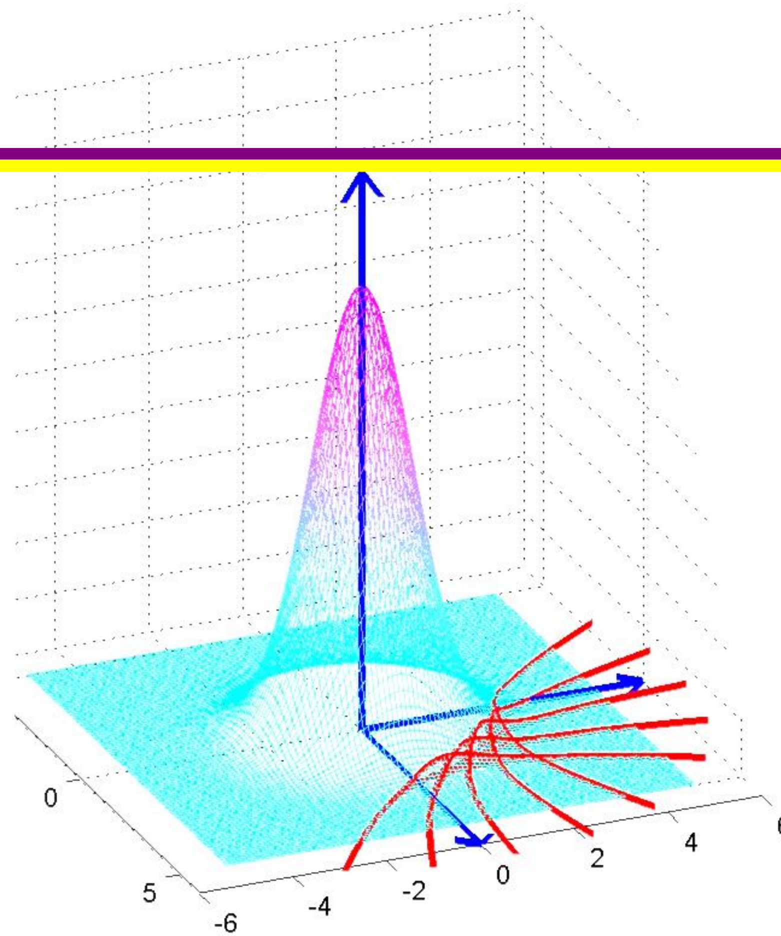
A Probabilistic Framework for solving Inverse Problems



Lambros S. Katafygiotis, Ph.D.

OUTLINE

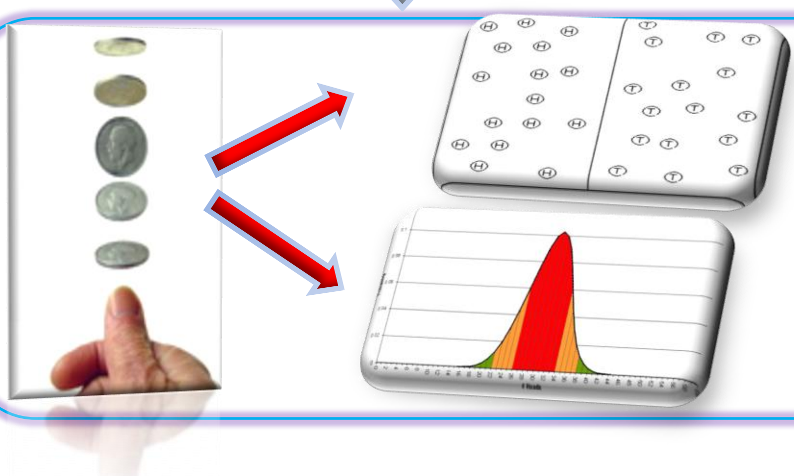
- Introduction to basic concepts of Bayesian Statistics
- Inverse Problems in Civil Engineering
- Probabilistic Model Updating
- Solutions based on Asymptotic Results
- Solutions based on MCMC simulations
- MCMC simulations for reliability estimation
- Damage Detection
- Concluding remarks

PROBABILISTIC MODELLING OF UNCERTAINTIES

Uncertainty Sources: (i) **Modeling error** of physical phenomena ;
(ii) Uncertainties due to **incomplete information**;
(iii)Uncertainties due to **measurement noise.**

‘Frequentist’ Statistics

Probability: A relative frequency of occurrences
(drawing conclusions from sample data)



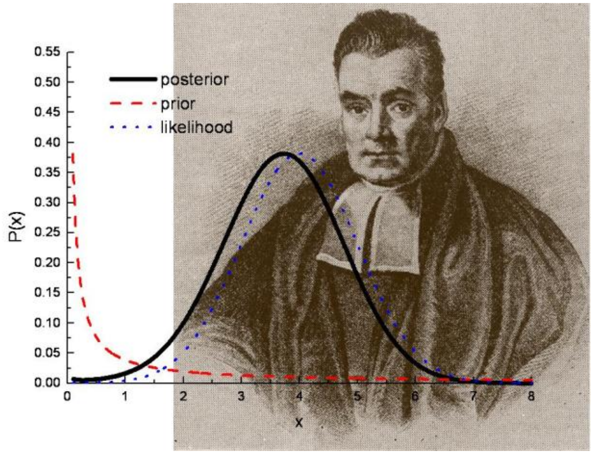
Bayesian Statistics (Cox 1961)

Probability: A measure of plausibility
(A personal degree of belief in a proposition)
Allows us to talk about probability of a parameter, or probability of a model

Bayes’ theorem plays an increasingly prominent role in statistical applications!

Bayes' Theorem

P(A,B)=P(A|B)P(B)=P(B|A)P(A)
Therefore, P(A|B)=P(B|A)P(A)/P(B)



$$p(\theta|D,M) = cp(D|\theta,M)p(\theta|M)$$

- $p(\theta|D,M)$ is the updated or "posterior" PDF of the model parameters given the measured modal data D and the assumed model class M.
- $p(D|\theta,M)$ is the likelihood function that represents the PDF of the observed data given the parameters θ and the model class M.
- $p(\theta|M)$ is the initial or "prior" PDF of the parameters given the model class.
- c is a normalizing constant that ensures that the posterior PDF integrates to one.

- Not only the optimal estimates (most probable values) can be determined but also their associated uncertainties can be quantified
- The associated uncertainties are expressed through a posterior PDF which is interpreted as a measure of plausibility
- Note that Bayes' formulas involve conditional probabilities

BAYESIAN MODEL SELECTION

The posterior probabilities of the various model classes given the data D is :(Beck and Yuen 2004)

Evidence of Model Class

$$P(M_i \mid D) = \frac{p(D \mid M_i)\, P(M_i)}{p(D \mid M_{Fam})}$$

Prior Probability

Optimal model class M_{best} is selected as the one that maximizes $P(M_i \mid D)$

Structural Monitoring and Model Updating

- Detect, localize and assess damage in structures
  - Inevitable aging and degradation resulting from operational environment
  - Damage due to natural disasters, such as earthquakes and hurricanes or man made actions, such as terrorist attacks, accidents, etc.
- Validate structural designs and evaluate structural performance
- Monitor and control construction process
- Characterize loads in situ and assess load carrying capacities
- Vibration control

