Evaluating and Quantifying Uncertainty

Antonio Possolo
NIST Fellow & Chief Statistician

Statistical Engineering Division
Information Technology Laboratory

January 14, 2016

NIST
National Institute of Standards and Technology

References

- M. J. Bayarri *et al.* (2007)
  A Framework for Validation of Computer Models
  Technometrics 49(2), 138–154
  `http://dx.doi.org/10.1198/004017007000000092`

- J. Palomo *et al.* (2015) SAVE: An R Package for the Statistical Analysis of Computer Models
  Journal of Statistical Software, 64(13), 1–23
  `http://dx.doi.org/10.18637/jss.v064.i13`

- A. Possolo (2015) *Simple Guide for Evaluating and Expressing the Uncertainty of NIST Measurement Results*
  NIST Technical Note 1900
  `http://dx.doi.org/10.6028/NIST.TN.1900`

Measurement Uncertainty — Pitot Tube
MEASUREMENT EQUATION

- Airspeed $v = \sqrt{\dfrac{2\Delta R_s T}{p}}$

Δ	Difference between total and static pressures
T	Air temperature
R_s	Specific gas constant for dry air
p	Static air pressure

Measurement Uncertainty — Pitot Tube
INPUTS & UNCERTAINTY EVALUATION

INPUTS

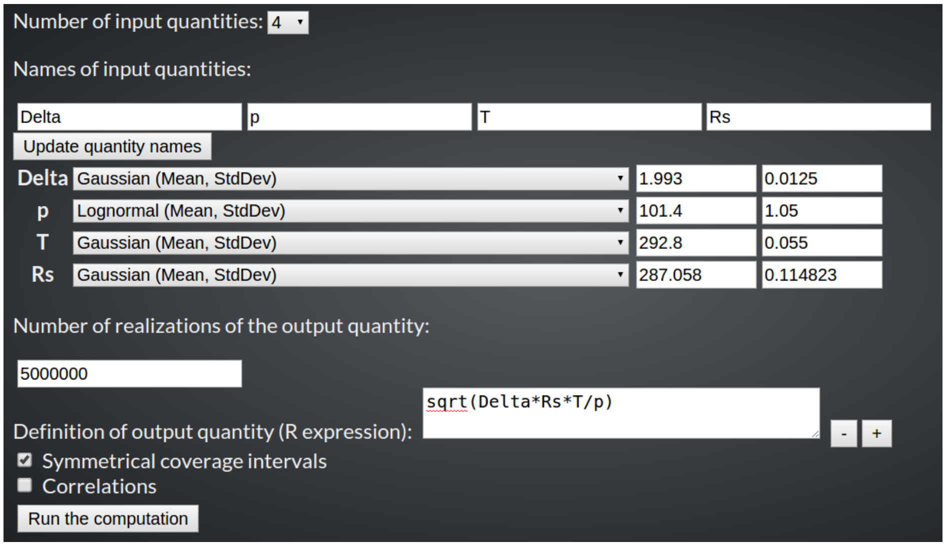
	ESTIMATE	STD. UNC.	MODEL
Δ	1.993 kPa	0.0125 kPa	Gaussian
p	101.4 kPa	1.05 kPa	Lognormal
T	292.8 K	0.055 K	Gaussian
R_s	287.058 $\mathrm{J\,kg^{-1}\,K^{-1}}$	0.114 823 $\mathrm{J\,kg^{-1}\,K^{-1}}$	Gaussian

EVALUATION

- NIST Uncertainty Machine (`uncertainty.nist.gov`)
- Gauss's Formula (GUM) and Monte Carlo Method (GUM-S1) produce same results: $v = 40.6\,\mathrm{m/s}$ and $u(v) = 0.25\,\mathrm{m/s}$

