Bayesian Methods - 2

Sankaran Mahadevan

Vanderbilt University, Nashville, TN

November 5, 2010

Email: sankaran.mahadevan@vanderbilt.edu

Outline

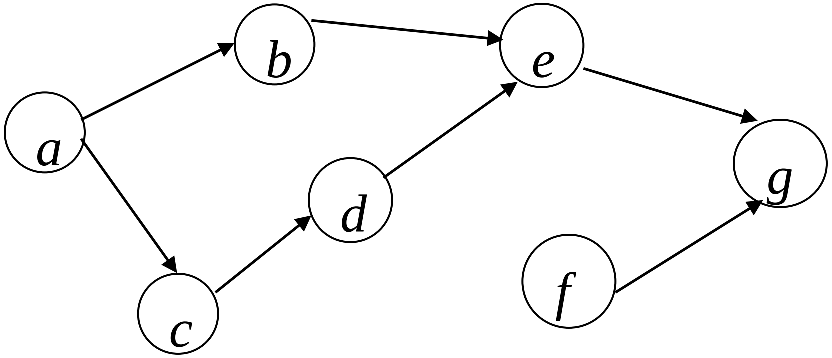
Lecture 1

- Bayes theorem
- Implementation
  - Conjugate distributions
  - Markov Chain Monte Carlo simulation
- Three types of updating
  - Distribution parameters, model inputs, model coefficients

Lecture 2

- Bayes networks
- Model calibration
- Model validation

Bayes network concept



$a, b, \ldots$ component nodes (model inputs, outputs, error terms, experimental data, events)

g - final model output

U - set of all nodes $\{ a, b, \ldots, g \}$

Joint PDF of all nodes

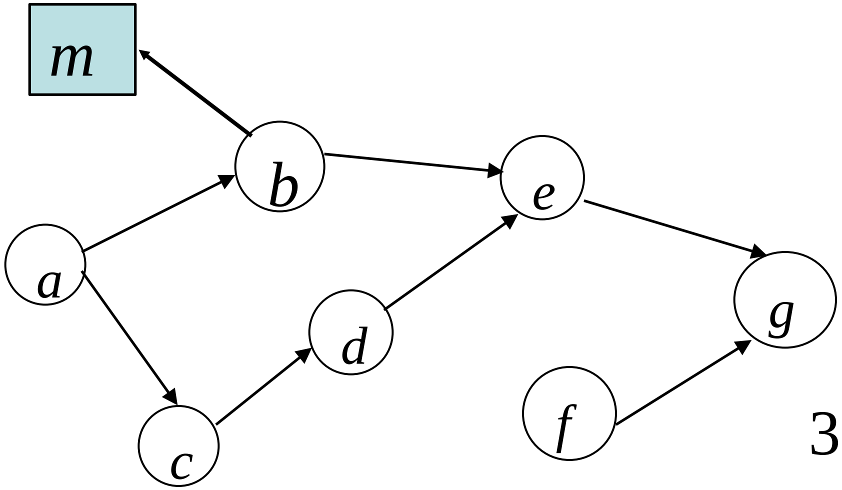
$$\pi(U) = \pi(a) \quad \pi(b|\, a) \quad \pi(c|\, a) \quad \pi(d|\, c) \quad \pi(e|\, b, d) \quad \pi(f\,) \quad \pi(g|\, e, f\,)$$

PDF of final output g

$$\pi(g) = \int \pi(U)\, \mathrm{d}a\, \mathrm{d}b \ldots \mathrm{d}f$$

With new observed data m

$$\pi(U, m) = \pi(U) \quad \pi(m|\, b)$$



Bayes network benefits for UQ

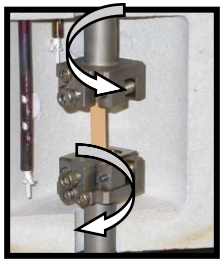
- Include various sources of uncertainty, errors, model predictions and experimental data
- Update all nodes based on new data at any node
- Assist in
  - model calibration
  - model validation
  - extrapolation
  - sensitivity analysis
- Include epistemic uncertainty
  - sparse data
  - imprecise data
  - measurement errors
  - subjective information
  - modeling errors

Example 1: Multiple levels of models & tests

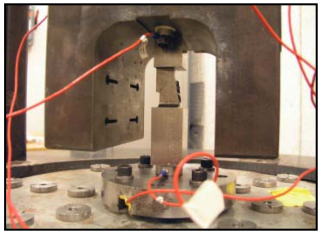
Foam

Joints

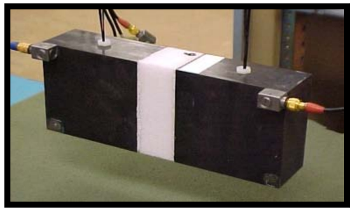
Level 0



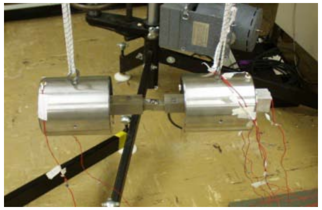
Material characterization



Level 1

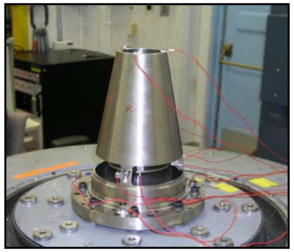


Component level

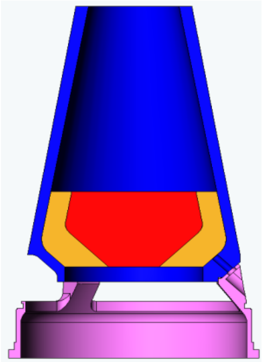


Level 2

Sub-system level



System level



Urbina, Paez, Mahadevan, SDM 2009, SDM 2010

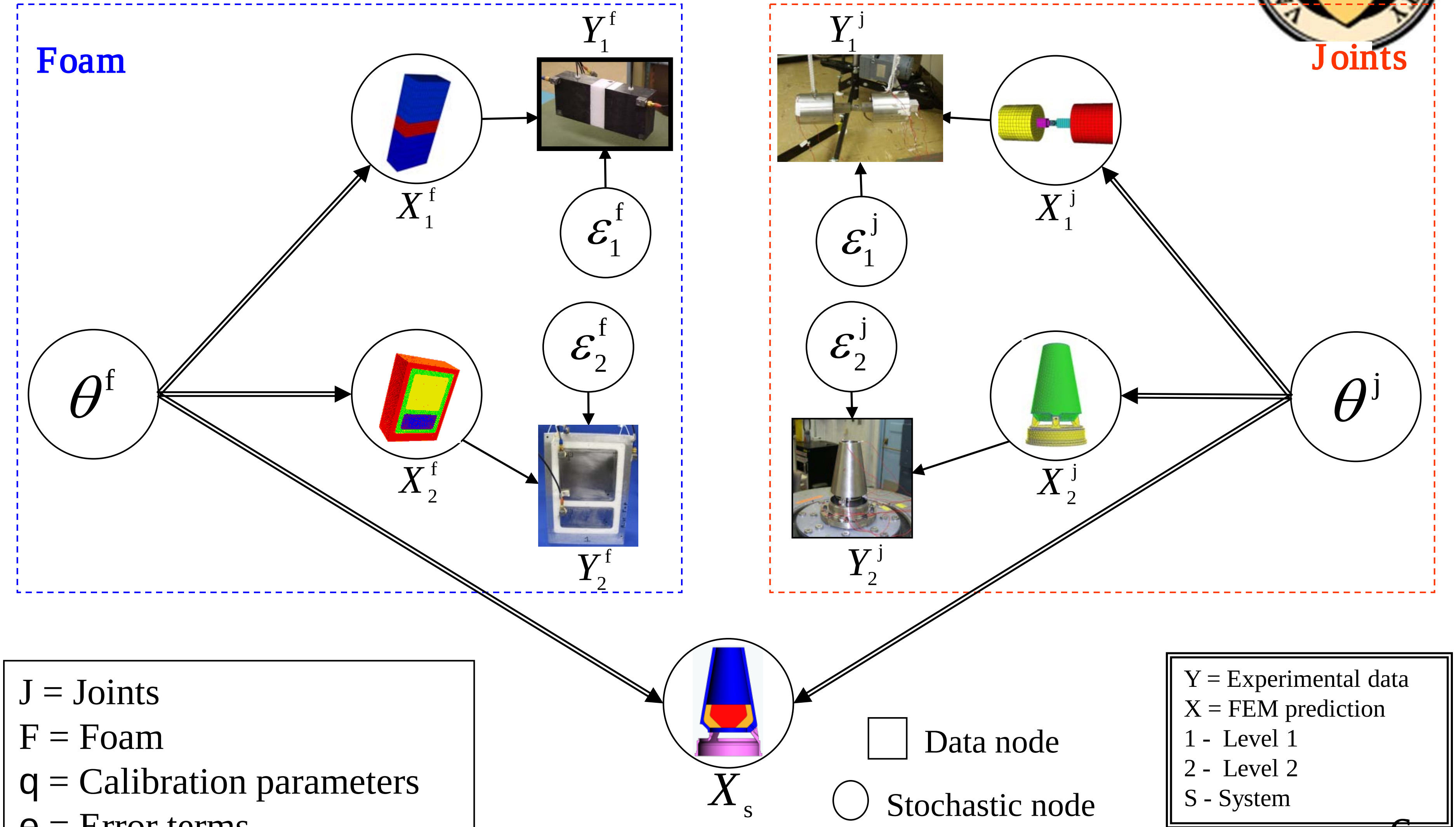
Hardware data and photos courtesy of Sandia National Laboratories