Structural Monitoring Employing Wireless Sensor Networks

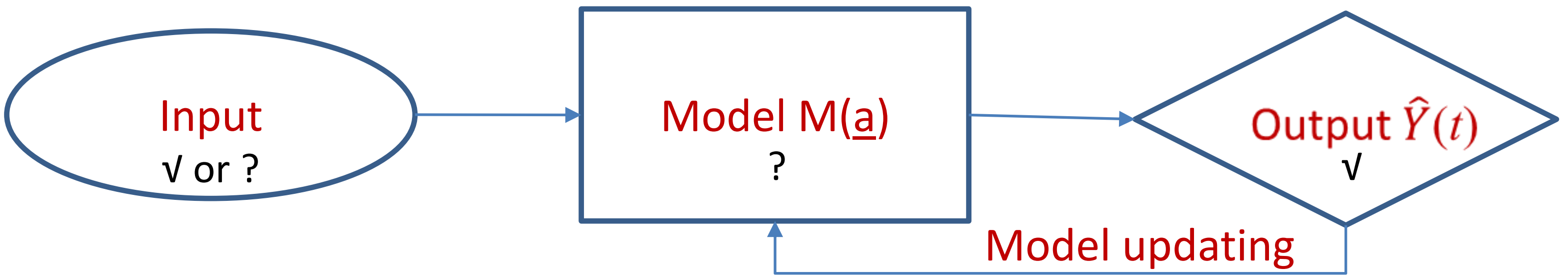
Wired structural monitoring system

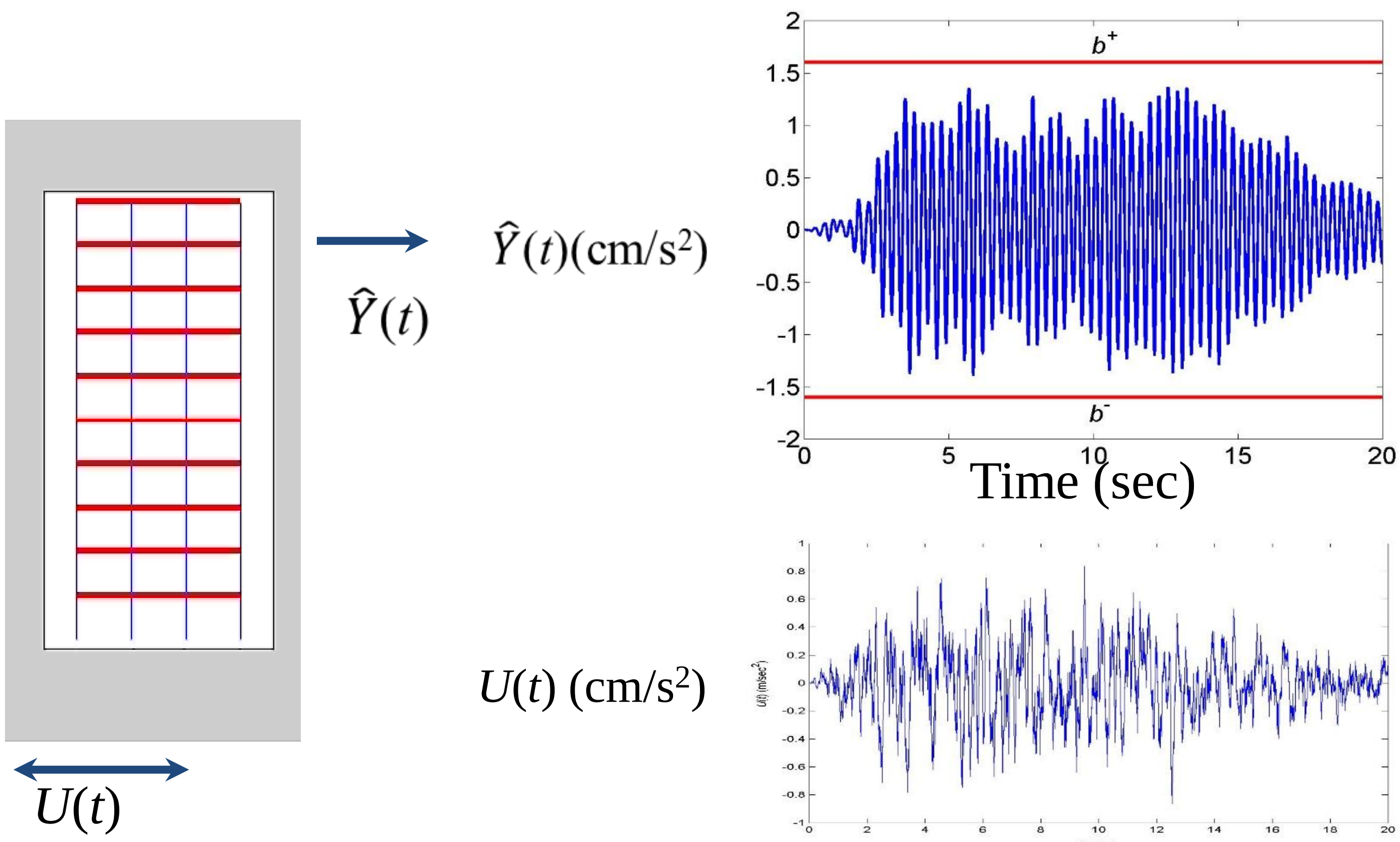
- Pros:
  - High-fidelity data can be achieved
- Cons:
  - Time consuming due to required cabling installation
  - All data are collected and processed at a central station, making it costly to store and process such huge amounts of data

Wireless structural monitoring system

- Pros:
  - Flexible and easy deployment
  - Able to locally process data within each sensor and transmit only important information to the central station
  - Reduced amount of data to be collected and able to distribute the computing burden to all sensors
- Cons:
  - Hardware limitations (limited processing and storage capacity)
  - Sensing synchronization errors
  - Power constraints due to battery limitations

Model Updating



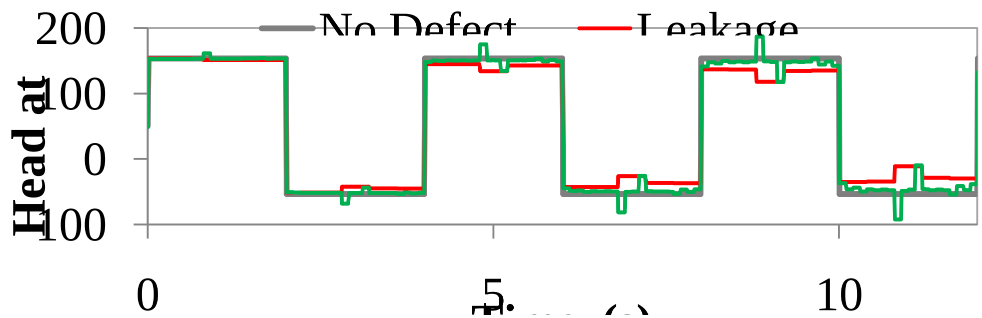


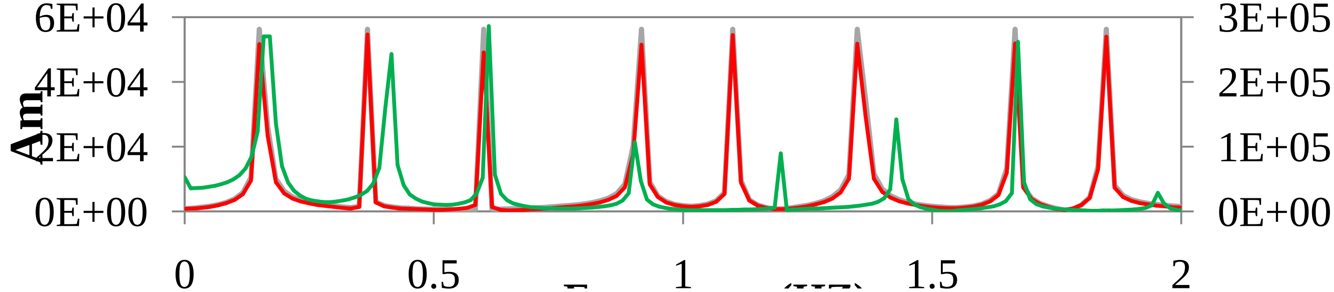
Different defects, different signatures!!

$$E = \sum (h_i^m - h_i)^2 = \text{minimum}$$

$$\frac{\partial H}{\partial t} + \frac{Q}{A}\frac{\partial H}{\partial x} + \frac{a^2}{gA}\frac{\partial Q}{\partial x} + \frac{a^2}{gA} Q_L \delta(x - x_L) = 0$$

$$\frac{\partial}{\partial t}(\rho A V) + \frac{\partial}{\partial x}(\rho A V^2) + A\frac{\partial p}{\partial x} + \frac{\rho A f V^2}{2D} = 0$$

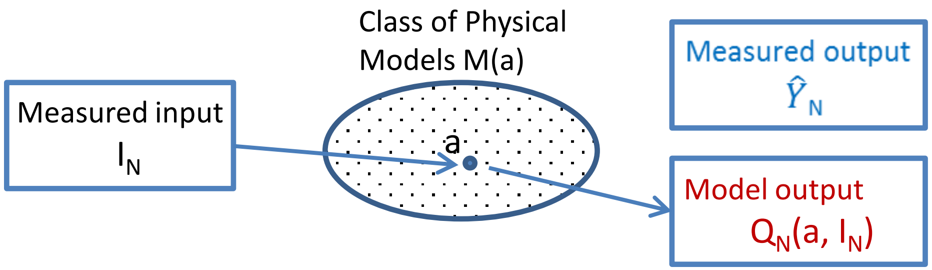


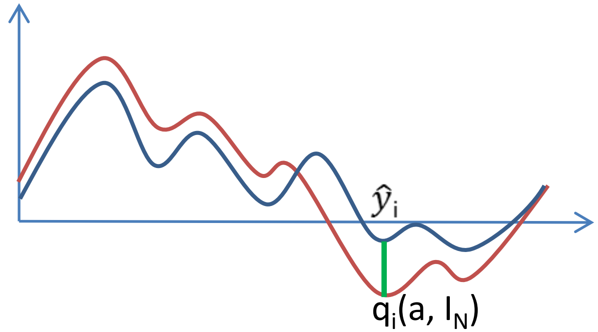


1. Blockages induce frequency shift
2. Leaks induce damped pattern, but no shift
3. Wall thinning; changes wave speed
4. Air blockages; changes wave speed and damp pressure

Model Updating Algorithms

❑ Deterministic model updating

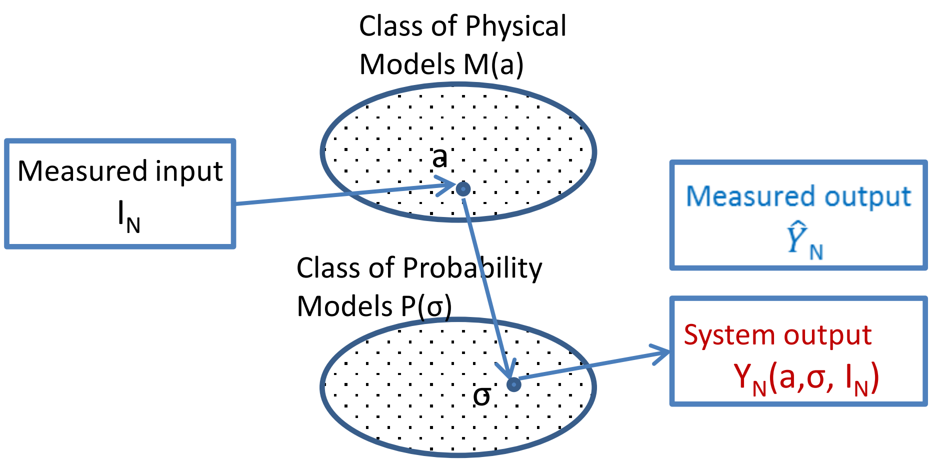


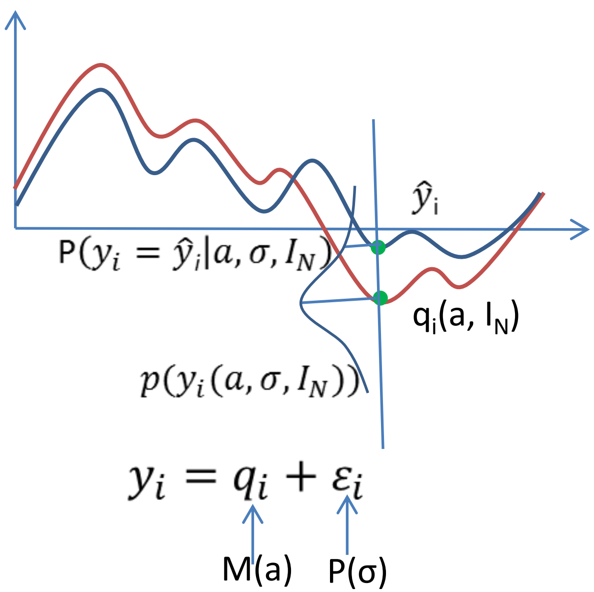


Least squares: $\arg\min_a \sum (\hat{y}_i - q_i(a, I_N))^2$

- Does not allow for explicit treatment of uncertainties (modeling error, parameter uncertainties)
- Unable to handle and interpret locally identifiable and unidentifiable cases

