Framework



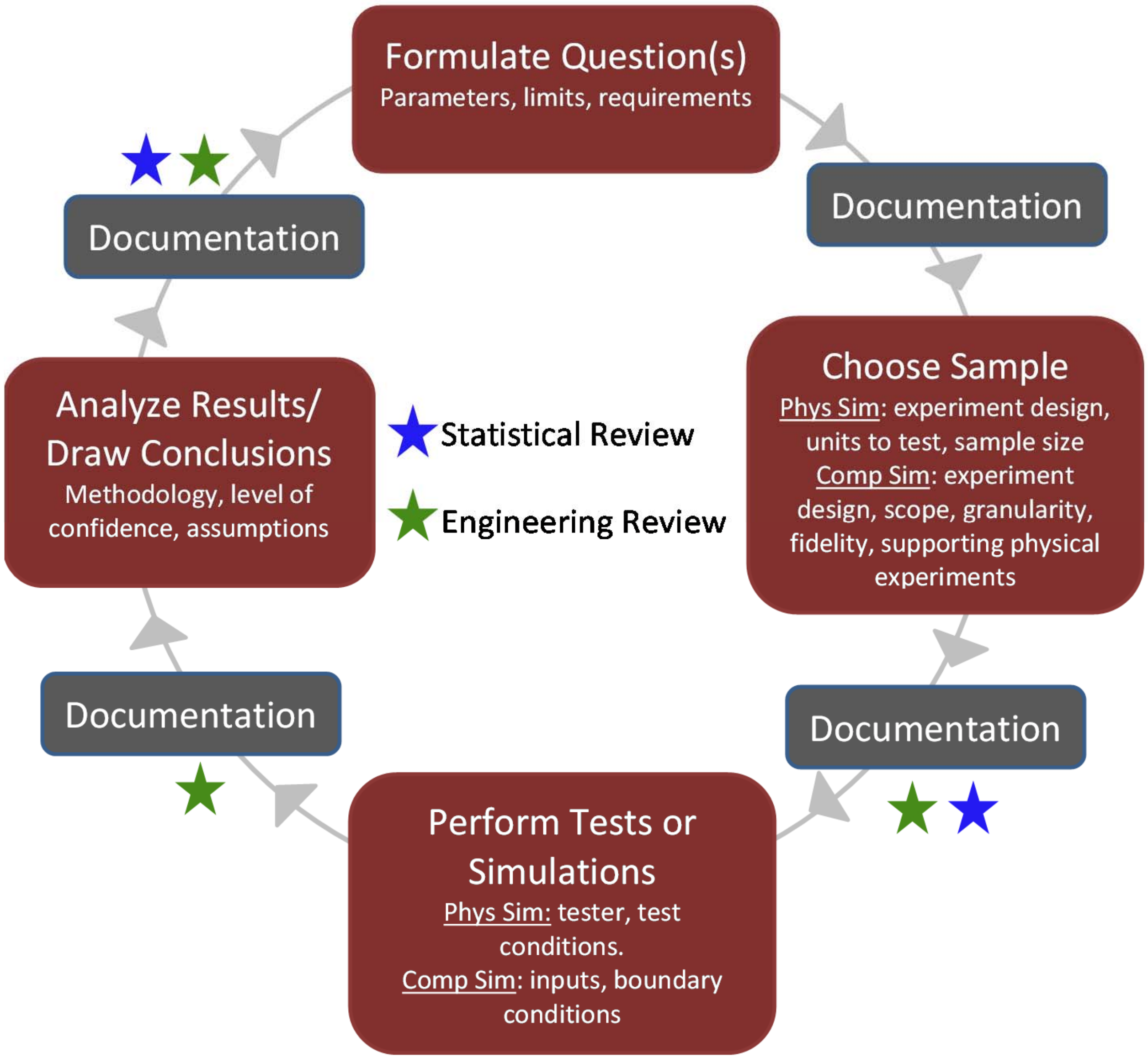
Conceptual Framework



Conceptual Framework



The QMU Process



Recommended Execution

- Begin with data reviews and screening analyses on a wide range of components and parameters
  - Quick looks at the data to better understand where more in-depth looks (Engineering Analyses) are needed. ***k*-factors can be used as a tool for screening.**
- Based on the previous step and other considerations documented by the Subject Matter Experts, prioritize analyses to look at in more depth (1-2 per year)
  - Prioritization could include a combination of physical and computational simulation QMU analyses
- Conduct engineering analyses with an interdisciplinary team and document the results
- Based on the engineering analysis documentation, determine whether a statistical analysis is merited
  - Engineering analyses may result in a conclusion that the data and/or current level of understanding of the data do not warrant a rigorous statistical analysis.
- If a rigorous statistical analysis is needed and the data is understood to an appropriate degree, perform rigorous statistical the analysis per the
- The completed QMU report should be independently reviewed and archived in the QMU workflow