System complexity → Increases

Sources of uncertainty → Increase

Amount of real data → **Decreases**

Bayes Network Implementation



J = Joints
F = Foam
q = Calibration parameters
e = Error terms

☐ Data node
○ Stochastic node

Y = Experimental data
X = FEM prediction
1 - Level 1
2 - Level 2
S - System

Bayesian Calibration

Continuous Case

$$\pi(\theta \mid \mathbf{D}) = \frac{\Pr(\mathbf{D} \mid \theta)\pi(\theta)}{\int \Pr(\mathbf{D} \mid \theta)\pi(\theta)d\theta}$$

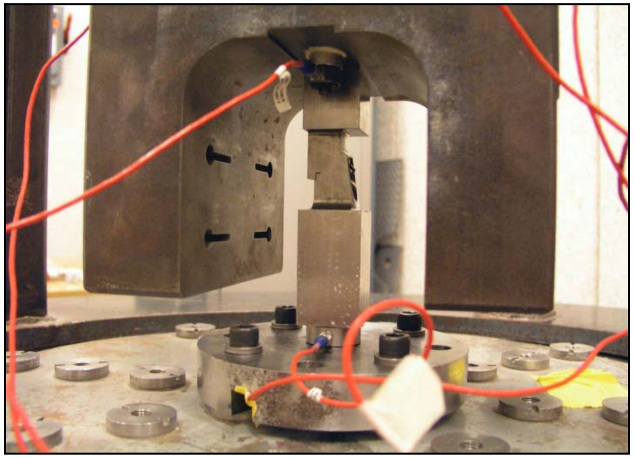
θ : Parameter Being Inferred

D : Observed data on Variable X

Three types of calibration:

1. θ is distribution parameter of random variable X
2. θ is input to a model, whose output is X
3. θ is coefficient of a model, whose output is X

Calibration parameters



Joints

Smallwood energy disssipation model

$$F_j = klin(d_j - d_i) - knon(d_j - d_i)^{npow} + F_i$$

Foam

Modulus of Elasticity E → No experimental data

→ Six expert-specified intervals

Epistemic Uncertainty: Johnson family of distributions

- Flexible family of distributions – normal, lognormal, bounded, unbounded
- Four parameters ξ, λ, γ, δ
- A transformation to standard normal distribution is defined for each of the above four distribution types

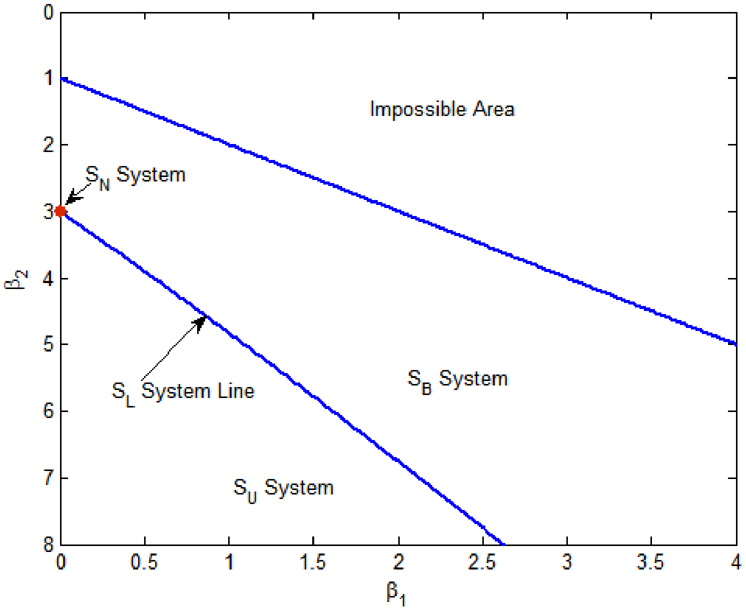
$$z = \gamma + \delta f\left(\frac{x-\xi}{\lambda}\right)$$

$$f(y) = \ln(y), \text{ for lognormal } (\mathrm{S_L}) \text{ distribution}$$

$$= \ln\left[y + \sqrt{y^2+1}\right], \text{ for unbounded } (\mathrm{S_U}) \text{ Jo}$$