❑ Probabilistic model updating





BAYES' THEOREM

$$p(\mathbf{a},\sigma \mid \mathcal{D}_N) = c\, p(\hat{\mathbf{Y}}_N \mid \mathbf{a},\sigma)\, \pi(\mathbf{a},\sigma)$$

$$p(\hat{Y}_N \mid \mathbf{a},\sigma) = \frac{1}{\left(\sqrt{2\pi}\sigma\right)^{NN_0}} \exp\left[-\frac{NN_0}{2\sigma^2} J(\mathbf{a})\right]$$

$$J(\mathbf{a}) = \frac{1}{NN_0} \sum_{n=1}^{N} \left\| \hat{\mathbf{y}}(n) - S_o \mathbf{q}(n;\mathbf{a}) \right\|^2$$

$$\boxed{p(\mathbf{a} \mid \mathcal{D}_N) = c_1 J(\mathbf{a})^{-N_J}\, \pi(\mathbf{a},\hat{\sigma}(\mathbf{a}))}$$

$$\hat{\sigma}^2(\mathbf{a}) = J(\mathbf{a}),\quad N_J = \frac{(NN_0 - 1)}{2}$$

where:

$\mathbf{a}$	:	structural parameters
σ	:	output-error parameter
$\hat{\mathbf{Y}}_\mathbf{N}$	:	$\{\hat{\mathbf{y}}(1),\cdots\hat{\mathbf{y}}(N)\}$
$\hat{\mathbf{y}}(n)$	:	observed response at $t = n\Delta t$
$\mathbf{q}(n;\mathbf{a})$	:	model response at $t = n\Delta t$
N	:	number of observed response vectors
N_0	:	number of observed DOF – dimension of $\mathbf{y}$
$\mathcal{D}_N$	:	$\{\hat{\mathbf{Y}}_N, I_N\}$ measured dynamic data (output and input)

OPTIMAL PARAMETER(S) $\hat{\mathbf{a}}$

The parameter(s) which maximize $p(\mathbf{a} | \mathcal{D}_N)$ or (assuming slowly varying prior $\pi(\mathbf{a},\sigma)$) the parameters that minimize $J(\mathbf{a})$, that is:

$$J(\hat{\mathbf{a}} | \mathcal{D}_N) = \min_{\mathbf{a} \in S(\mathbf{a})} J(\mathbf{a})$$

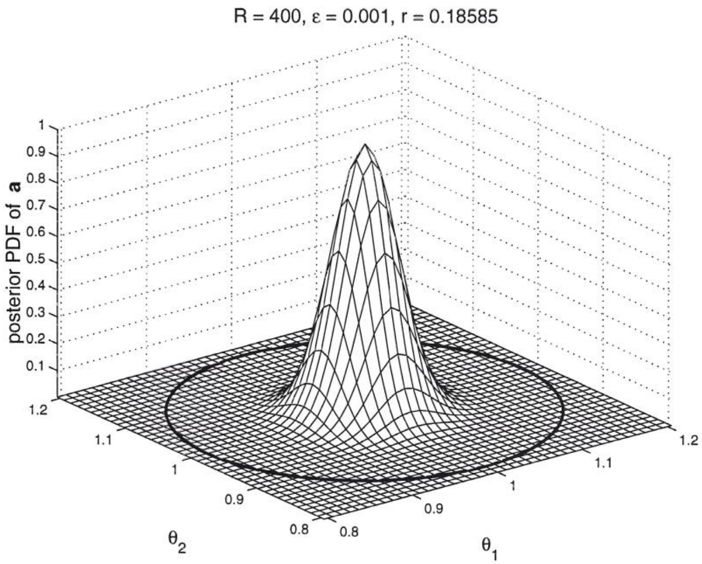
Note:
Only Models $M(\mathbf{a})$ with $J(\mathbf{a})$ equal (or "very close") to $J(\hat{\mathbf{a}})$ have significant probability.

"Very close" means $J(\mathbf{a}) \leq J(\hat{\mathbf{a}}) \varepsilon^{-\frac{1}{N_J}}$

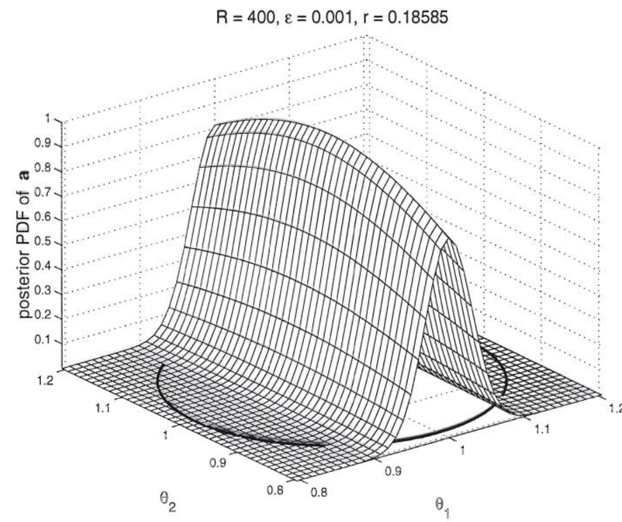
where ε is a threshold of "significant" relative PDF values [e.g. $\varepsilon = 10^{-3}$, $N_J = 1000$, $J(\mathbf{a}) \leq 1.007 J(\hat{\mathbf{a}})$]

Thus, the updated PDF is concentrated in a very small sub-region of the parameter space.

IDENTIFIABILITY (DEFINITION)



$p(\mathbf{a} \mid \mathcal{D}_N, \mathcal{M}_P)$ with $\lambda_1(A_N(\hat{\mathbf{a}})) = \lambda_2(A_N(\hat{\mathbf{a}})) = 1000$



$p(\mathbf{a} \mid \mathcal{D}_N, \mathcal{M}_P)$ with $\lambda_1(A_N(\hat{\mathbf{a}})) = 1000$ and $\lambda_2(A_N(\hat{\mathbf{a}})) = 10$

IDENTIFIABILITY (DEFINITIONS)

IDENTIFIABILITY (Weak Definition)

There exists a finite number of optimal parameters
$\hat{\mathbf{a}}^{(k)},\ k = 1, \ldots, K$

IDENTIFIABILITY OF ORDER R (Strong Definition)

There exists a finite number of optimal parameters
$\hat{\mathbf{a}}^{(k)},\ k = 1, \ldots, K$
and
In addition, $p(\mathbf{a} | \mathcal{D}_N)$ decays "rapidly" in all directions in the neighborhood of each optimal point, as one moves away from that point.

"Rapidly" means: $$\boxed{\min \lambda_i(A(\hat{\mathbf{a}})) > R \geq 0} \qquad (*)$$

$$A(\mathbf{a}) = \nabla^2 [N_J \ln J(\mathbf{a})]$$

<u>Note</u>:

- Equation $(*)$ implies that if one moves away from any optimal point a distance x in any direction, the PDF will decay faster than $\exp(-Rx^2/2)$.
- It also implies that the region of "significant" probabilities (relative PDF larger than a chosen threshold $0 < \varepsilon < 1$) consists of neighborhoods of the optimal points contained within spheres of radius $r = \sqrt{-2 \ln \varepsilon / R}$.

ASYMPTOTIC APPROXIMATIONS FOR IDENTIFIABLE CASES

The following approximations are valid for a strongly identifiable case.

UPDATED PDF

$$p(\mathbf{a} \mid \mathcal{D}_N) \approx \sum_{k=1}^{K} w_k N\left(\mathbf{a}; \hat{\mathbf{a}}^{(k)}, A^{-1}\left(\hat{\mathbf{a}}^{(k)}\right)\right)$$

$A(\mathbf{a}) = \nabla^2[N_J \ln J(\mathbf{a})]$ (assume slowly varying $\pi(\mathbf{a}, \sigma)$)

$$w_k = \pi\left(\hat{\mathbf{a}}^{(k)}\right)\left|A_N\left(\hat{\mathbf{a}}^{(k)}\right)\right|^{-\frac{1}{2}}$$

PREDICTIVE RESPONSE

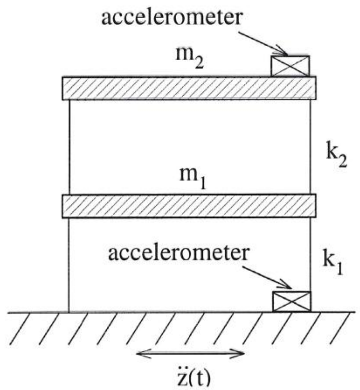
$$p\left(\mathbf{Y}_{N+1,M} \mid \mathcal{D}_N\right) \approx \sum_{k=1}^{K} w_k \, p\left(\mathbf{Y}_{N+1,M} \mid \hat{\mathbf{a}}^{(k)}\right)$$

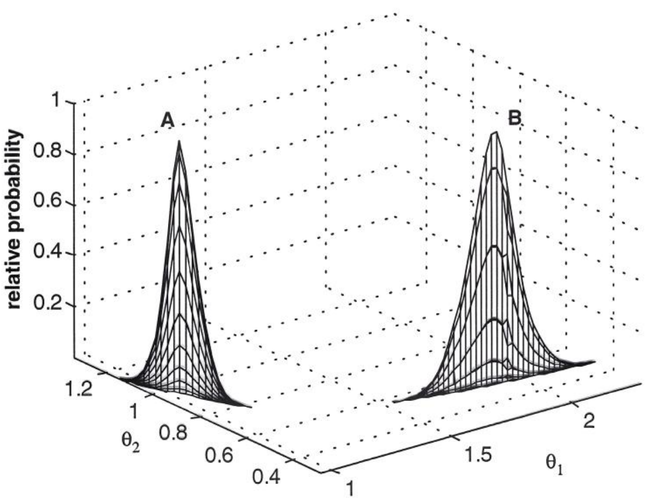
Remark:

In an identifiable case all optimal models are output-equivalent. An algorithm for finding the set of all optimal solutions, given one of them, can be found in [Katafygiotis and Beck 1997]

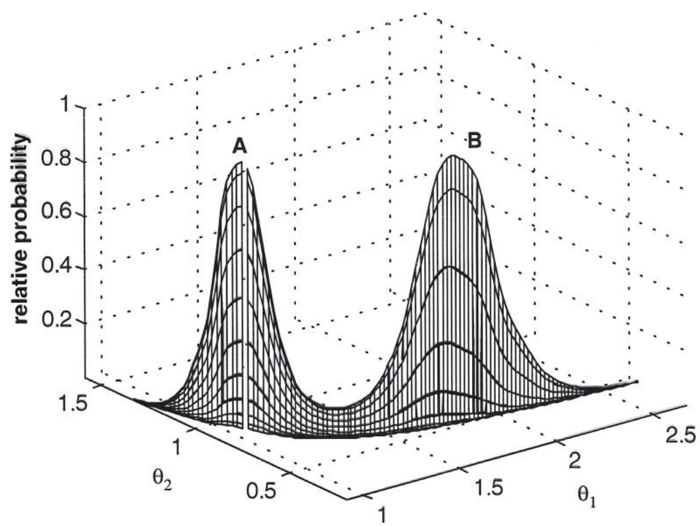
Katafygiotis L. S. and Beck J. L., "Updating models and their Unertainties. II: Model Identifiability", Journal of Engineering Mechanics, ASCE, 124(4), 463-467, 1998.