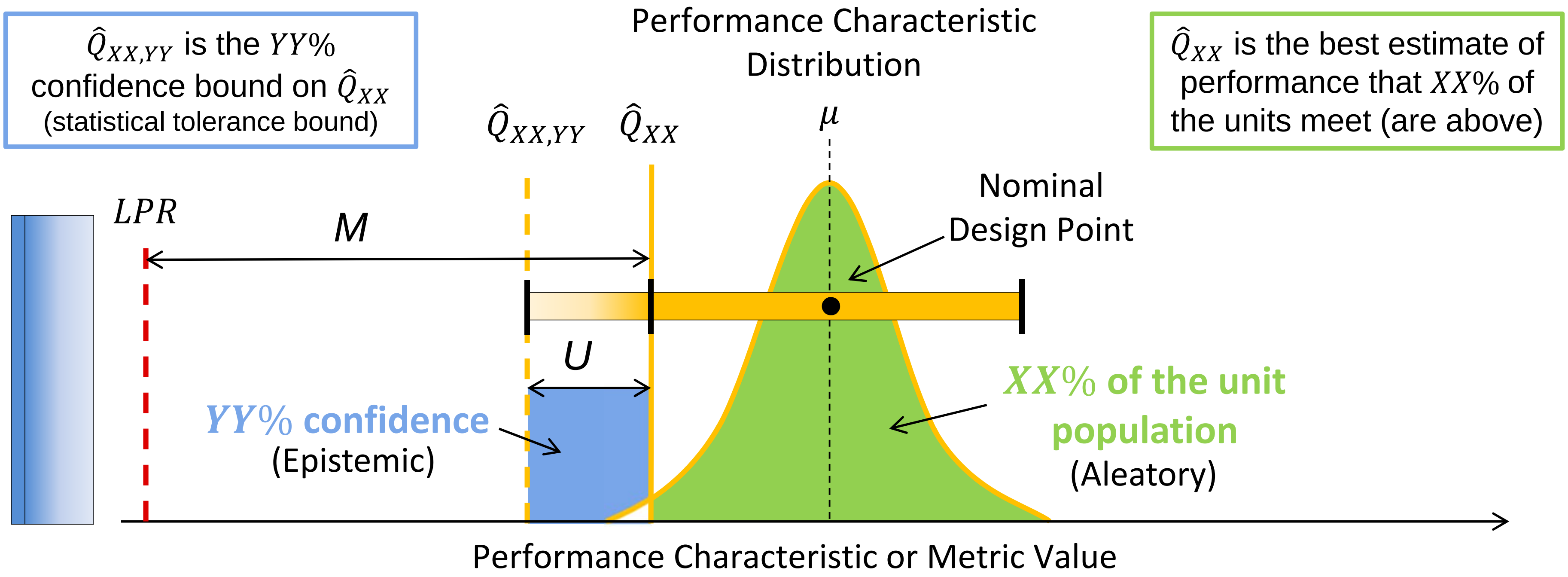
Understand the Problem and Data

Statistical Analysis

Document and Report Conclusions

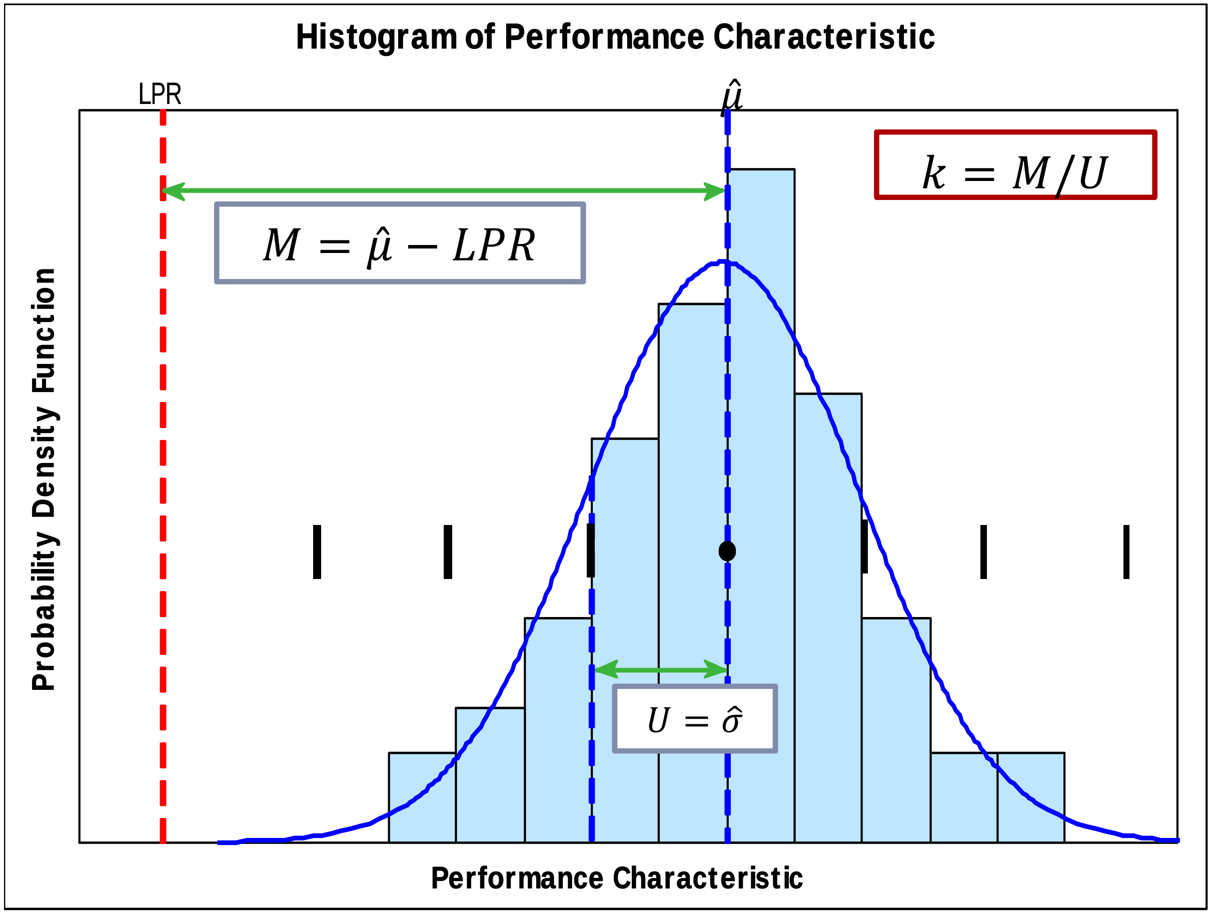
Statistical Framework for QMU Questions

- Are we $YY\%$ confident that **at-least** $XX\%$ of the unit population will yield a response **greater than** the lower performance requirement (LPR)?
  - The values of XX, YY, and LPR and the comparisons '**at-least**' versus '**at-most**' and '**greater than**' versus '**less than**' are all parameters of the requirement.
  - Provides a framework consistent with the Academy of Sciences recommendations*