$$= \ln[y/(1-y)] \text{ for bounded } (\mathrm{S_B}) \text{ Johnson}$$

$$= y, \text{ for Normal } (\mathrm{S_N}) \text{ distribution}$$

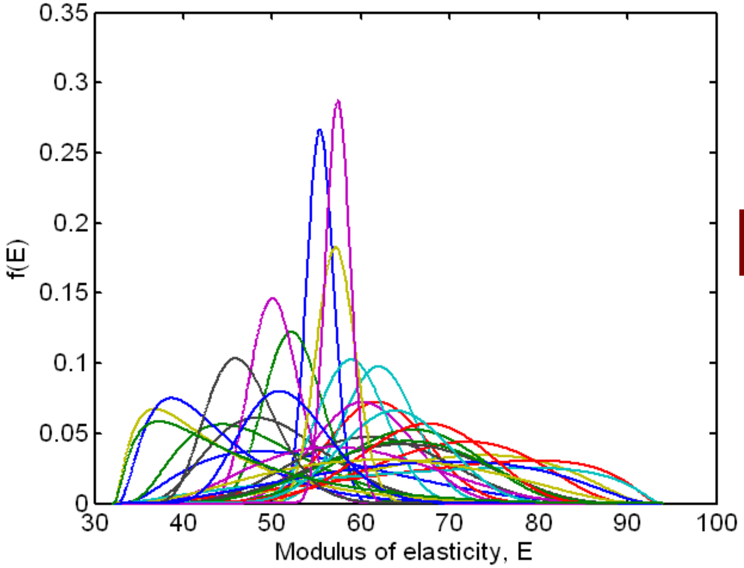


$$\boxed{y = (x-\xi)/\lambda}$$

$$\beta_1 \equiv m_3^2 / m_2^3 \qquad \beta_2 \equiv m_4 / m_2^2$$

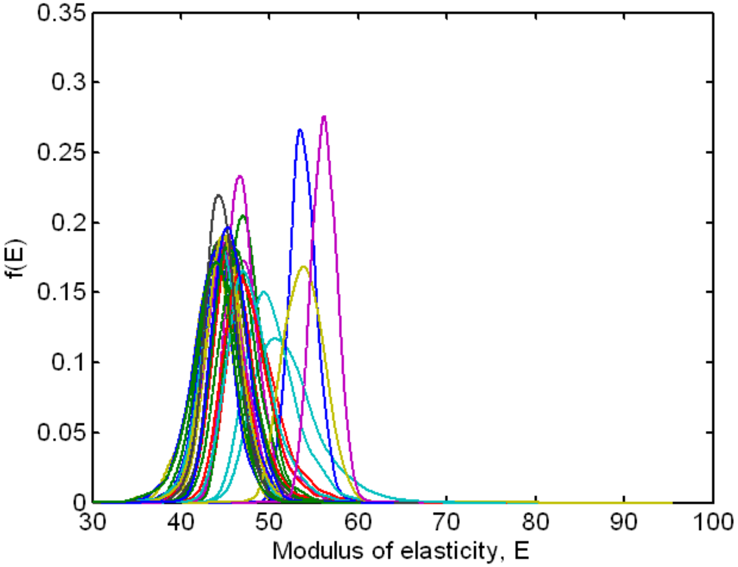
Uncertainty propagation

Prior distributions of E

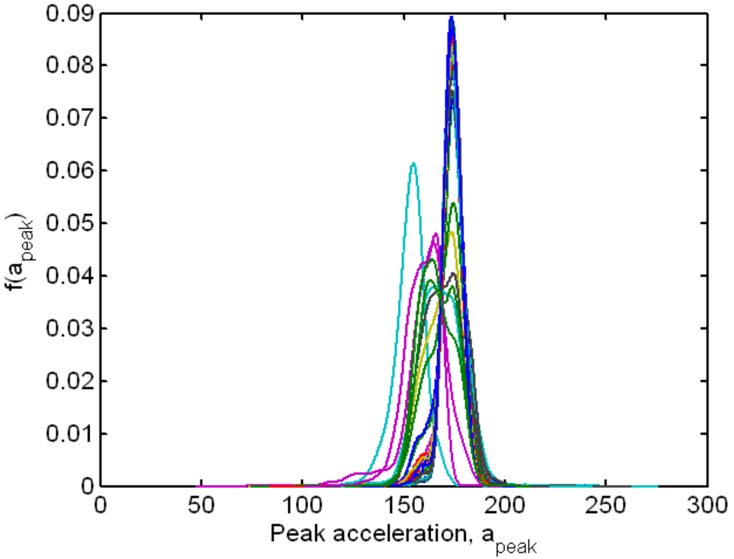


Bayesian updating

Posterior distributions of E



Posterior distribution of system level response



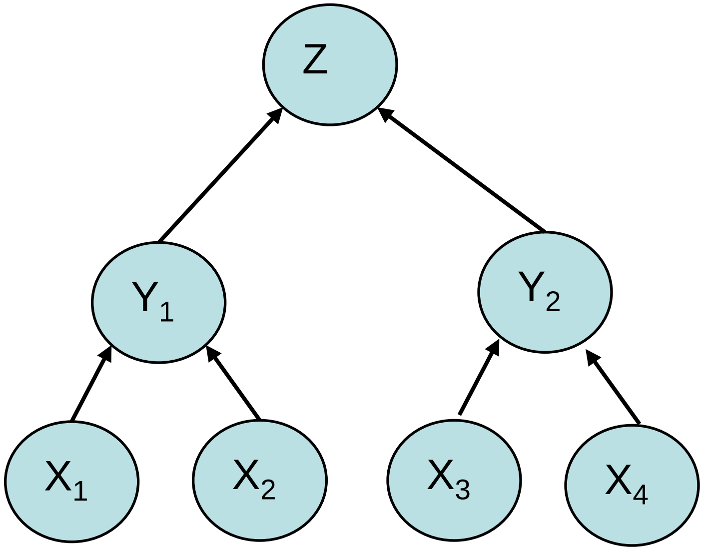
System level response

Resource Allocation using Bayes network

- N_1 , N_2 → number of tests at levels 1 and 2
- Reduction in variance of system output $\Delta V = f(N_1, N_2)$
- Optimization
  - Subject to total cost
  - Find N's so that ΔV is maximum
- Future test → generated by random sampling
- Need to consider multiple random samples
- For each sample → calculate reduction in variance of system output

Example 2: Multi-level models

- $Y_1 = X_1 + X_2$, $Y_2 = X_3 + X_4$
- $Z = Y_1 – Y_2$
- Decide number of tests for Y_1 and Y_2
- Minimize variance of Z, subject to testing budget
- X_1 and X_3 have precise distributions → no update
  - $X_1 \sim N(100,5)$
  - $X_3 \sim N(10, 1)$
- X_2 and X_4 have prior distributions → updated with new data
  - $X_2 \sim N(50, 10)$
  - $X_4 \sim N(15, 4)$



Cost of each test on Y_1 → 10

Cost of each test on Y_2 → 5

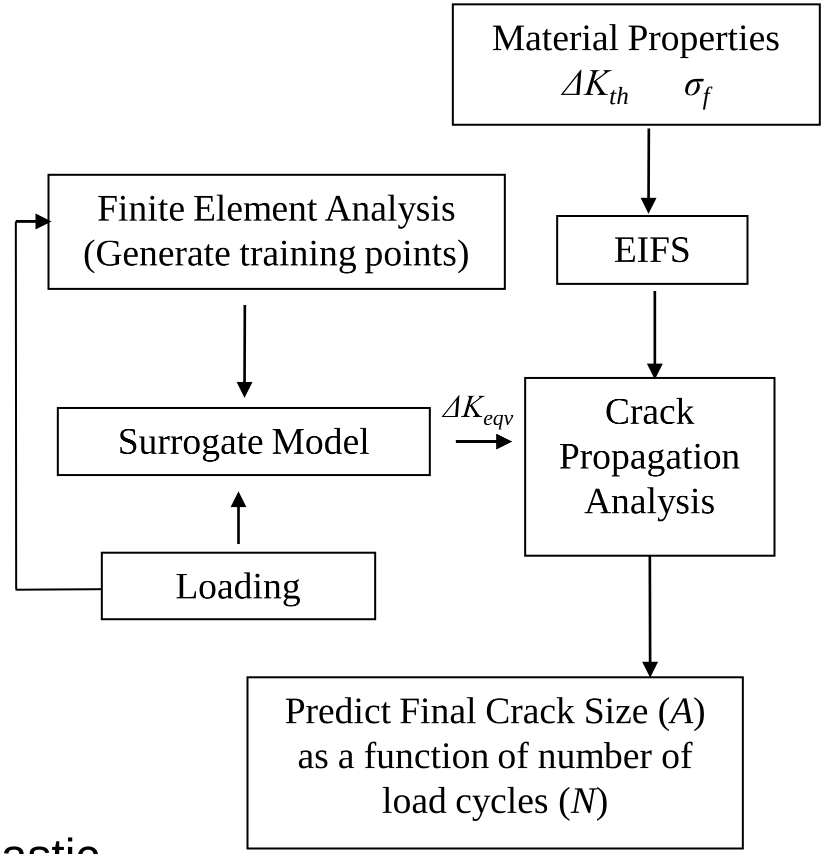
Total budget → 50

OPTIMAL TESTING

$n_1 = 4, n_2 = 2$

Example 3: Connecting multiple models for crack growth prediction

- Physical variability
  - Loading
  - Material Properties
- Data uncertainty
  - Sparse input data
  - Output measurement
- Model uncertainty/errors
  - Finite element discretization error
  - Gaussian process surrogate model
  - Crack growth law
- Complicated interactions
  - Some errors deterministic, some stochastic
  - Combinations could be non-linear, nested, or iterative



Material Properties ΔK_{th} σ_f

Finite Element Analysis (Generate training points)