Measurement Uncertainty — Pitot Tube

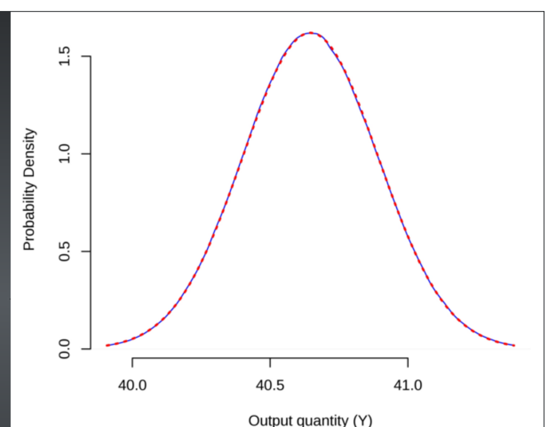
NIST UNCERTAINTY MACHINE — INPUT



Measurement Uncertainty — Pitot Tube

NIST UNCERTAINTY MACHINE — OUTPUT

```
===== RESULTS ==============================

Monte Carlo Method

Summary statistics for sample of size 5000000

ave     = 40.6462
sd      = 0.2462
median  = 40.6456
mad     = 0.246

Symmetrical coverage intervals

99% (      40,     41.3)        k =      2.6
95% (    40.2,     41.1)        k =      1.8
90% (    40.2,     41.1)        k =      1.8
68% (    40.4,     40.9)        k =        1

ANOVA (% Contributions)

            w/out Residual w/ Residual
Delta                26.73       26.73
p                    73.14       73.13
T                     0.02        0.02
Rs                    0.11        0.11
Residual                NA        0.01

-------------------------------------------

Gauss's Formula (GUM's Linear Approximation)

      y  = 40.6448
    u(y) = 0.2462
```

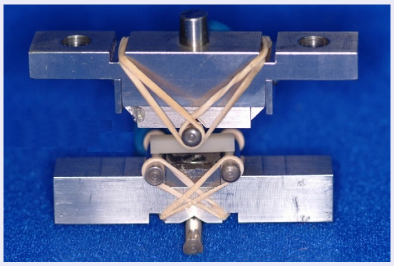


Download binary R data file with Monte Carlo values of output quantity
Download a text file with Monte Carlo values of output quantity
Download text file with numerical results shown on this page
Download JPEG file with plot shown on this page
Download configuration file

Measurement Uncertainty — Tensile Strength
OBSERVATION EQUATION

- *Observation equation* expresses measurand as known function of parameters of probability distribution of inputs

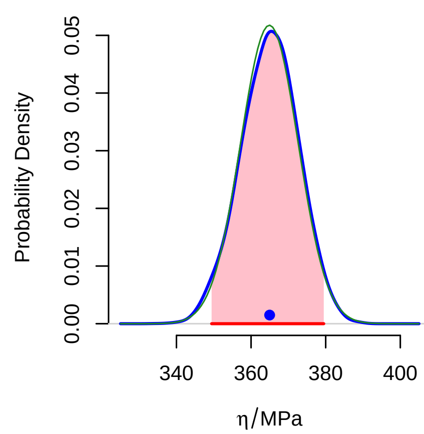
ALUMINA COUPONS



- Rupture stress in flexure test modeled as Weibull random variable with shape α and scale σ_C
- MEASURAND Weibull mean value $\eta = \sigma_C \Gamma(1 + 1/\alpha)$

Measurement Uncertainty — Tensile Strength
RESULTS

- Maximum likelihood estimates $\widehat{\alpha} = 10.1$, $\widehat{\sigma}_C = 383\,\text{MPa}$
- $\widehat{\eta} = \widehat{\sigma}_C \Gamma(1 + 1/\widehat{\alpha}) = 365\,\text{MPa}$
- Uncertainty evaluation
  - Monte Carlo (exact)
  - Statistical Theory (approximate)