*NAS/NRC (National Academy of Science/National Research Council) (2008), "Evaluation of quantification of margins and uncertainties for assessing and certifying the reliability of the nuclear stockpile", National Academy Press, Washington, DC

k-Factor Methodology



k = 1: We estimate (50% confidence) that 84% of the unit population will yield a response **greater than** the performance requirement ***LPR***

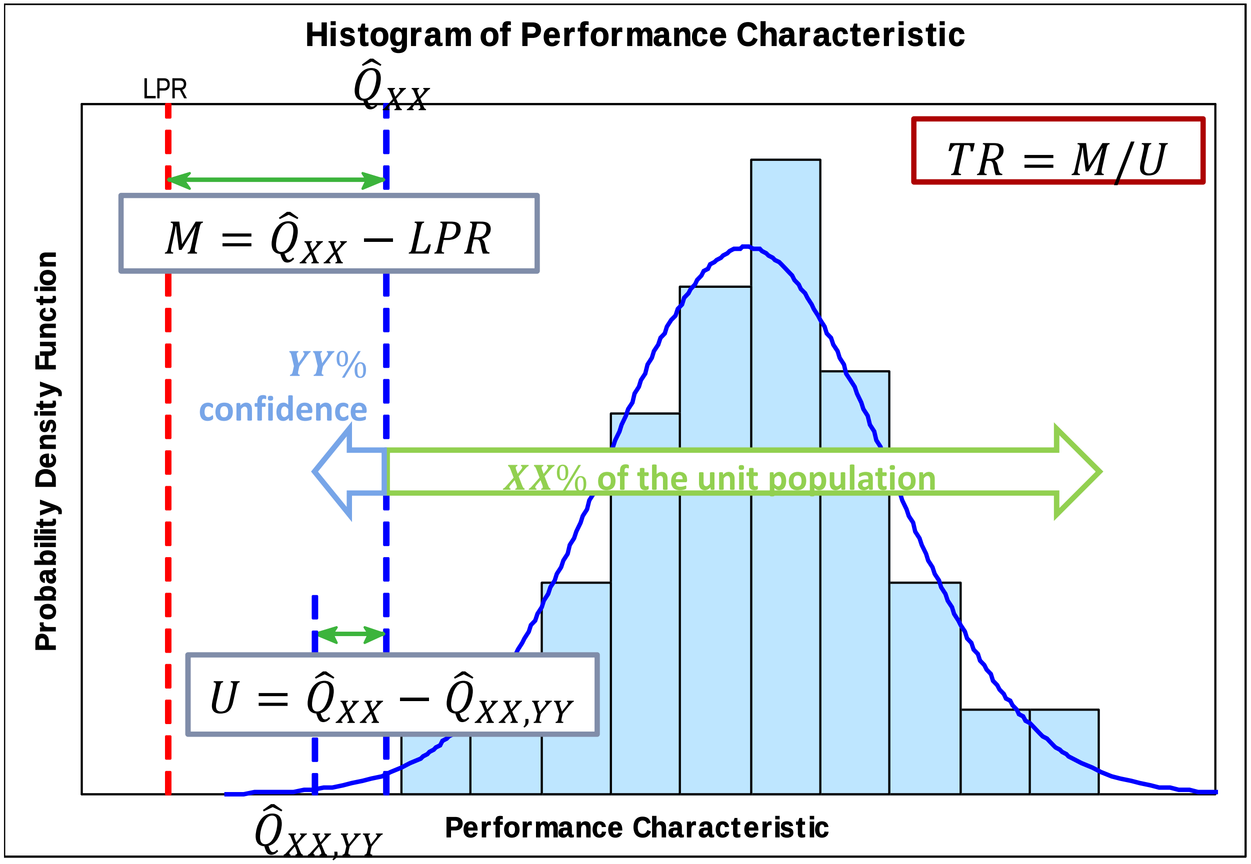
k = 2: We estimate (50% confidence) that 98% of the unit population will yield a response **greater than** the performance requirement ***LPR***

k = 3: We estimate (50% confidence) that 99.9% of the unit population will yield a response **greater than** the performance requirement ***LPR***

This is easy to estimate and is a great screening tool.
K-factor uses mean and standard deviation.