EIFS

Surrogate Model

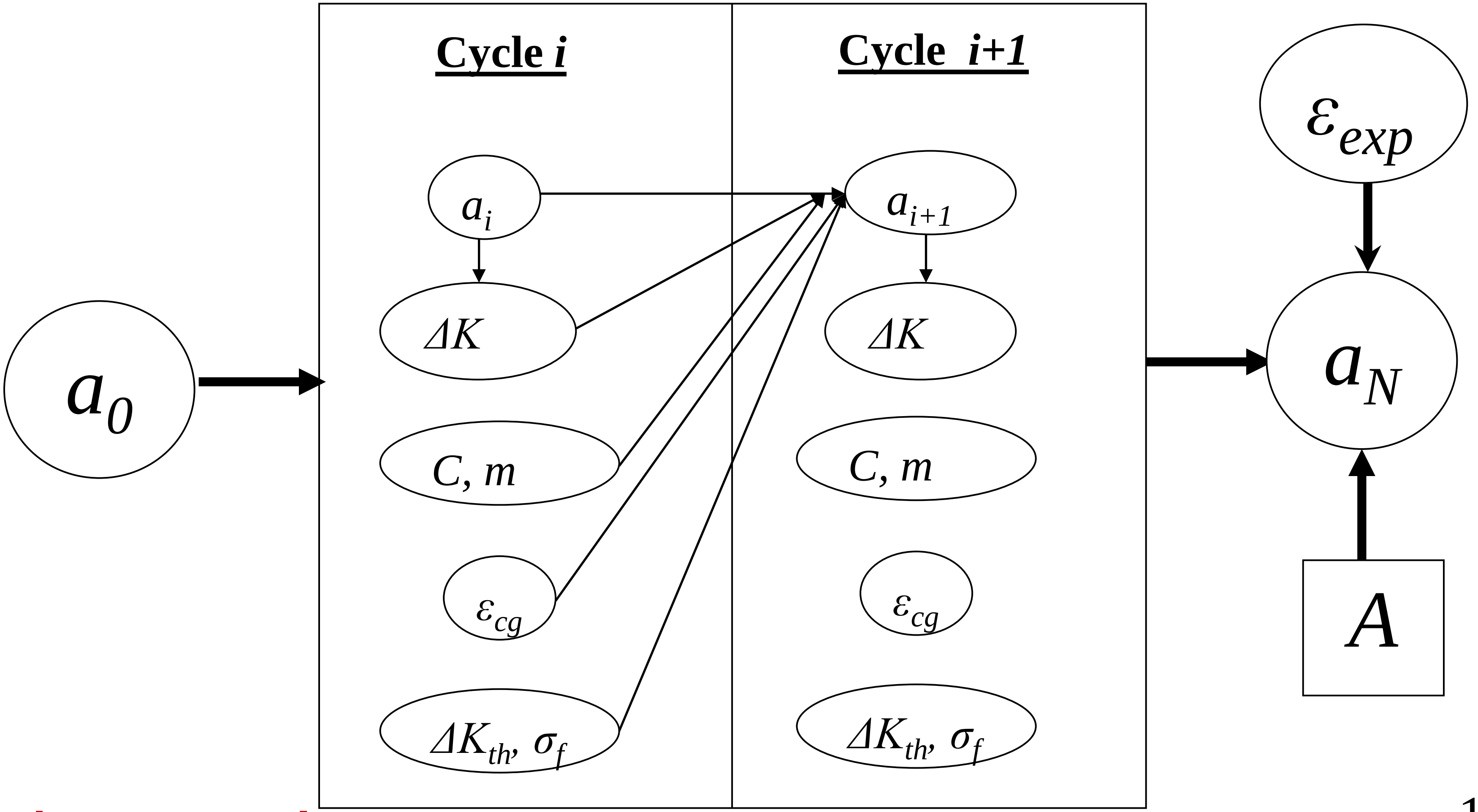
ΔK_{eqv}

Crack Propagation Analysis

Loading

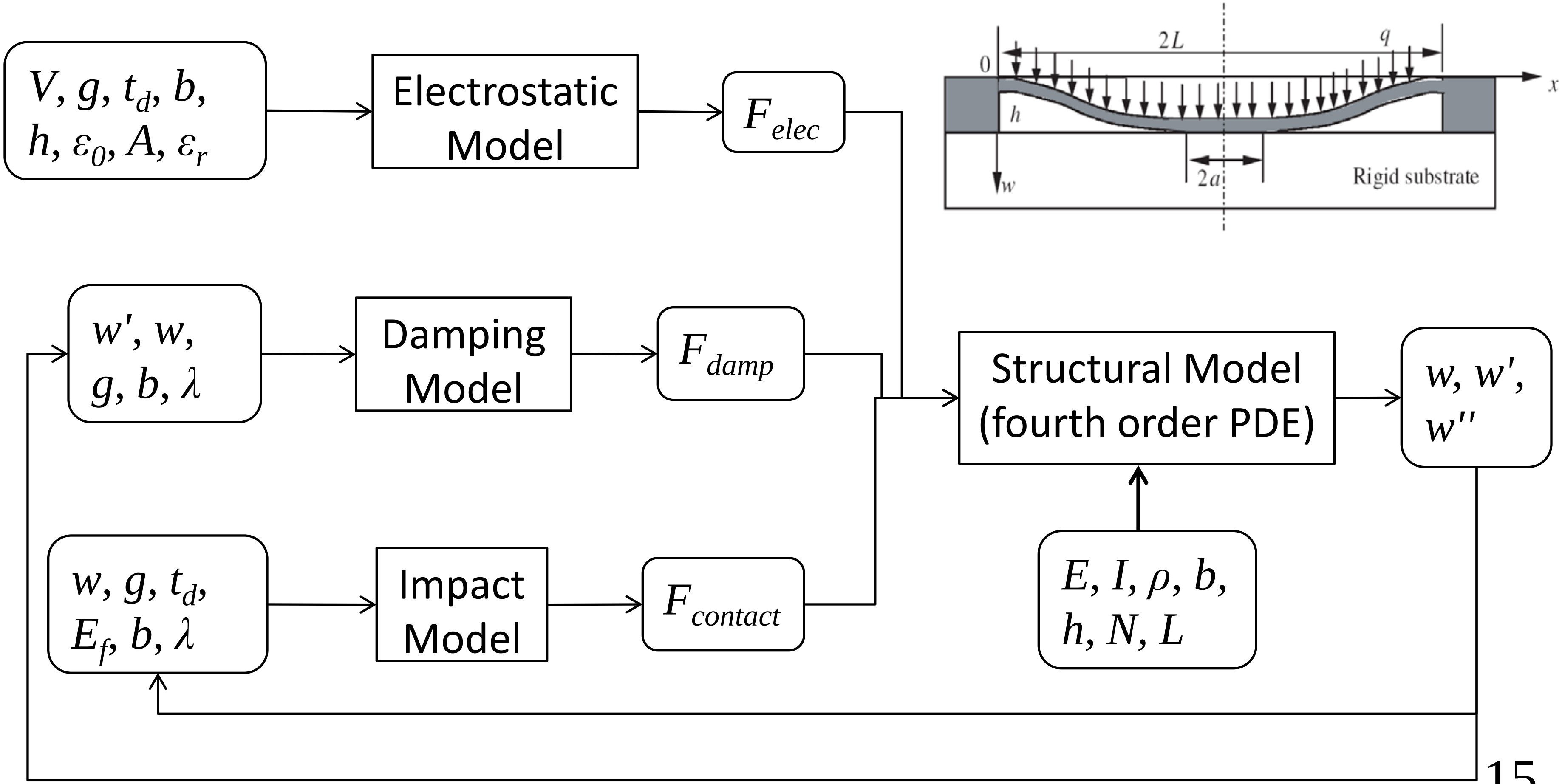
Predict Final Crack Size (*A*) as a function of number of load cycles (*N*)

Dynamic Bayes network for crack growth UQ, calibration, validation



Sankararaman et al, 2010

Example 4: 1-D RF Switch Model



$V, g, t_d, b,$
$h, \varepsilon_0, A, \varepsilon_r$

Electrostatic Model

F_{elec}

$w', w,$
g, b, λ

Damping Model

F_{damp}

$w, g, t_d,$
E_f, b, λ

Impact Model

$F_{contact}$

Structural Model (fourth order PDE)

$E, I, \rho, b,$
h, N, L

$w, w',$
w''

Equations of 1-D RF Switch Model

$$\rho A \frac{\partial^2 w(x,t)}{\partial t^2} + EI \frac{\partial^4 w(x,t)}{\partial x^2} = f(x,t)$$

$$f(x,t) = F_{elec} - F_{damp}$$

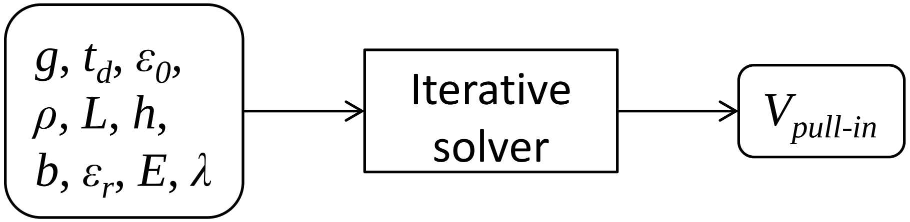
$$F_{elec} = \frac{\varepsilon_0 V^2 A}{2(g + \frac{t_d}{\varepsilon_r} - w)^2}$$

$$F_{damp} = c_f(w)w'$$

$$g - t_d - w \leq 2t_{asperity}$$

E : Young’s modulus of beam
ρ : Density of beam
A : area
V : input voltage, step function
g : air gap
t_d : dielectric thickness
ε_0 : permittivity of free space
ε_r : dielectric constant
w : vertical displacement of mass
w' : vertical velocity of mass
$t_{asperity}$: asperity thickness
$c_f(w)$: damping coefficient