EXAMPLE: TWO-STORY SHEAR BUILDING





20% noise



40% noise

NON-IDENTIFIABLE CASE - MANIFOLD

NON-IDENTIFIABILITY

There exist a finite or infinite number of optimal solutions. Given an optimal solution $\hat{\mathbf{a}}$, the PDF $p(\mathbf{a} | \mathcal{D}_N)$ does not decay "rapidly" in all directions in the neighborhood of $\hat{\mathbf{a}}$. The PDF is concentrated in the neighborhood of a manifold of dimension $0 < N_G < N_a$.

The dimension N_G of the manifold is determined by the number of eigenvalues satisfying:

$$\lambda_i(A(\hat{\mathbf{a}})) \leq R \quad i = 1, \ldots, N_G < N_a$$

The corresponding eigenvectors span a hyperplane tangential to the manifold at $\hat{\mathbf{a}}$.

Note:

1) The identifiable case is a special case, where $N_G = 0$.
2) The manifold can be connected or disconnected.
3) Optimization algorithms will usually yield just one point, not necessarily optimal, on the manifold.

CHALLENGES

- Representation of manifold.
- Asymptotic approximations for updated PDF and predictive response.

NON-IDENTIFIABLE CASE - MANIFOLD

Representation of Manifold

Represent the manifold by a finite set of points $\{\mathbf{a}^{(l)},\ l=1,\ldots,L\}$ located on it. The set of points should:

1) represent the entire manifold, not only part of it.
2) be almost uniformly distributed.

Each point $\mathbf{a}^{(l)}$ is assigned a weighting w_l representing the volume of the PDF in its neighborhood.

$$w_l = c_2 \cdot J\left(\mathbf{a}^{(l)}\right)^{-N_J} \pi\left(\mathbf{a}^{(l)}, \hat{\sigma}\left(\mathbf{a}^{(l)}\right)\right) \cdot \left[\lambda_{N_G+1}^{(l)} \cdot \lambda_{N_G+2}^{(l)} \cdots \lambda_{N_a}^{(l)}\right]^{-\frac{1}{2}} \cdot I\left(\mathbf{a}^{(l)}\right)$$

where:

c_2 : normalizing constant so that $\sum_{l=1}^{N_l} w_l = 1$

$\lambda_{N_G+1}^{(l)}, \lambda_{N_G+2}^{(l)}, \cdots, \lambda_{N_a}^{(l)}$: the (N_a-N_G) largest eigenvalues of $A\left(\mathbf{a}^{(l)}\right)$

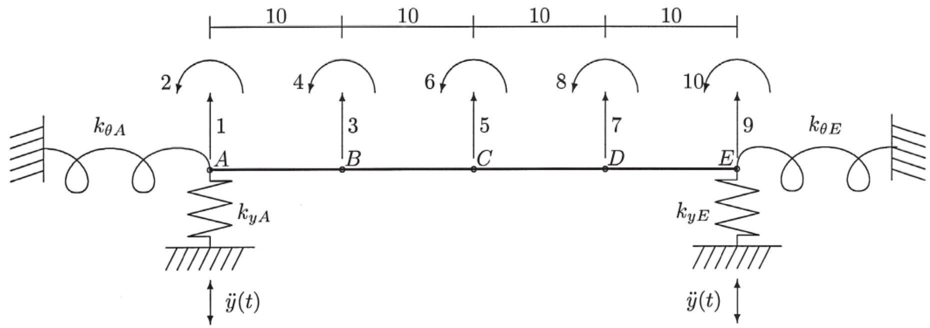
$I\left(\mathbf{a}^{(l)}\right)$: coefficient proportional to the tributary area of the manifold corresponding to $\mathbf{a}^{(l)}$ [e.g. $I\left(\mathbf{a}^{(l)}\right)=1$ for uniform spacing.]

Asymptotic Approximations

$$p\left(\mathbf{a} \mid \mathrm{D}_N\right) = \sum_{l=1}^{L} w_l \delta\left(\mathbf{a} - \mathbf{a}^{(l)}\right)$$

$$p\left(\mathbf{Y}_{N+1,M} \mid \mathrm{D}_N\right) = \sum_{l=1}^{L} w_l\, p\left(\mathbf{Y}_{N+1,M} \mid \mathbf{a}^{(l)}\right)$$

Example: Elastically supported bridge



	k_{yA} 10^7	k_{yB} 10^7	$k_{\theta A}$ 10^5	$k_{\theta B}$ 10^5	EI_{AB} 10^6	EI_{BC} 10^6	EI_{CD} 10^6	EI_{DE} 10^6
nominal undamaged	1.10	1.20	0.90	0.85	0.95	1.05	0.90	0.95
nominal damaged	1.10	1.20	0.90	0.85	0.76	1.05	0.90	0.95
model 1	θ_1	θ_1	θ_2	θ_2	θ_3	θ_3	θ_3	θ_3
model 2	θ_1	θ_2	θ_3	θ_3	θ_4	θ_4	θ_4	θ_4
model 3	θ_1	θ_1	θ_2	θ_2	θ_3	θ_3	θ_4	θ_4
model 4	θ_1	θ_2	θ_3	θ_4	θ_5	θ_5	θ_5	θ_5

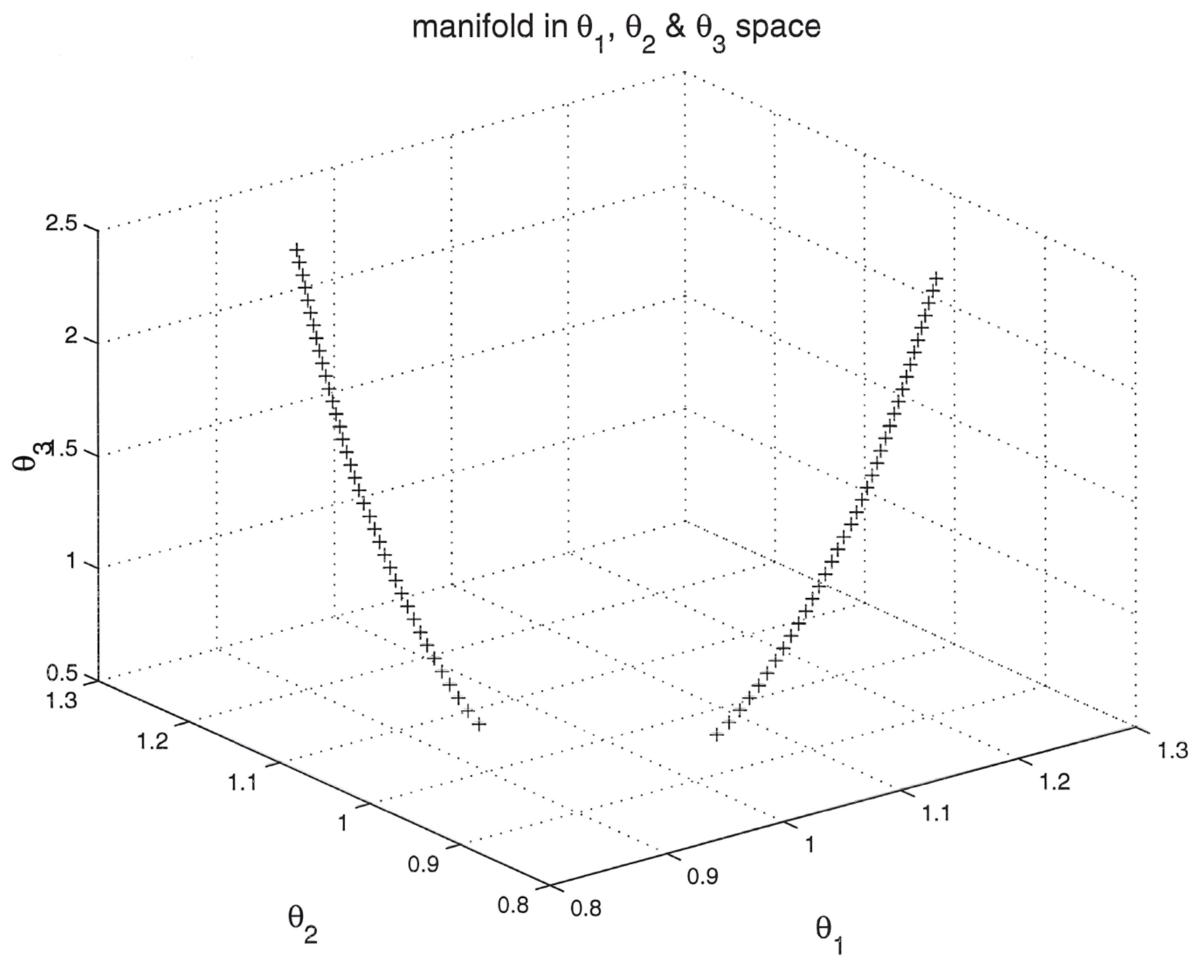
Note:

model 1: identifiable

model 2 & 3: one-dimensional manifold

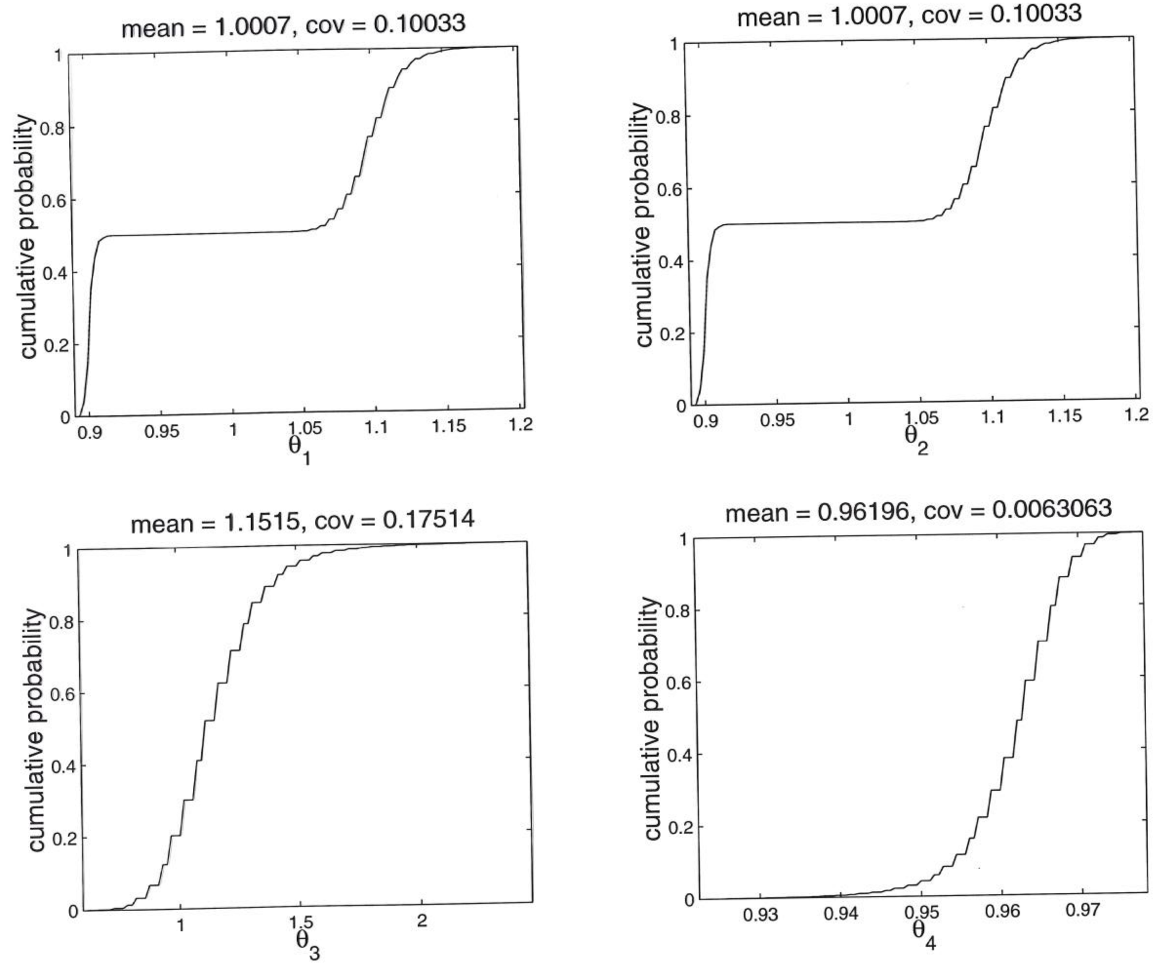
model 4: two-dimensional manifold

MANIFOLD

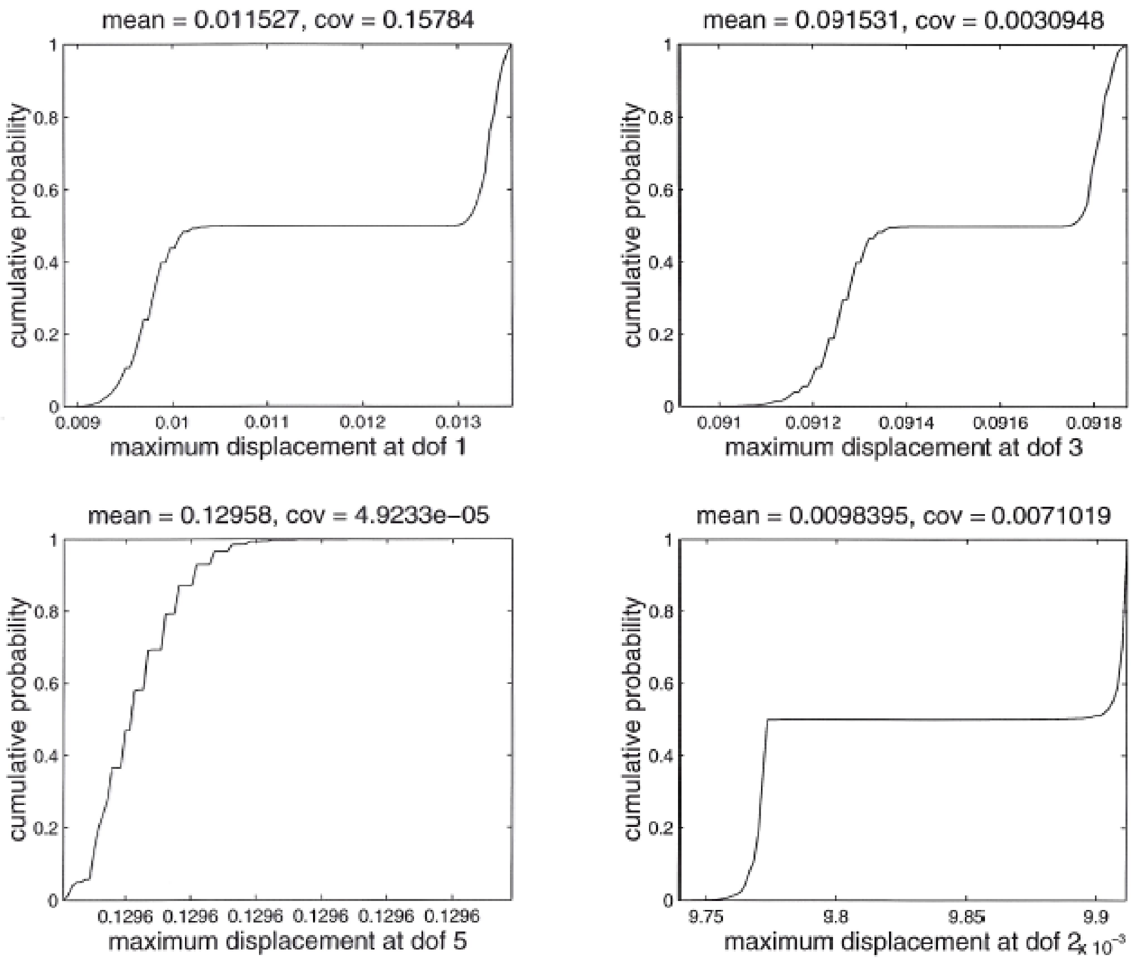


manifold in θ1, θ2 & θ3 space
2.5
2
1.5
1
0.5
θ3
1.3
1.2
1.1
1
0.9
0.8
θ2
0.8
0.9
1
1.1
1.2
1.3
θ1

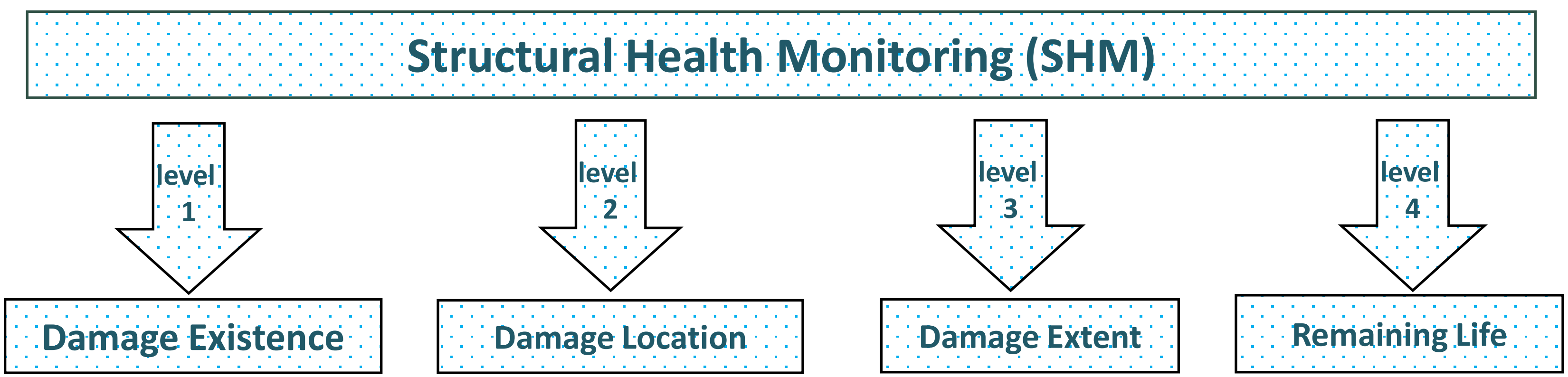
Cumulative probability for θs



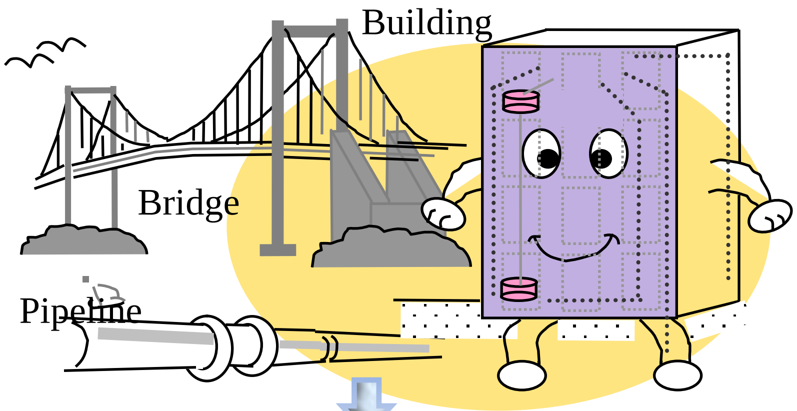
Cumulative probability for maximum responses