Graphical Depiction of the k-factor QMU Metrics

Tolerance Bound (TB) Methodology

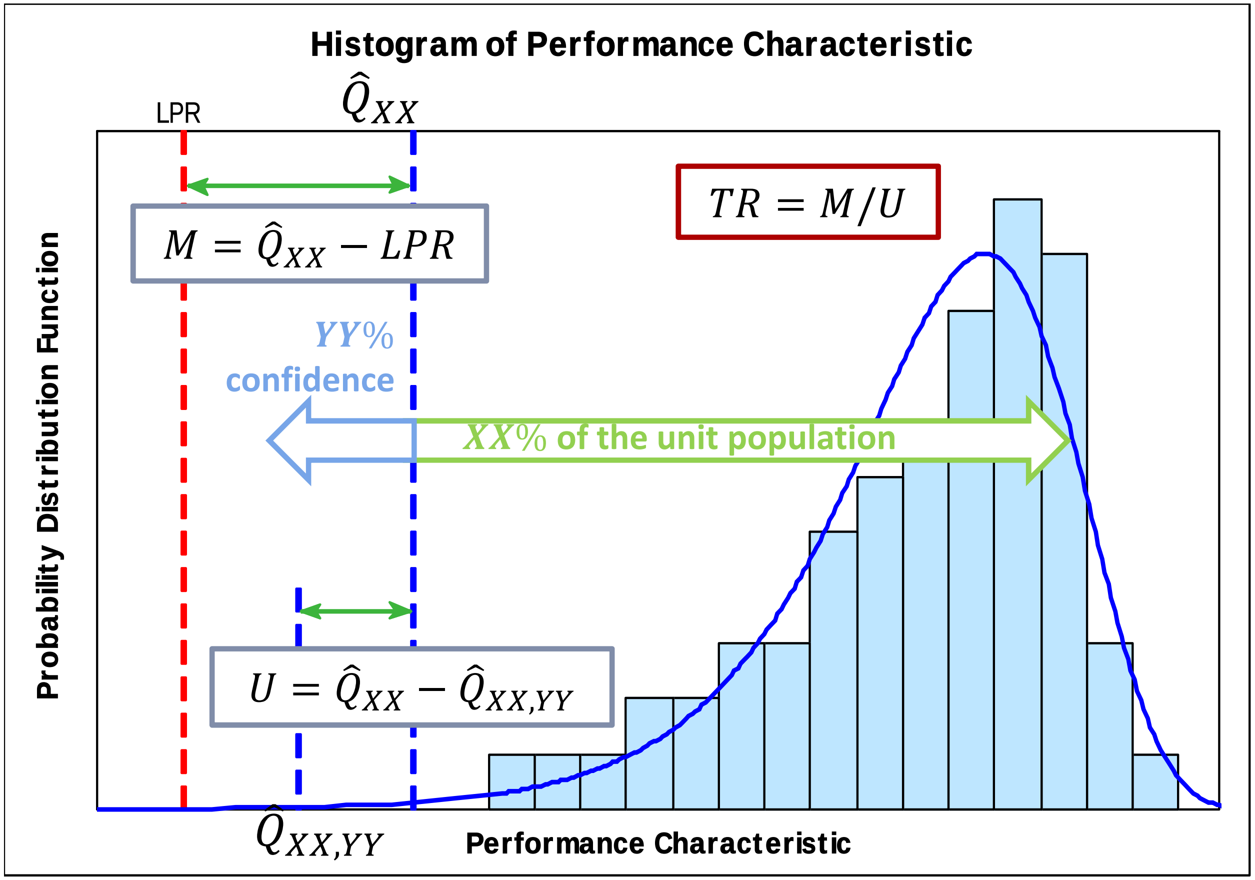


TR > 1: We are $YY\%$ confident that **at-least** $XX\%$ of the unit population will yield a response **greater than** the performance requirement $\boldsymbol{LPR}$.

Tolerance Ratio (TR) uses percentiles and tolerance bounds.

Graphical Depiction of the TB QMU Metrics Relative to a Normal Distribution

Tolerance Bound (TB) Methodology

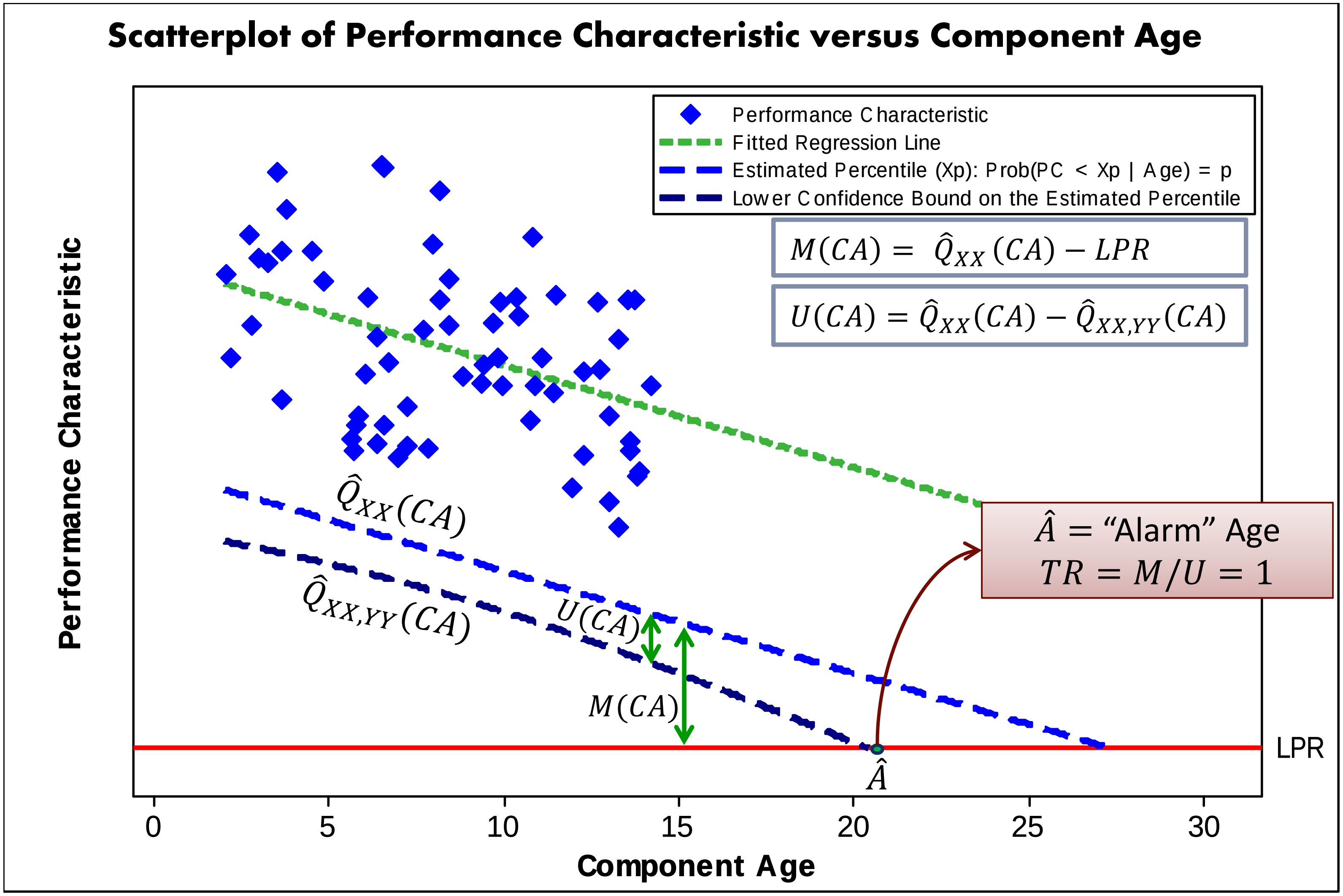


TR > 1: We are $YY\%$ confident that **at-least** $XX\%$ of the unit population will yield a response **greater than** the performance requirement LPR

Tolerance Ratio (TR) takes into account population distribution.