Black Box Representation of 1-D RF Switch Pull-In Voltage Calculation

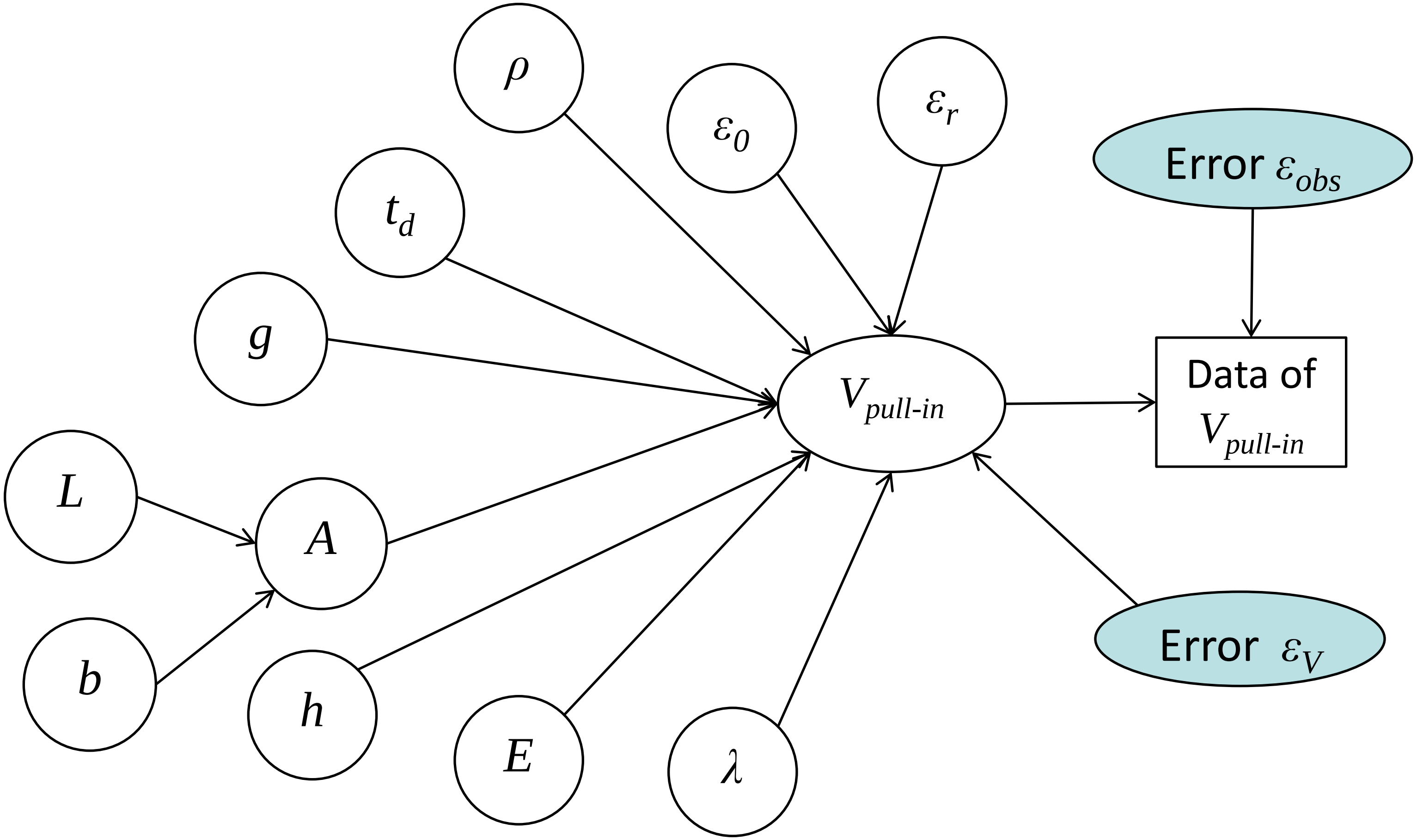


g, t_d, ε_0,
ρ, L, h,
b, ε_r, E, λ

Iterative solver

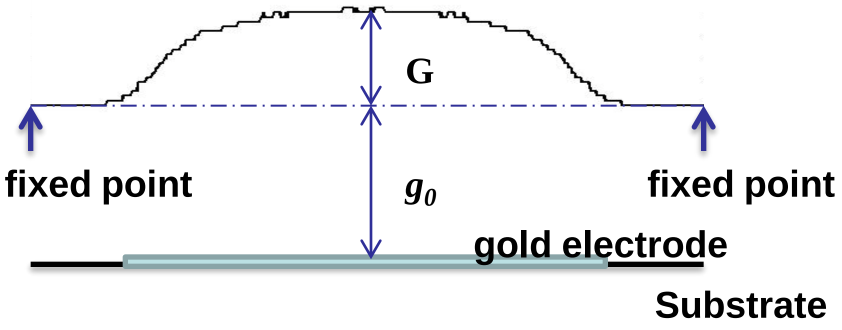
$V_{pull\text{-}in}$

Bayes Network for a 1-D Switch – Black Box Version



Calibration of Young's Modulus *E*

- Fixed-fixed beam, distribution of *E* is unknown



$$V_{pl} = f(L,b,h,g;E) + \varepsilon_f$$
$$V_{pl-obs} = V_{pl} + \varepsilon_{obs}$$

  - L, b, h : geometry of the beam (length, width, thickness)
  - g : initial gap = $(G + g_0)$
  - ε_f : calibration residual ~ N(0,σ_f), σ_f is unknown and needs calibration
  - ε_{obs} : measurement error ~ N(0,σ_{obs})

Likelihood Function: Model Residual and Measurement Error

- Construct the conditional probability of model prediction V_{pl} given E and σ_f

$$V_{pl} = f(L, b, h, g; E) + \varepsilon_f$$

$$\rightarrow V_{pl} \sim \mathrm{N}(f(L, b, h, g; E), \sigma_f) \rightarrow \pi(V_{pl} \mid E, \sigma_f)$$

- Construct the conditional probability of observation given V_{pl}

$$V_{pl-obs} = V_{pl} + \varepsilon_{obs}$$

$$\rightarrow V_{pl\text{-}obs} \sim \mathrm{N}(V_{pl}, \sigma_{obs}) \rightarrow \pi(V_{pl\text{-}obs} \mid V_{pl})$$

- Likelihood function of E and σ_f

$$L(E, \sigma_f) = \Pr(V_{pl-obs} \mid E, \sigma_f) = \int \pi(V_{pl-obs} \mid V_{pl}) \pi(V_{pl} \mid E, \sigma_f) dV_{pl}$$

Imprecise Experimental Data – Interval Data

- Voltage is increased in 5-volt steps during the measurement, and the pull-in voltage is reported within the range recorded → interval data
- Incorporate interval data into likelihood function:

$$L(\theta) = \Pr(\mathbf{D} \mid \theta) = \prod_{i=1}^{N} \Pr(D_i \mid \theta)$$

$$= \prod_{i=1}^{N} \Pr(V_{1i} \le V_{pl-obs} \le V_{2i} \mid \theta)$$

$$= \prod_{i=1}^{N} \int_{V_{1i}}^{V_{2i}} \pi(V_{pl-obs} \mid \theta) dV_{pl-obs}$$

Sankararaman and Mahadevan, RESS, 2011

Likelihood Function for Interval Data

- Interval data $[V_{1i}, V_{2i}]$ ← interval width = 5 volts
- Joint likelihood function of E and σ_f

$$L(E,\sigma_f) = \Pr(\mathbf{D} \mid E,\sigma_f)$$

$$= \prod_{i=1}^{N} \Pr(D_i \mid E,\sigma_f)$$

$$= \prod_{i=1}^{N} \int_{V_{1i}}^{V_{2i}} \left(\int \pi(V_{pl-obs} \mid V_{pl})\pi(V_{pl} \mid E,\sigma_f)dV_{pl} \right) dV_{pl-obs}$$