Measurement

Experimental or computational process
that, by comparison with a standard,
produces an estimate of the true value of a property
of a material or virtual object or collection of objects,
or of a process, event, or series of events,
together with an evaluation of the uncertainty associated with that estimate,
and intended for use in support of decision-making

— NIST TN 1900, §2

Measurement Uncertainty

HEAT CAPACITY OF AMMONIA

We think our reported value is good to 1 *part in* 10 000

We are willing to bet our own money at even odds that it is correct to 2 *parts in* 10 000

Furthermore, if by any chance our value is shown to be in error by more than 1 *part in* 1000*, we are prepared to eat the apparatus and drink the ammonia*

— C. H. Meyers, 1930s

Told by D. P. Johnson, reported by H. Ku, 1973
Quoted by T. Doiron & J. Stoup, 1997

Measurement Uncertainty

> Doubt about the true value of the measurand that remains after making a measurement
>
> — NIST TN 1900, §3

- Measurement uncertainty described fully and quantitatively by probability distribution on set of values of measurand
- Probability distribution represents state of knowledge:
  - Subjective construct that expresses how firmly metrologist believes she knows measurand's true value
  - Characterizes how degree of her belief varies over set of possible values of measurand

Uncertainty Quantification

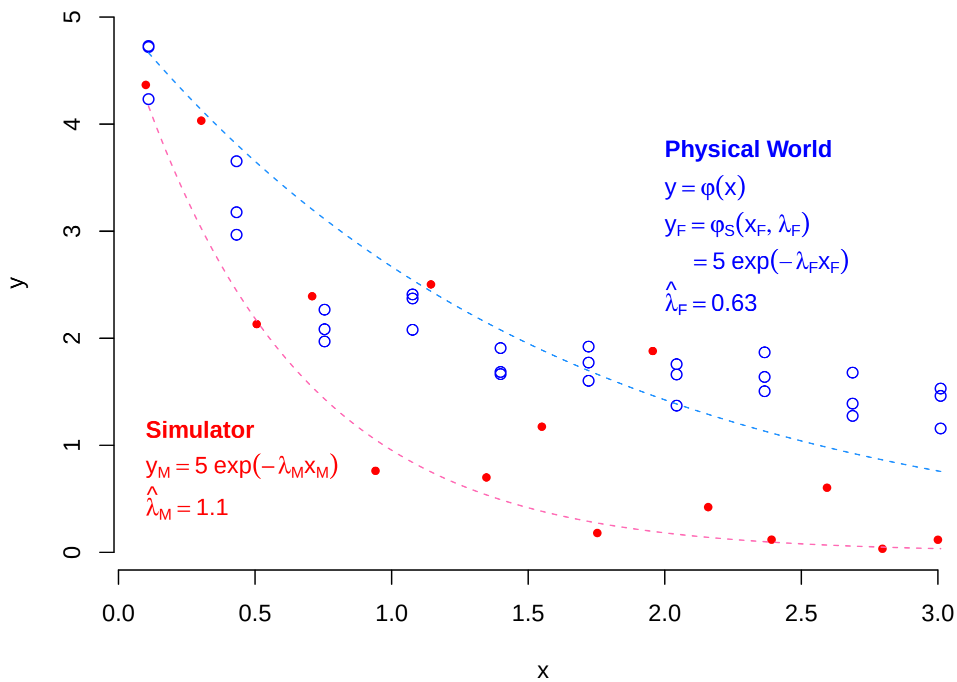
MATHEMATICAL MODELS & COMPUTER CODES

- *Simulator* $\mathcal{S}$ is model for system $\mathcal{W}$ in physical world
- Since $\mathcal{S}$ reproduces $\mathcal{W}$ only imperfectly or incompletely, its output is surrounded by margin of doubt (*uncertainty*)
- In many cases, each evaluation of $\mathcal{S}$, and each observation of $\mathcal{W}$, are very costly: impracticable to characterize uncertainty using conventional Monte Carlo methods
  - Build *Emulator* $\mathcal{E}$ that approximates $\mathcal{S}$ and can express all recognized sources of uncertainty in play