Graphical Depiction of the TB QMU Metrics Relative to a Non-Normal Distribution

Aging Trend with a Tolerance Bound



Graphical Depiction of the QMU Metrics with a Linear Regression Model

Physical vs. Computational Simulation

- Predicted responses obtained through computational simulation (compsim) differ from physical simulation response data in a number of ways. Three main differences are listed below.

  1. In most compsim applications, there are a relatively large number of controllable inputs representing both model parameters and boundary conditions.

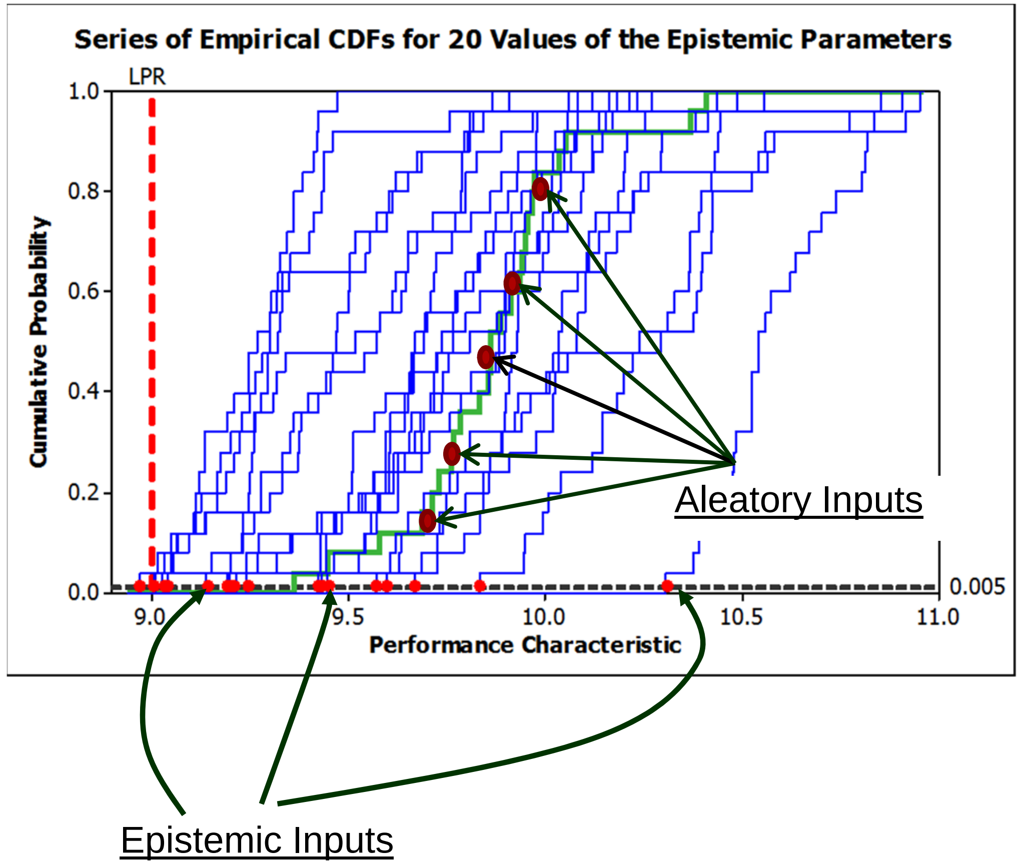
  2. In compsim applications, the response is deterministic (given the same inputs the code will yield the same responses).

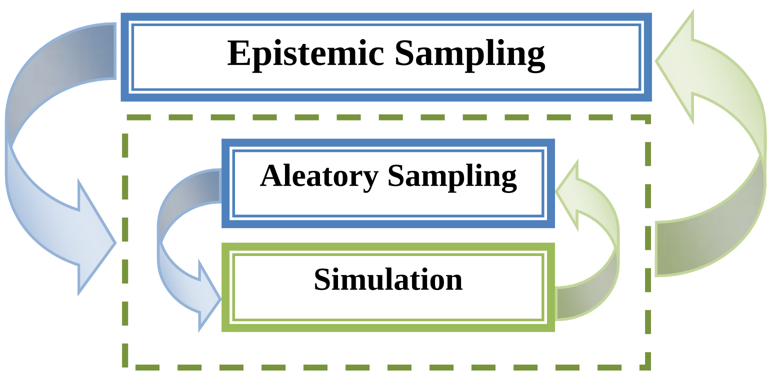
  3. Usually, there is a substantial amount of additional model-related uncertainty in compsim applications.
     - Mesh convergence uncertainty associated with incomplete mesh convergence
     - Model form uncertainty associated with approximate and incomplete models of the physics

- These differences drive a need for different technical methodologies.