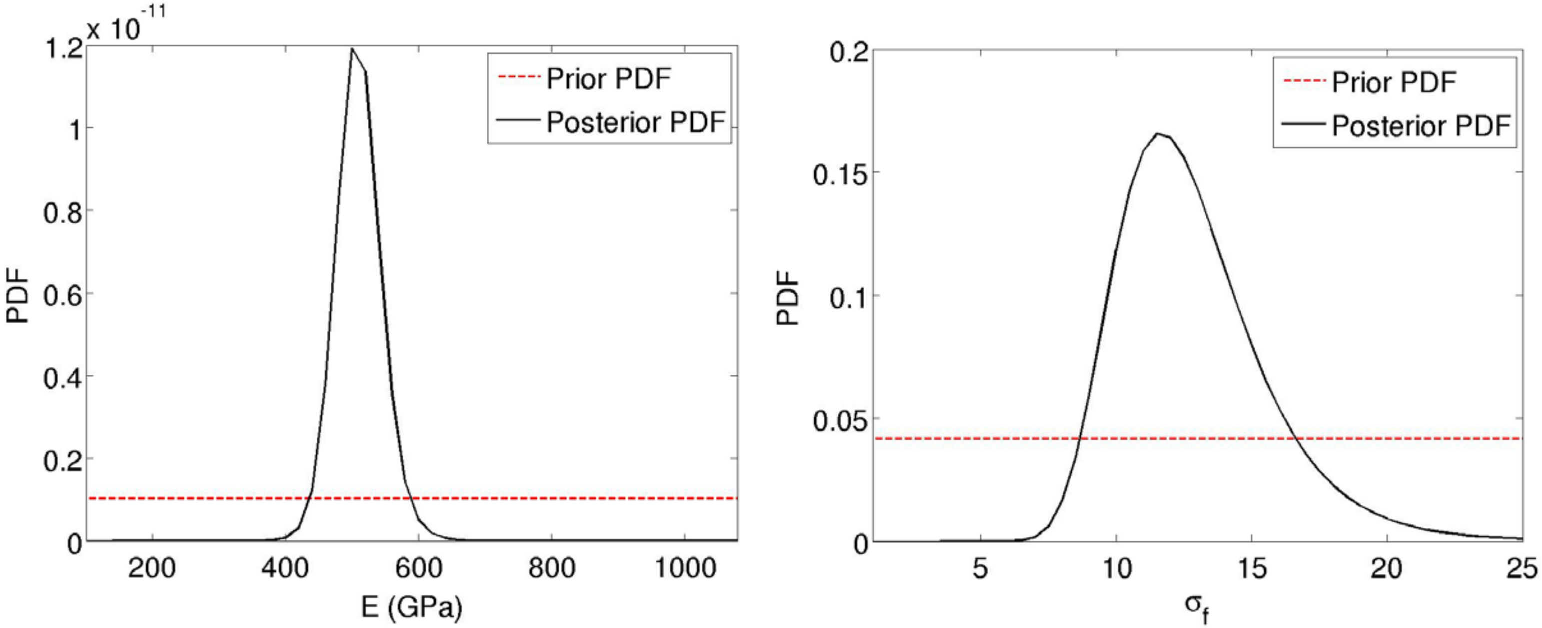
- Marginal likelihoods of E and σ_f

$$L(E) = \int L(E,\sigma_f)\pi(\sigma_f)d\sigma_f$$

$$L(\sigma_f) = \int L(\sigma_f, E)\pi(E)dE$$

$$\xrightarrow{\text{Calibration}} \quad \pi(\theta \mid \mathbf{D}) = \frac{\Pr(\mathbf{D} \mid \theta)\pi(\theta)}{\int \Pr(\mathbf{D} \mid \theta)\pi(\theta)d\theta}$$

Results of Bayesian Model Calibration

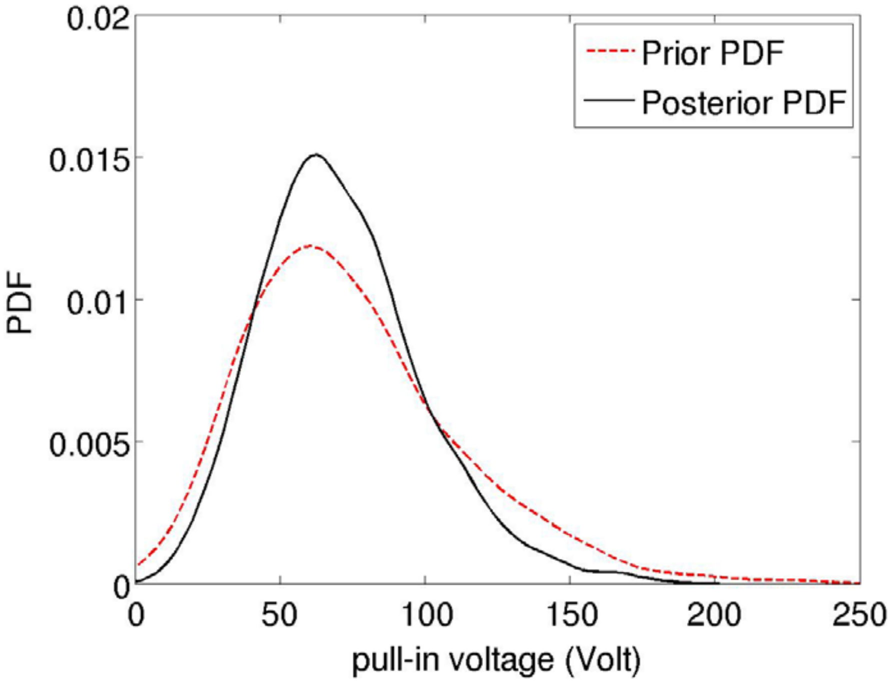


Note:

The prior PDF's of E and σ_f are assumed uniform (non-informative priors)

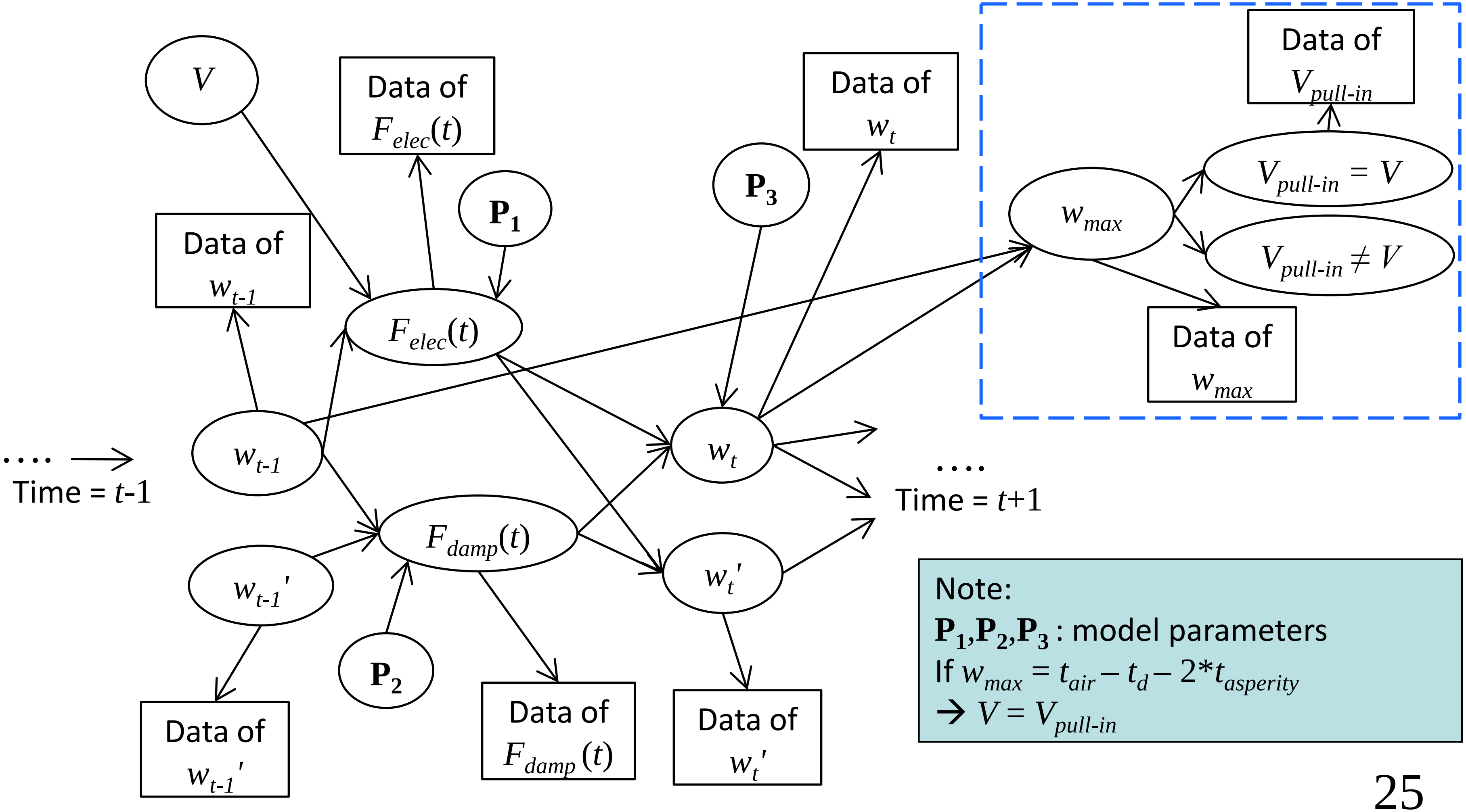
Maximum likelihood estimates: E_{MLE} = 500 GPa, $\sigma_{f\text{-}MLE}$ = 11.5