Structural Health Monitoring



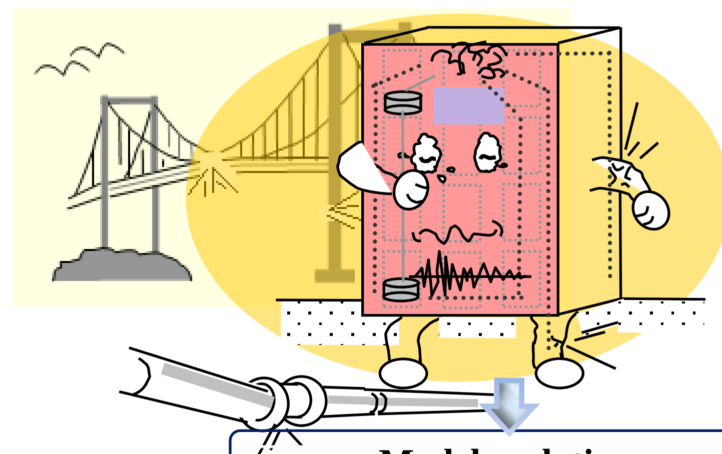
Baseline Condition



Model updating

PDF of stiffness parameters

Damage Condition



Model updating

PDF of stiffness parameters

Based on the Gaussian approximation of stiffness scaling factors in two states(θ_i^{ud} and θ_i^{pd}), the probability of damage (Vanik 1997) is

$$P_i^{dam}(d) \approx \Phi\left(\frac{(1-d)\hat{\theta}_i^{ud} - \hat{\theta}_i^{pd}}{\sqrt{(1-d)^2\left(\hat{\sigma}_i^{ud}\right)^2 + \left(\hat{\sigma}_i^{pd}\right)^2}}\right)$$

Risk Analysis, Reliability Analysis, and Decision Making

ELSEVIER

Probabilistic Engineering Mechanics 18 (2003) 315–327

PROBABILISTIC
ENGINEERING
MECHANICS

www.elsevier.com/locate/probengmech

Leakage detection in water pipe networks using a Bayesian probabilistic framework

Z. Poulakis, D. Valougeorgis, C. Papadimitriou*

Department of Mechanical and Industrial Engineering, University of Thessaly, Pedion Areos, Volos 38334, Greece

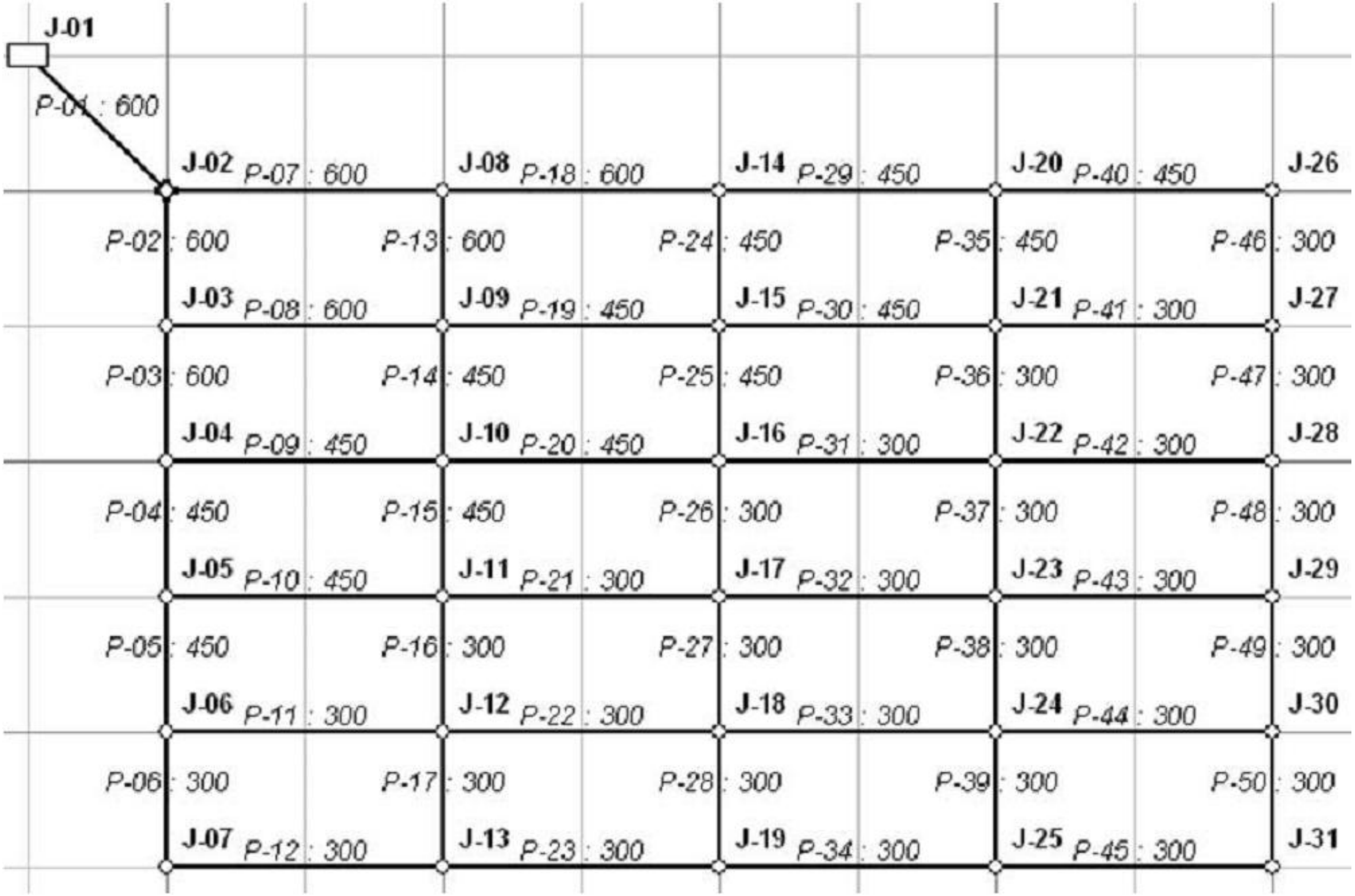
Received 17 November 2002; revised 1 July 2003; accepted 1 July 2003

Abstract

A Bayesian system identification methodology is proposed for leakage detection in water pipe networks. The methodology properly handles the unavoidable uncertainties in measurement and modeling errors. Based on information from flow test data, it provides estimates of the most probable leakage events (magnitude and location of leakage) and the uncertainties in such estimates. The effectiveness of the proposed framework is illustrated by applying the leakage detection approach to a specific water pipe network. Several important issues are addressed, including the role of modeling error, measurement noise, leakage severity and sensor configuration (location and type of sensors) on the reliability of the leakage detection methodology. The present algorithm may be incorporated into an integrated maintenance network strategy plan based on computer-aided decision-making tools.

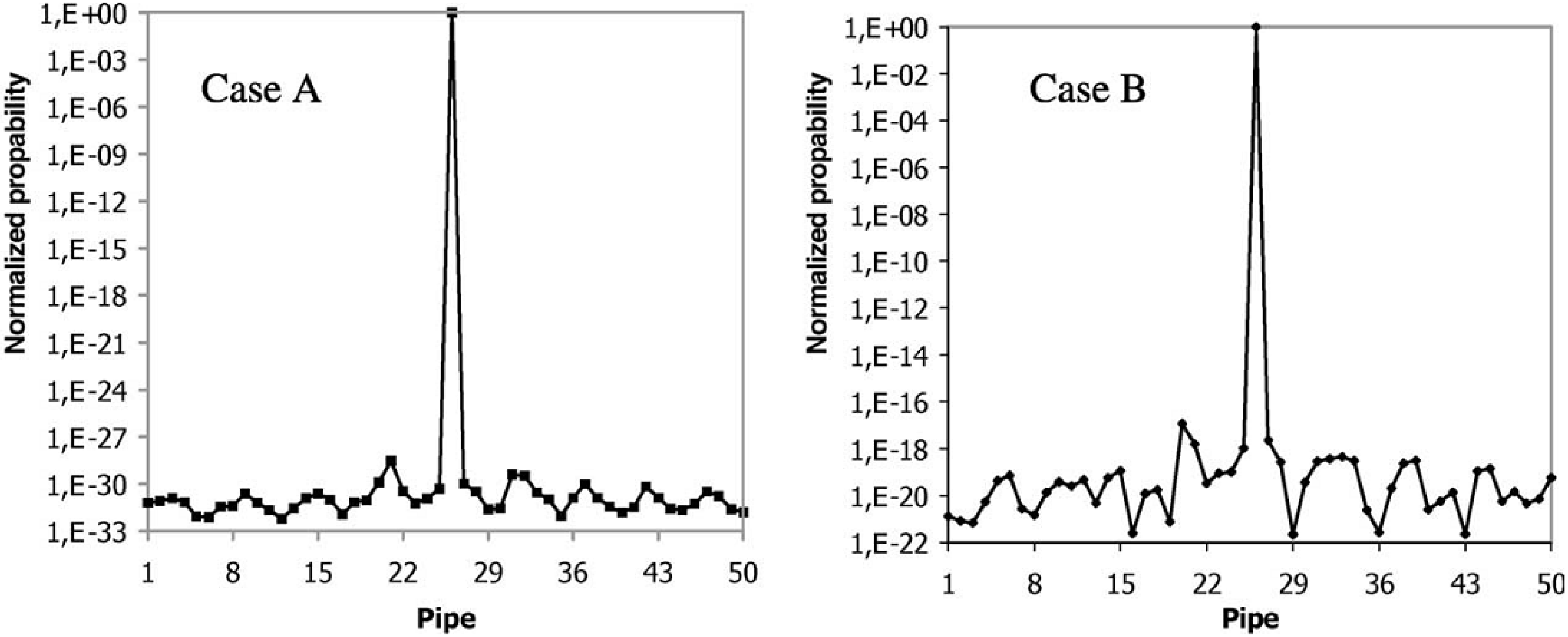
Keywords: System identification; Bayesian method; Leakage detection; Water pipe networks

Poulakis, Z., Valougeorgis, D., & Papadimitriou, C. (2003). Leakage detection in water pipe networks using a Bayesian probabilistic framework. *Probabilistic Engineering Mechanics*, *18*(4), 315-327.