QMU for Computational Simulation

- Probability of Frequency Methodology. Also referred to as "Second-Order" Probability
- Nested sampling combines epistemic and aleatory uncertainty
- For each outer loop sample a set of epistemic variables, run an inner loop Uncertainty Quantification study over aleatory (probability) variables

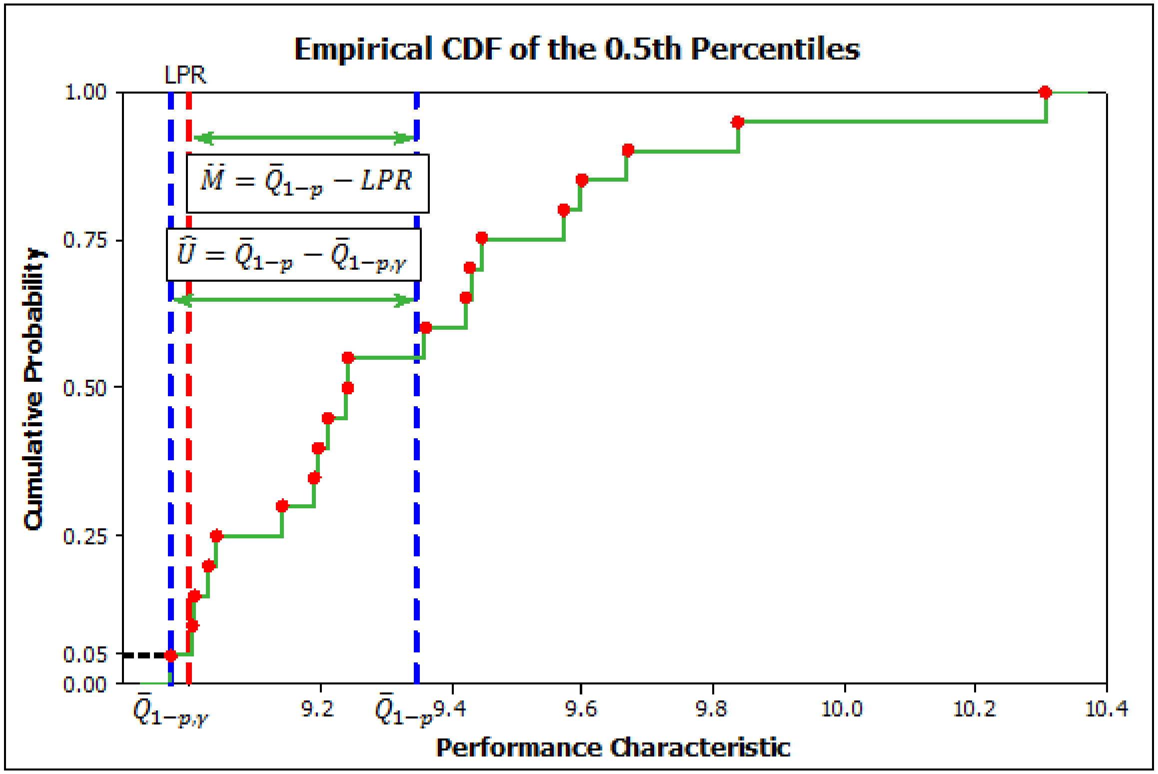




"Envelope" of CDF traces represents response epistemic uncertainty

QMU for Computational Simulation

- The average value from the estimated 0.5th percentile is used as the point estimate of the percentile: $\bar{Q}_{1-p}$
- To account for the uncertainty in the point estimate, a confidence bound is obtained from the distribution: $\bar{Q}_{1-p,\gamma}$
- In this example the mean of the 0.5th percentile distribution is 9.35 and the 95% confidence level is 8.97.



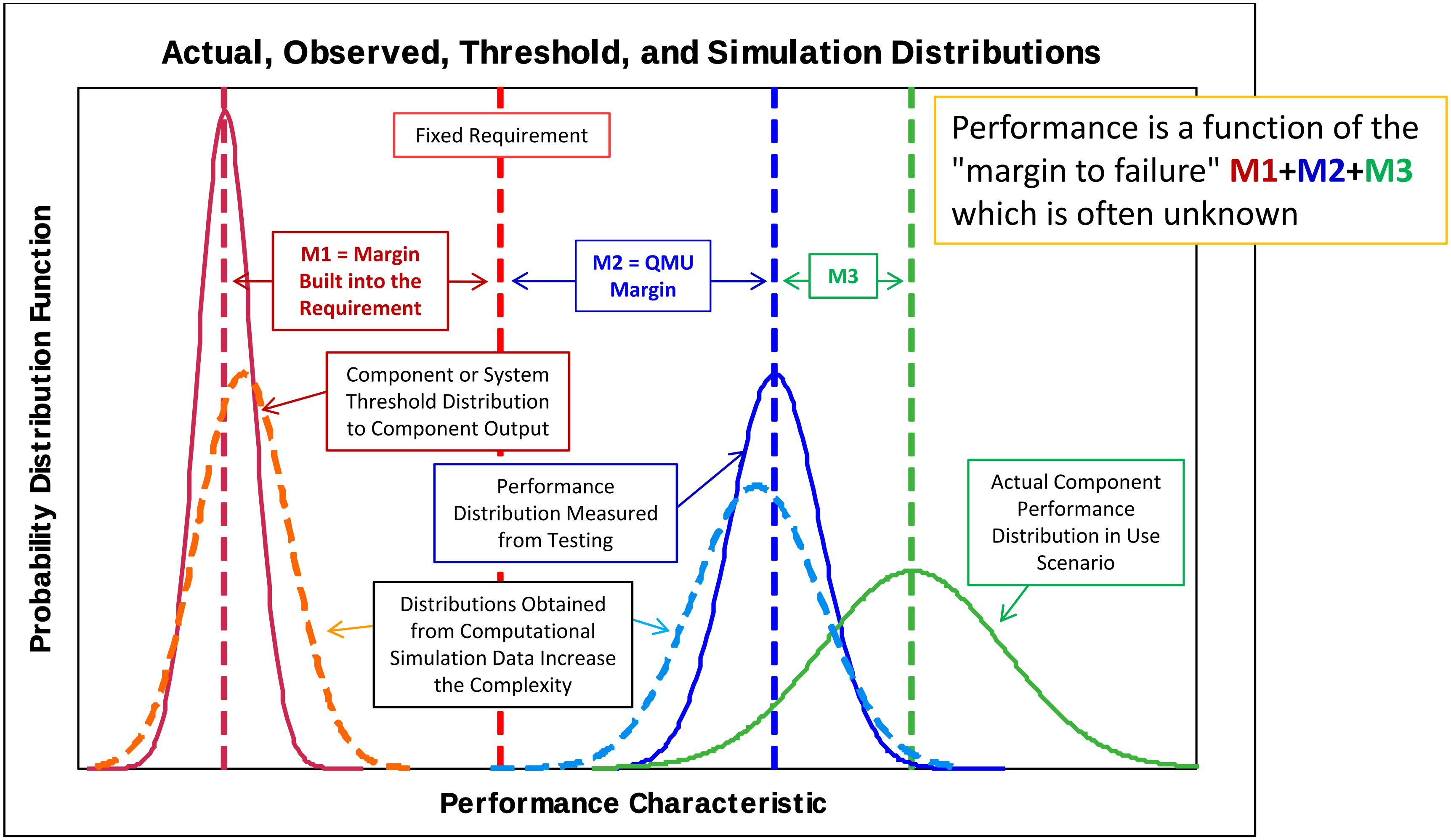
$$\hat{M} = \bar{Q}_{1-p} - LPR = 9.35 - 9 = 0.35$$