    Including model uncertainty and uncertainty associated with model implementation in computer code
  - Calibrate emulator using only modest number of runs of $\mathcal{S}$ or observations of $\mathcal{W}$

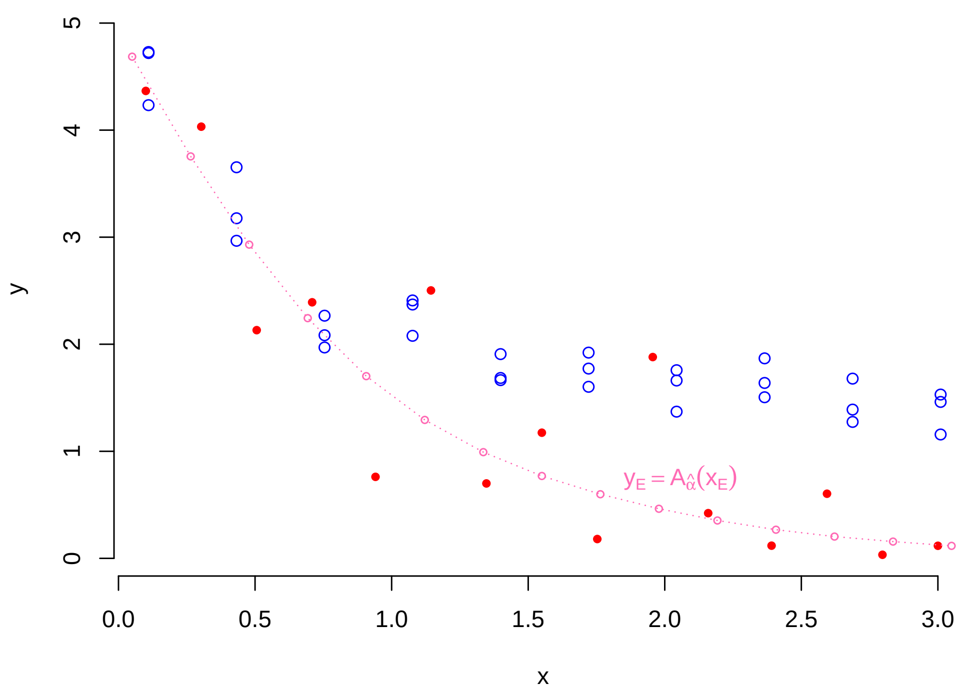
Simulator

- Costly evaluations of simulator $\varphi_{\mathcal{S}}$ to be calibrated using costly observations of physical world