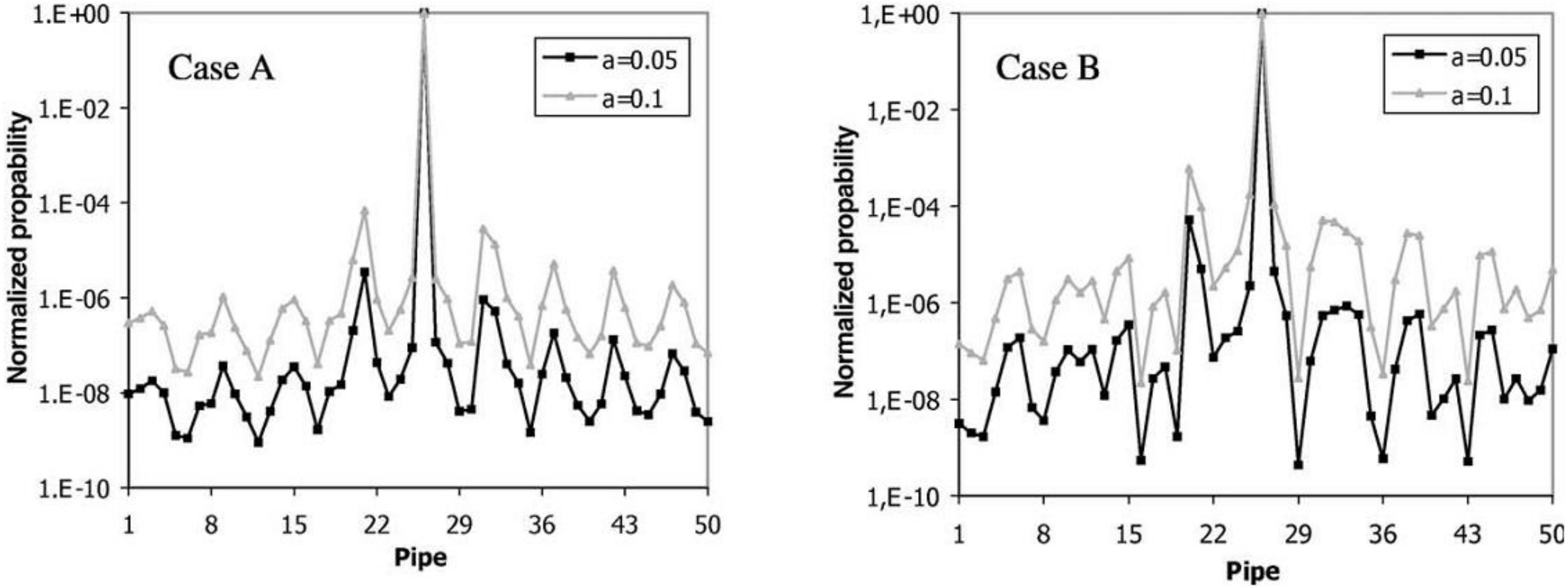


Water pipe network configuration

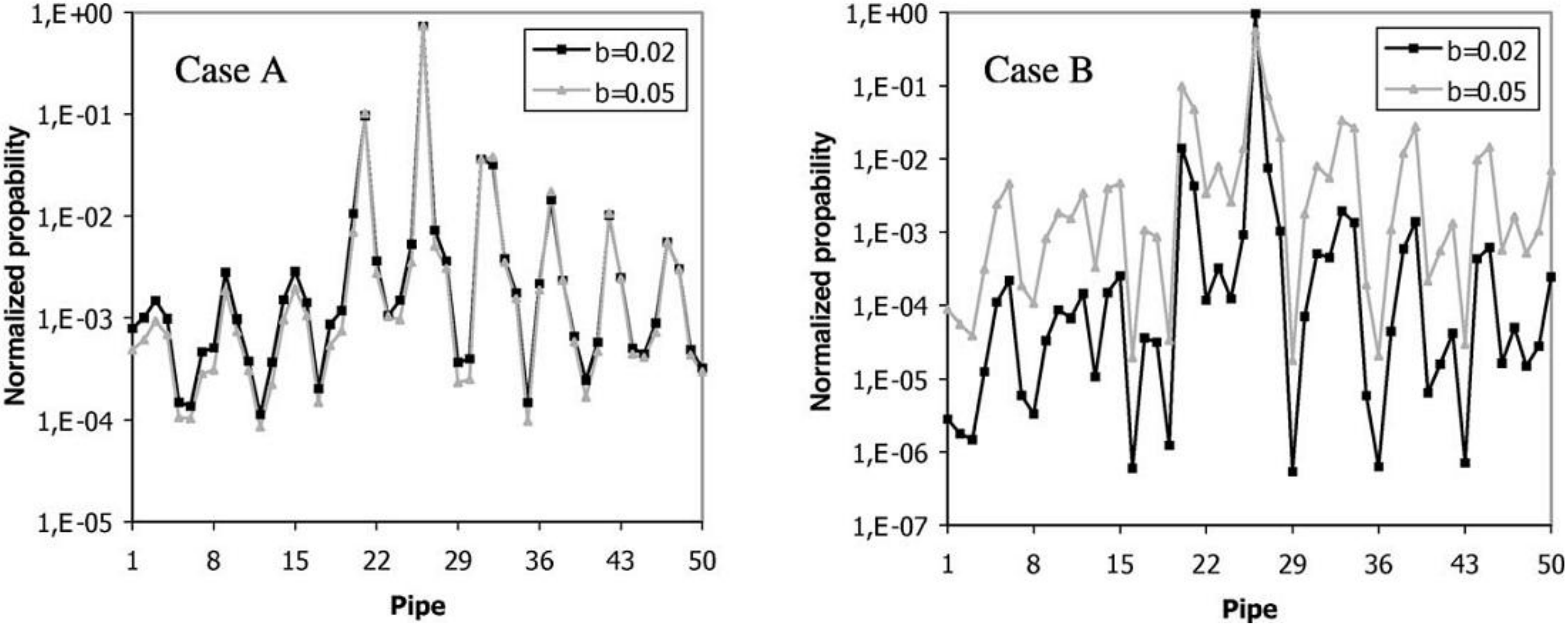
- Particular example involves a mixture of discrete and continuous optimizations. Discrete parameters may grow rapidly. Discrete optimizations can be treated using genetic optimization algorithms.



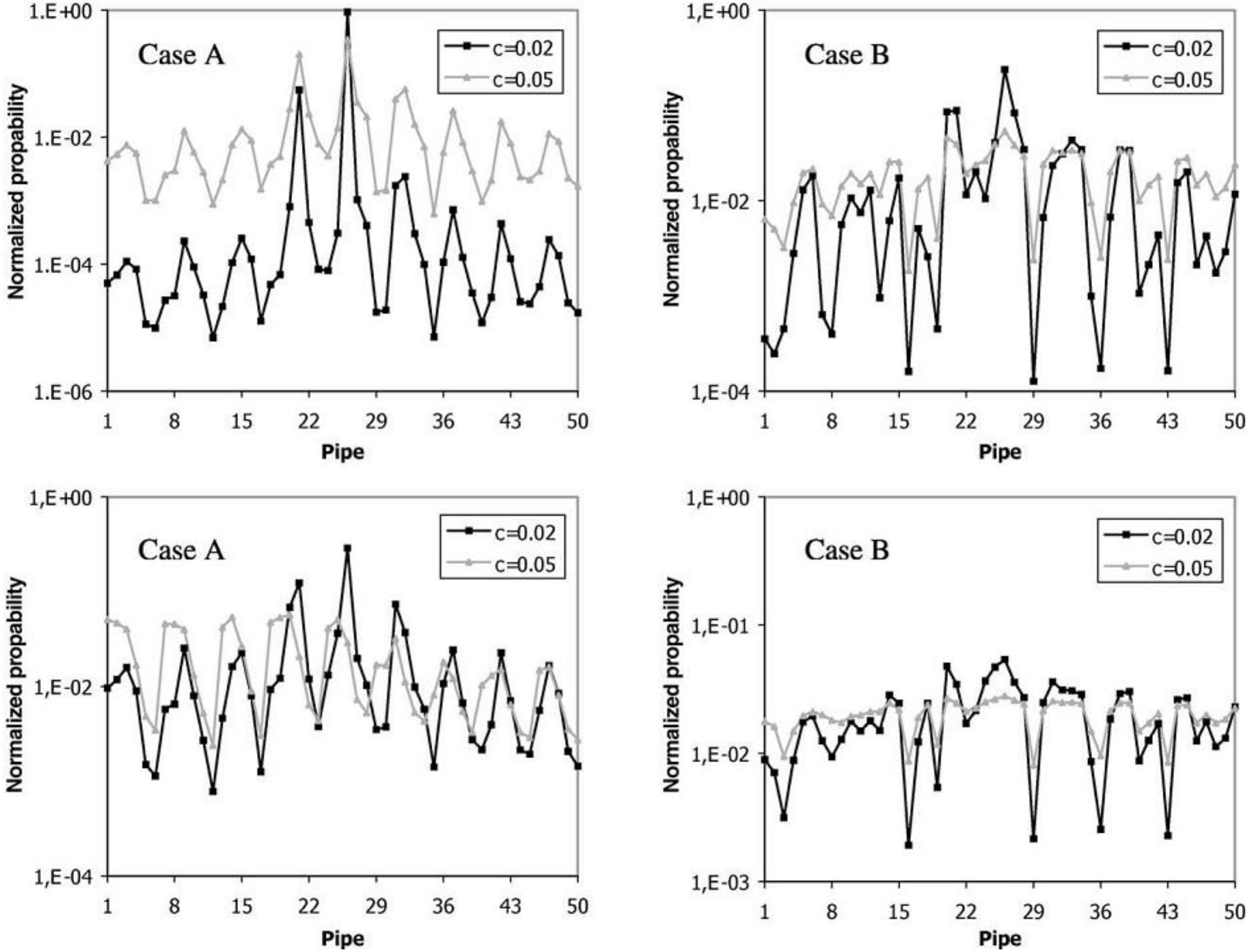
Peak values of normalized PDF at each pipe section using (A) manometers and (B) flow meters. Leakage is located at pipe 26 with severity equal to 22.8 l/s (1.5% of the total water volume).