$$\hat{U} = \bar{Q}_{1-p} - \bar{Q}_{1-p,\gamma} = 9.35 - 8.97 = 0.38$$

$$TR = \hat{M}/\hat{U} = 0.92$$

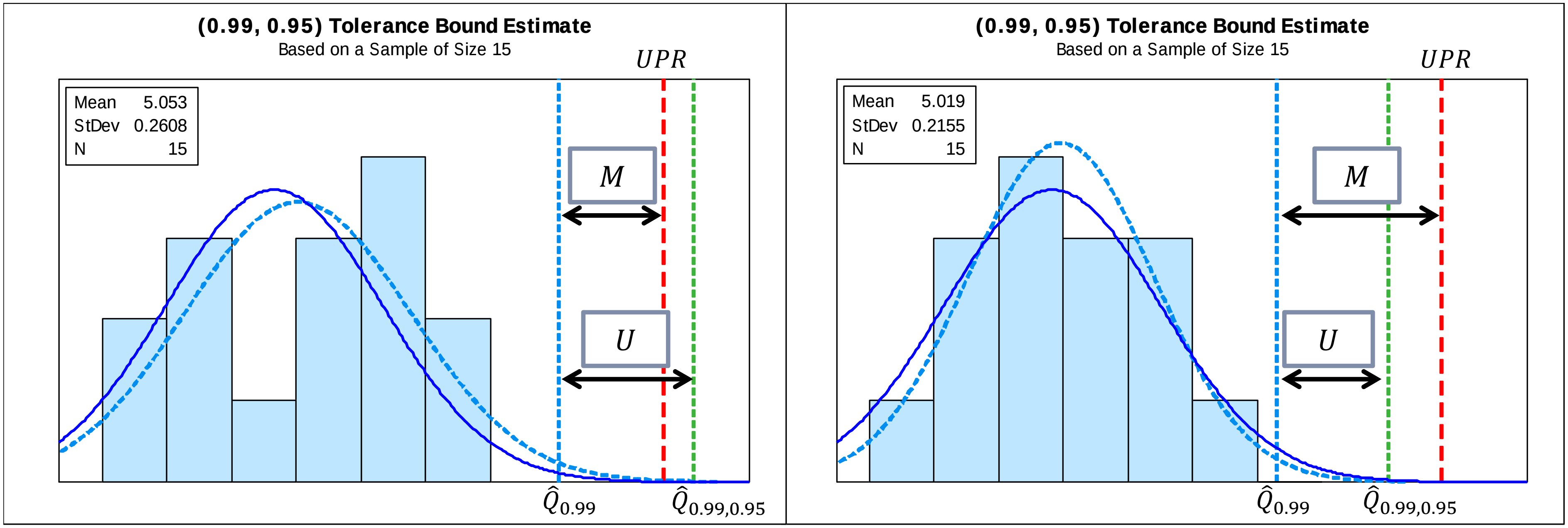
Graphical Depiction of the Probability of Frequency QMU Metrics

- Is our testing representative to use conditions and is our performance requirement a realistic measure of component function impacting system performance?