Prior and Posterior Prediction of Pull-In Voltage



	Mean	Standard Deviation
Prior Prediction (Volt)	74.67	36.80
Posterior Prediction (Volt)	71.26	27.67

Detailed Bayes Network without Error Terms



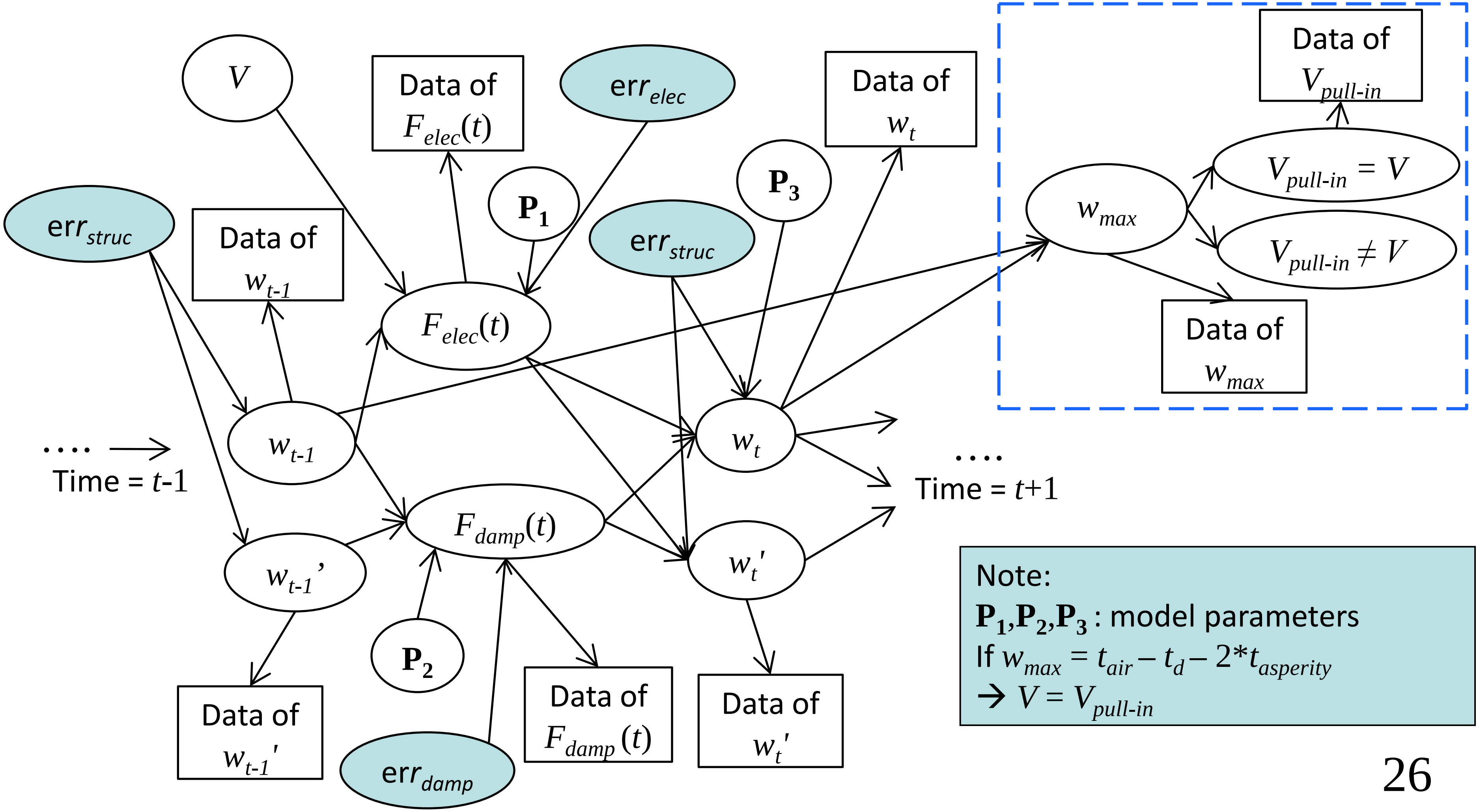
Note:

$\mathbf{P_1}$,$\mathbf{P_2}$,$\mathbf{P_3}$: model parameters

If $w_{max} = t_{air} - t_d - 2 * t_{asperity}$

→ $V = V_{pull\text{-}in}$

Bayes Network with Error Terms



Note:

$\mathbf{P_1}$,$\mathbf{P_2}$,$\mathbf{P_3}$: model parameters

If $w_{max} = t_{air} - t_d - 2*t_{asperity}$

→ $V = V_{pull\text{-}in}$

Model Validation

- Are model predictions consistent with the experimental observations, given a defined level of uncertainty?
- Graphical comparison
  - Compare model predictions and experimental data graphically
- Classical hypothesis testing
  - z-test or t-test (to compare mean values)
  - Chi-Square test (to compare variances)
  - Chi-square, K-S, Anderson-Darling, Cramer tests (for distributions)
- Bayesian hypothesis testing
  - Bayes factor (applicable for well-characterized experiments and uncharacterized experiments)

Bayesian Hypothesis Testing

- From Bayes theorem

$$\frac{\Pr(\mathrm{H_0}:\text{model is correct}\,|\,\mathbf{D})}{\Pr(\mathrm{H_1}:\text{model is incorrect}\,|\,\mathbf{D})}=\frac{\Pr(\mathbf{D}\,|\,\mathrm{H_0}:\text{model is correct})}{\Pr(\mathbf{D}\,|\,\mathrm{H_1}:\text{model is incorrect})}*\frac{\Pr(\mathrm{H_0})}{\Pr(\mathrm{H_1})}$$

- Validation metric → **Bayes factor**

$$B=\frac{\Pr(\mathbf{D}\,|\,\mathrm{H_0}:\text{model is correct})}{\Pr(\mathbf{D}\,|\,\mathrm{H_1}:\text{model is incorrect})}$$

$$=\frac{\prod_i^N \Pr(D^i\,|\,\mathrm{H_0})}{\prod_i^N \Pr(D^i\,|\,\mathrm{H_1})}=\prod_i^N B^i$$

Confidence measure $= B\,/\,B+1$

  - $B > 1$ → data supports the model
  - $B < 1$ → data does not support the model