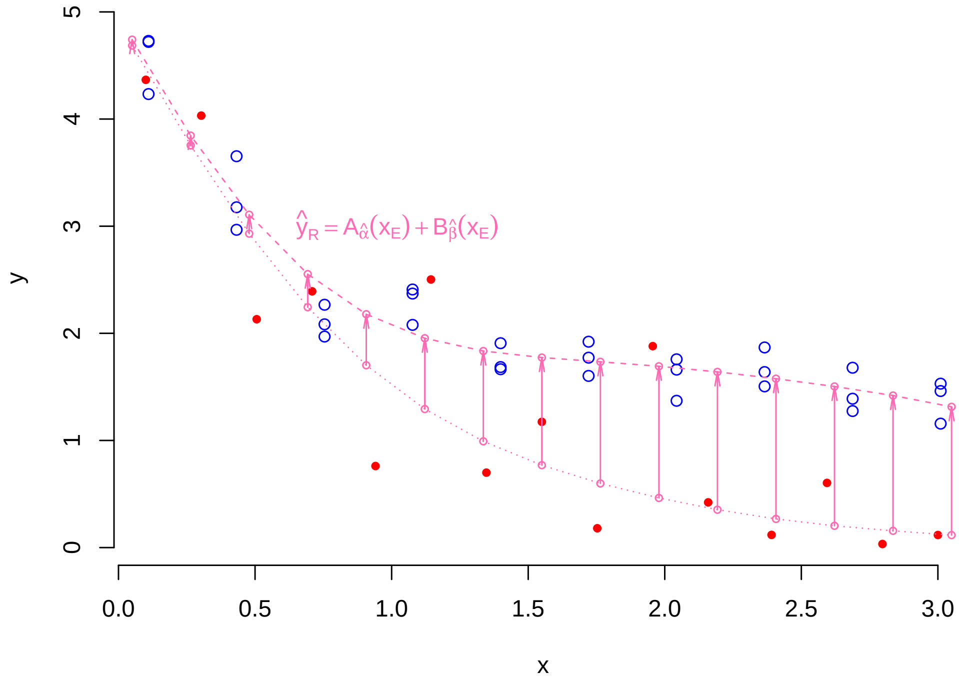
Emulator

- Gaussian random function A_α whose evaluations are much less expensive than simulator's — inherits simulator's bias



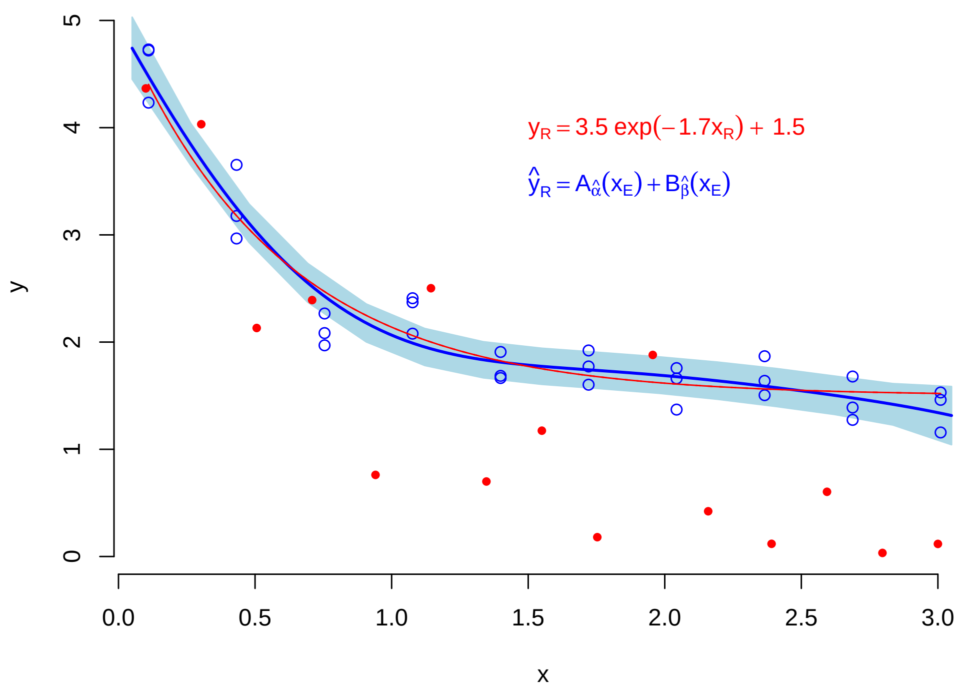
Bias Estimation & Correction

- Bias is persistent effect modeled as another Gaussian random function B_β that is used to correct emulator



Bias-Corrected Emulator

- Uncertainty quantification is by-product of Bayesian procedure used to estimate bias-corrected emulator



Summation

- Uncertainty quantification (for mathematical models and computer codes) can be done using same technical devices used to characterize measurement uncertainty
- Gaussian random functions (of several variables) provide flexible, general purpose emulators
- Bayesian approach (typically employing Markov Chain Monte Carlo sampling) enables uncertainty quantification relying on modest numbers of evaluations of costly simulator and of costly observations of physical world system