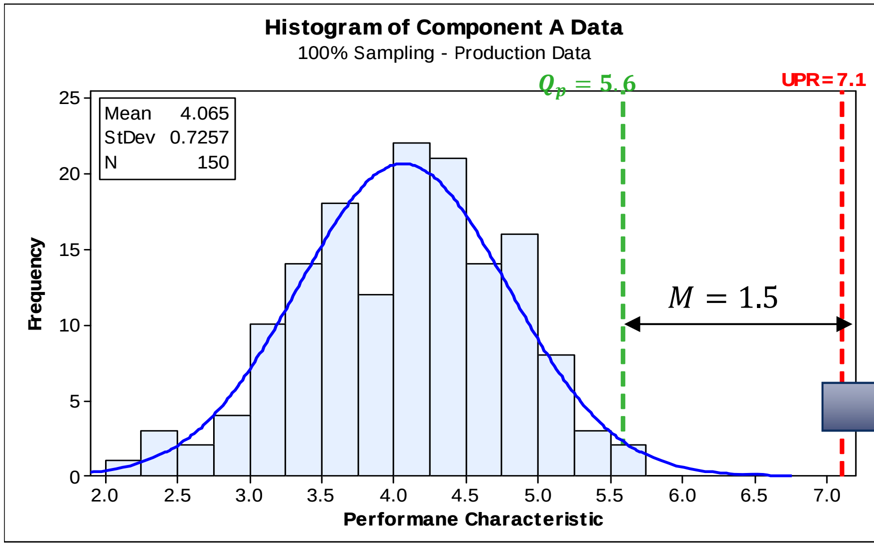
QMU Used for Margin Based Sample Size

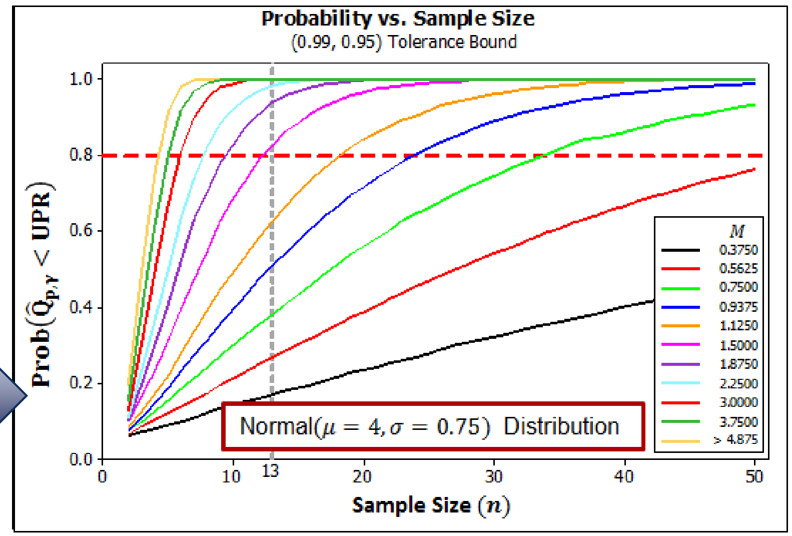
- "How many samples do I need to **demonstrate** the performance characteristic has sufficient margin to requirements <u>with a high probability</u>, provided the true underlying distribution has positive margin?"
  - Samples below are from the same population, same sample size, but produce different results.
  - Smaller sample sizes create more risk of not being able to make high confidence statements.



Margin Based Sample Sizes - Example

- Suppose we are interested in re-characterizing a component at some point in its lifetime and there was 100% sampling conducted during production (or previous testing). The initial population size consisted of 150 units, all of which were tested.





Curves are generated via Monte Carlo Simulation

$p = 0.99, \gamma = 0.95$			
M	Minimum sample size for $Prob(Q_{p,\gamma} < UPR) \geq 0.80$	M	Minimum sample size for $Prob(Q_{p,\gamma} < UPR) \geq 0.80$
0.375	110	1.875	10
0.5625	55	2.25	8
0.75	34	2.625	7
0.9375	24	3.0	6
1.125	19	3.75	5
1.5	13	≥4.875	4

A sample of size $n = 13$ is required to have an 80% probability that the estimated $(p = 0.99, \gamma = 0.95)$ upper tolerance bound will be less than the upper performance requirement provided the observed margin is at least $M = 1.5$

The Future Vision for QMU at Sandia

- More rigorous engineering analyses and data reviews to better understand what is being tested, what it means, and what the limitations of the available data are.
  - Are there gaps in the current surveillance testing?
  - Are we measuring what's important?
- QMU analyses at the subsystem level that look at interactions between components and performance characteristics.
  - How do we "roll up" analyses on several characteristics?
  - Does the current testing exercise potential failure mechanisms at the interfaces between components?
- Failure testing and analyses to better understand the thresholds for critical performance characteristics.
  - If an analysis indicates low margin relative to a requirement, what does that mean?
  - Do we feel the system will fail if a unit performs outside of its requirements?
  - Do we understand the point at which it will begin to fail?
- More consistent use of methodologies and integration of both physical simulation and computational simulation QMU analyses.
  - Our ability to communicate the results of QMU analyses should be independent of the types of methodologies used.

References for QMU

1. J.T. Newcomer and K.E. Freeland (2013), *A Margin Based Approach to Determining Sample Sizes via Tolerance Bounds,* SAND2013-8168, September 2013.
2. J.T. Newcomer, B.M. Rutherford, E.V. Thomas, R.L. Bierbaum, L.M. Hickman, J.W. Lane, S Fitchett, A Urbina, A.A. Robertson, and L.P. Swiler, *Handbook of Statistical Methodologies for QMU*, SAND2012-9603, November 2012.
3. J.T. Newcomer, *A New Approach to QMU for Physical Simulation Data*, SAND2012-7912, Sandia National Laboratories, Albuquerque, NM, September 2012.
4. *Special Issue on Quantification of Margins and Uncertainties*, Reliability Engineering and System Safety, Volume 96, Issue 9, Pages 959-1256, Edited by J.C. Helton and M. Pilch, September 2011.
5. K. Krishnamoorthy and T. Mathew, *Statistical Tolerance Regions, Theory, Applications, and Computation*, Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., 2009.
6. K.V. Diegert, S.E. Klenke, G. Novotny, R.A. Paulsen, M. Pilch, and T.G. Trucano, *Toward a More Rigorous Application of Margins and Uncertainties within the Nuclear Weapons Lifecycle—A Sandia Perspective*, SAND2007-6219, Sandia National Laboratories, Albuquerque, NM, December 2007.
7. M. Pilch, T.G. Trucano, and J.C. Helton, *Ideas Underlying Quantitative Margins and Uncertainty (QMU): A White Paper*, SAND2006-5001, Sandia National Laboratories, Albuquerque, NM, September 2006.
8. Goodwin BT, Juzaitis RJ. *National certification methodology for the nuclear weapons stockpile*, UCRL-TR-223486. Lawrence Livermore National Laboratory, Livermore, CA, 2006.
9. G.E. Box, W.G. Hunter, J.S. Hunter, *Statistics for Experimenters: Design, Innovation, and Discovery*, 2nd Edition, Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., 2005.
10. W.Q. Meeker and L. A. Escobar, *Statistical Methods for Reliability Data*, Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., 1998.
11. R.H. Myers, *Classical and Modern Regression with Applications*, Second Edition, Duxbury Press, 1990.
12. J.K. Patel, C.H, Kapadia, and D.B. Owen, *Handbook of Statistical Distributions*, Marcel Dekker, Inc., 1976.
13. D.B. Owen, *Factors for One-Sided Tolerance Limits and for Variables Sampling* Plan, Sandia Corporation Monograph SCR-607, Sandia National Laboratories, Albuquerque, NM, March 1963.