Bayesian Hypothesis Testing Implementation

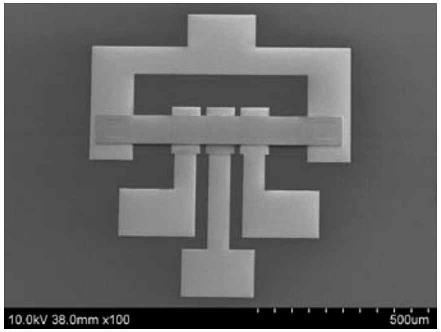
Well-characterized

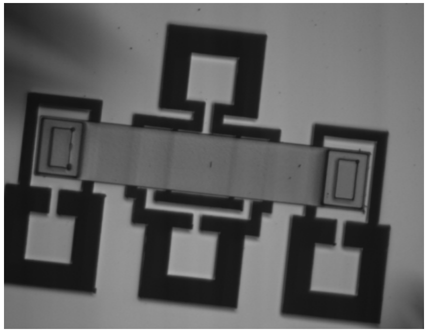
Uncharacterized

$$B^i = \frac{\Pr(D^i \mid H_0)}{\Pr(D^i \mid H_1)} = \frac{\int_{y_1}^{y_2} \Pr(y_D^i \mid y)\pi_0(y \mid \theta)dy}{\int_{y_1}^{y_2} \Pr(y_D^i \mid y)\pi_1(y \mid \theta)dy} = \frac{\iint \Pr(y_D^i \mid y)\pi_0(y \mid \theta_0)\pi(\theta_0)dyd\theta_0}{\iint \Pr(y_D^i \mid y)\pi_1(y \mid \theta_1)\pi(\theta_1)dyd\theta_1}$$

- $\pi_0(y|\theta)$: conditional PDF of y under the null hypothesis H_0 given input θ
  - $\pi_0(y|\theta)$ is a continuous function if model output is uncertain given θ
  - $\pi_0(y|\theta)$ is a Dirac Delta function if model output is deterministic given θ
- $\pi_1(y|\theta)$: conditional PDF of y (assumed uniform) under the alternative hypothesis H_1 given input θ
- $\Pr(y^i_D|y)$: Likelihood function → conditional probability of observing the data y^i_D given model prediction y (since $y_D = y + \varepsilon_{obs}$)
- y_1, y_2: integration limits, affecting the value of Bayes factor due to the uniform assumption for $\pi_1(y|\theta)$

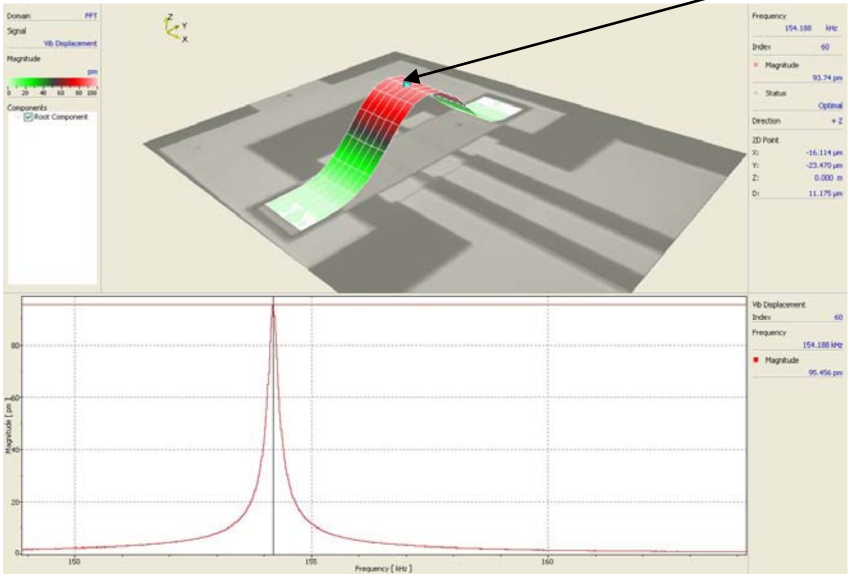
Example 5: Validation of damping factor prediction





Well-characterized experiments

Measurement made at center of beam for ring down data; beam excited electrostatically



Domain FFT
Signal Vib Displacement
Magnitude pm
Components
☑ Root Component

Frequency 154.188 kHz
Index 60
Magnitude 93.74 pm
Status Optimal
Direction + Z
2D Point
X: -16.114 µm
Y: -23.470 µm
Z: 0.000 m
D: 11.175 µm

Vib Displacement
Index 60
Frequency 154.188 kHz
Magnitude 95.456 pm

- **7 Devices**
- **4 pressures from 18784-66612 Pa**
- **5 replicates of data**
- **Nickel beam**
  Length = 395.34 um
  Width = 120 um

Damping Factor Model

- Surrogate model of damping factor – Third order polynomial chaos expansion

$$C = \sum_{i}^{n} a_i \phi_i(t, g, f) + \varepsilon_r$$

t: thickness
g: air gap
f: frequency

a_i: coefficient
Φ_i: i-th basis
ε_r: residual

$$\boldsymbol{\varphi}(t, g, f) = [1, t, g, f, t^2, g^2, f^2, t*g, t*f, g*f, \ldots$$
$$t^3, g^3, f^3, t^2*g, t*g^2, t^2*f, t*f^2, g^2*f, g*f^2]$$

Construction of likelihood functions

- Damping model

$$C_m = G(t, g, f) + \varepsilon_r$$

  - ε_r is the residue term ~ N(0,σ_r)
  - Input variables t, g, and f are measured for each data point
  - Model output C_m ~ N($G(t, g, f)$, σ_r) → $\pi_0(C_m \mid t, g, f)$

- Construct likelihood function of model prediction

$$C_{obs} = C_m + \varepsilon_{obs}$$

  - ε_{obs} is the measurement error ~ N(0,σ_{obs})
  - C_{obs} ~ N(C_m, σ_{obs}) → Pr($C_{obs} \mid C_m$) , for a particular value of C_m

- $y_1 = 0$, $y_2 = 1$ ← under-damped system

Bayes Factor Results (well-characterized experiment)

- Four different surrogate models corresponding to four different pressures
- B^i is computed at every experimental site (140 in total, 35 for each pressure)
- $N_{B_i} > 1$ = Number of experimental sites at which B^i is larger than 1

	18784 Pa	28643 Pa	43564 Pa	66612 Pa
min(B^i)	2.40e-6	0.001	0.202	1.48e-4
Max(B^i)	17.5	18.4	18.4	18.4
$N_{B_i} > 1$	30	30	33	21
B	1.33e7	1.92e26	1.17e32	2.28

- **RESULT**: Models are supported by data at most experimental sites

Summary

- Bayes network incorporates models and available data at different levels of complexity
- Both aleatoric and epistemic uncertainties, and errors, can be included
- System level model prediction uncertainty quantified
- Can be used for resource allocation
- Calibration can use data at multiple levels
- Validation → Bayes factor gives a quantitative measure of confidence in model prediction

References

1. Jaynes, E. T., *Probability Theory: The Logic of Science*, Cambridge University Press, 2003.
2. Lee, P. M., *Bayesian Statistics: An Introduction*, Second Edition. Wiley, 1997.
3. Leonard, T., and J. S. J. Hsu, *Bayesian Methods*. Cambridge University Press, 1999.
4. Mahadevan, S., Zhang, R., and N. Smith, “Bayesian Networks for System Reliability Reassessment,” *Structural Safety*, Vol. 23 (3), pp. 231-251, 2001.
5. Zhang, R., and S. Mahadevan, "Bayesian Methodology for Reliability Model Acceptance," *Reliability Engineering and System Safety*, Vol. 80, No. 1, pp. 95-103, 2003.
6. Mahadevan, S., and Rebba, R., “Validation of Reliability Computational Models Using Bayes Networks,” *Reliability Engineering and System Safety*, Vol. 87, No. 1, pp. 223-232, 2005.
7. Rebba, R., and S. Mahadevan, "Validation and Error Estimation of Computational Models," *Reliability Engineering and System Safety*, Vol. 91, No. 10-11, pp. 1390-1397, 2006.
8. Jiang, X., and S. Mahadevan, “Bayesian Risk-Based Decision Method for Model Validation Under Uncertainty,” *Reliability Engineering and System Safety*, Vol. 92, No. 6, pp. 707-718, 2007.
9. Jiang, X. and S. Mahadevan, "Bayesian Validation Assessment of Multivariate Computational Models," *Journal of Applied Statistics*, Vol. 35, No. 1, pp. 49-65, 2008.
10. McDonald, M.P., Zaman, K., Rangavajhala, S., and S. Mahadevan, “A Probabilistic Approach for Representation of Interval Uncertainty,” Reliability Engineering and System Safety, 2010 (available online).