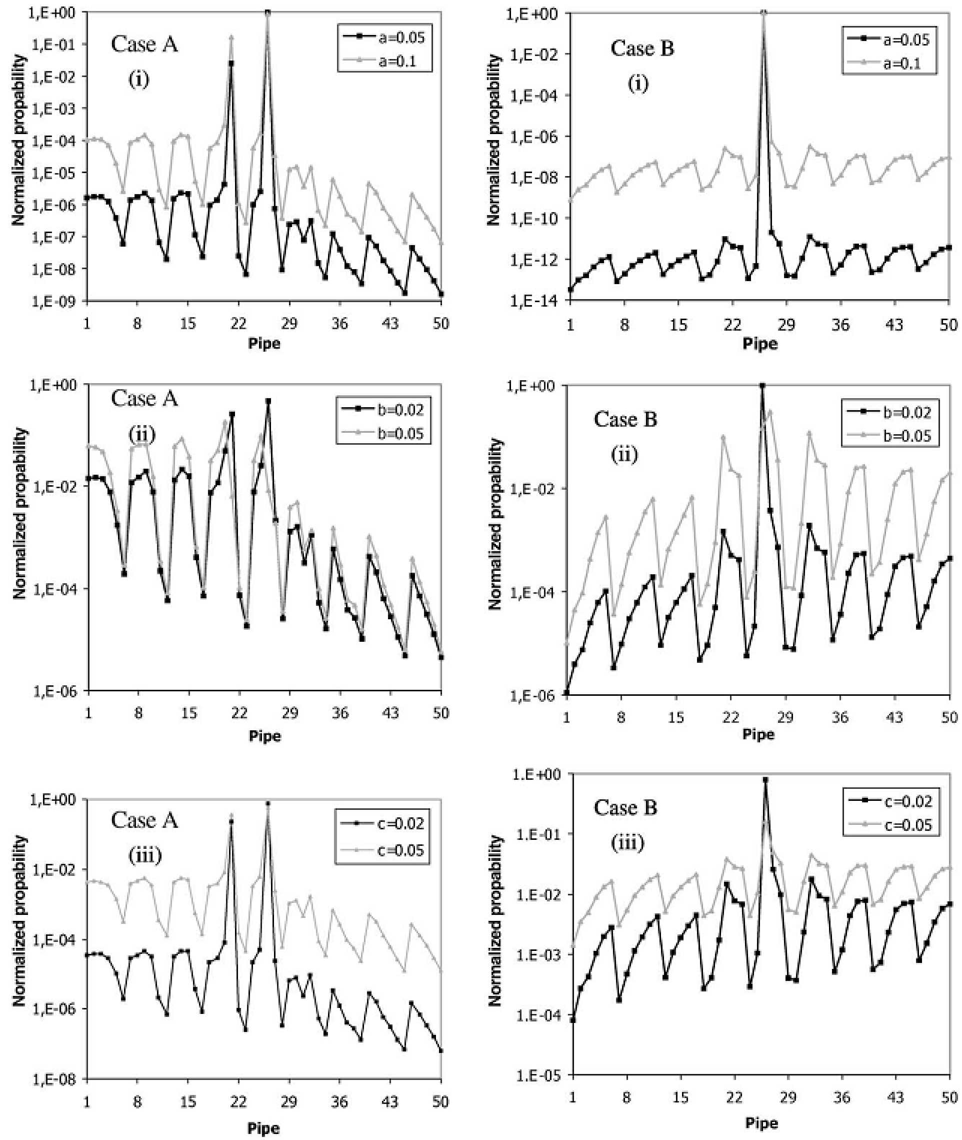
Peak values of normalized PDF at each pipe section using (A) manometers and (B) flow meters. Leakage is located at pipe 26 with severity equal to 22.8 l/s (1.5% of the total water volume). A perturbation a ¼ 5 and 10% is assumed in the piping roughness coefficient.



Peak values of normalized PDF at each pipe section using (A) manometers and (B) flow meters. Leakage is located at pipe 26 with severity equal to 22.8 l/s (1.5% of the total water volume). A perturbation b = 2 and 5% is assumed in the demands.



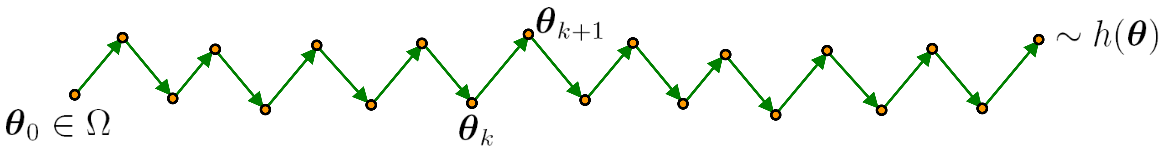
Peak values of normalized PDF at each pipe section using (A) manometers and (B) flow meters. Leakage is located at pipe 26 with severity equal to (i) 57.0 and (ii) 22.8 l/s (3.7 and 1.5% of the total water volume). A perturbation c = 2 and 5% is assumed in the modeled measurements.



Peak values of normalized PDF at each pipe section using (A) manometers in the nodes (17, 18, 19, 23, 24, 25, 31) and (B) flow meters in the pipe sections (1, 2, 3, 7, 18, 25, 26). Leakage is located in pipe 26 with severity equal to 57.0 l/s.

SIMULATION BASED APPROACHES

Metropolis-Hastings algorithm



$\boldsymbol{\theta}_k \rightsquigarrow \boldsymbol{\theta}_{k+1}$ with symmetric "proposal PDF" $p^*(\boldsymbol{\xi}|\boldsymbol{\theta})$

1. *Generate candidate state* $\tilde{\boldsymbol{\theta}}$

Simulate $\boldsymbol{\xi}$ according to $p^*(\boldsymbol{\xi}|\boldsymbol{\theta}_k)$ and compute the ratio $r = h(\boldsymbol{\xi})/h(\boldsymbol{\theta}_k)$

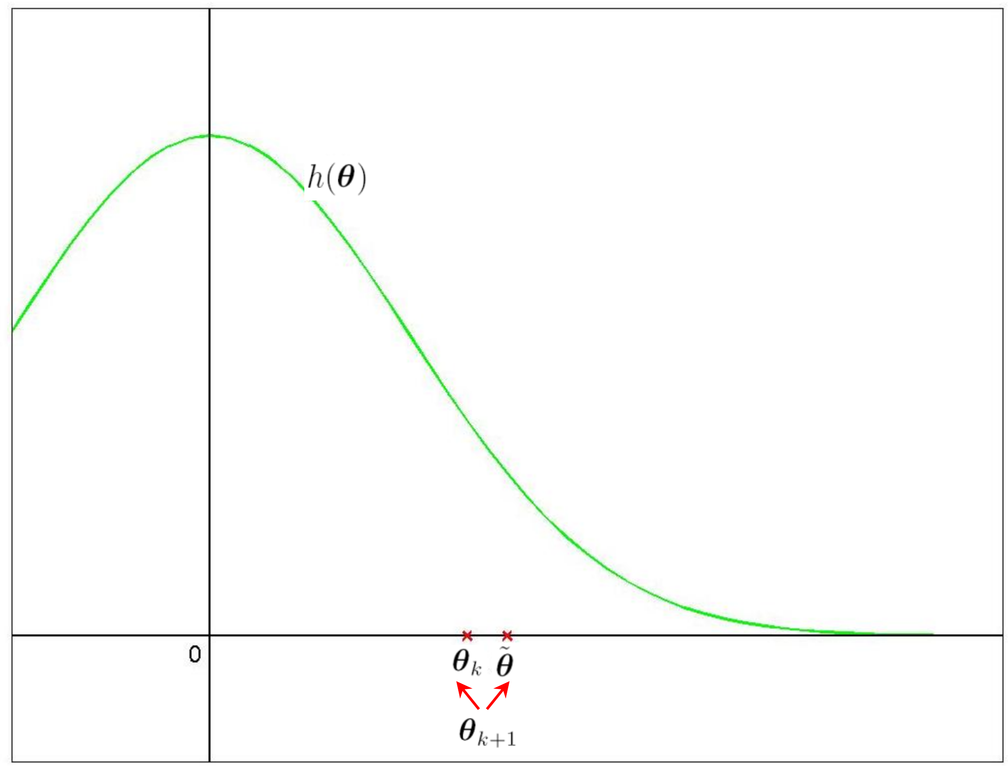
Set $\tilde{\boldsymbol{\theta}} = \boldsymbol{\xi}$ if $r \geq 1 (c(\boldsymbol{\xi}) \leq c(\boldsymbol{\theta}_k))$

otherwise,

$$\text{Set } \tilde{\boldsymbol{\theta}} = \begin{cases} \boldsymbol{\xi} & \text{with } r \\ \boldsymbol{\theta}_k & \text{with } 1 - r \end{cases}$$

2. *Accept/Reject* $\tilde{\boldsymbol{\theta}}$

$$\boldsymbol{\theta}_{k+1} = \begin{cases} \tilde{\boldsymbol{\theta}} & \text{if } \tilde{\boldsymbol{\theta}} \in \Omega \\ \boldsymbol{\theta}_k & \text{otherwise} \end{cases}$$

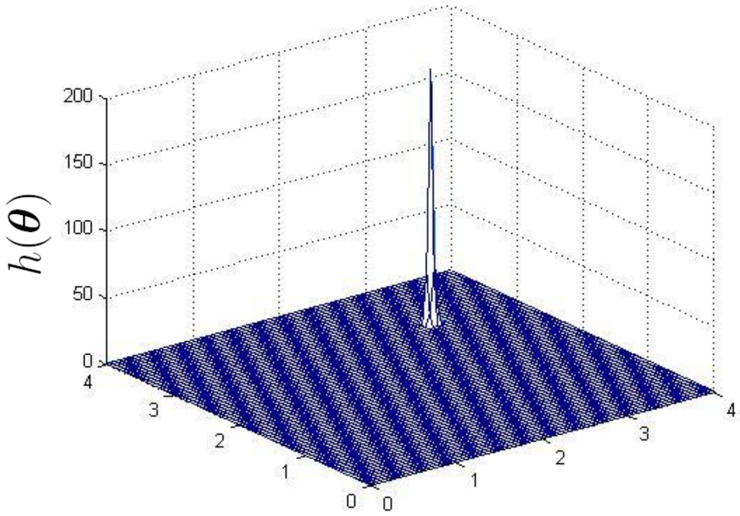


Metropolis-Hastings algorithm

Difficult to choose the spread of the $p^*(\boldsymbol{\xi}|\boldsymbol{\theta})$ to ensure:

1. The acceptance rate is not too small

2. The concentration volume can be effectively explored



Not suitable for populating the distribution concentrated on a small volume

Transitional Markov chain Monte Carlo (Ching & Chen 2007)

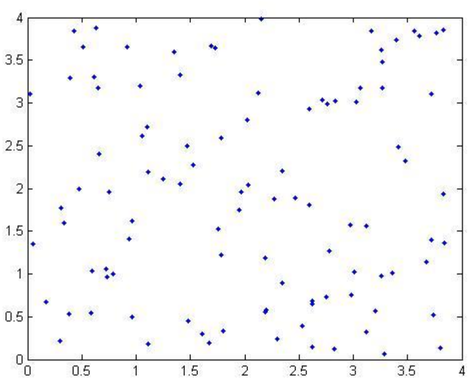
Introduce a series of intermediate PDFs
in the feasible design space:

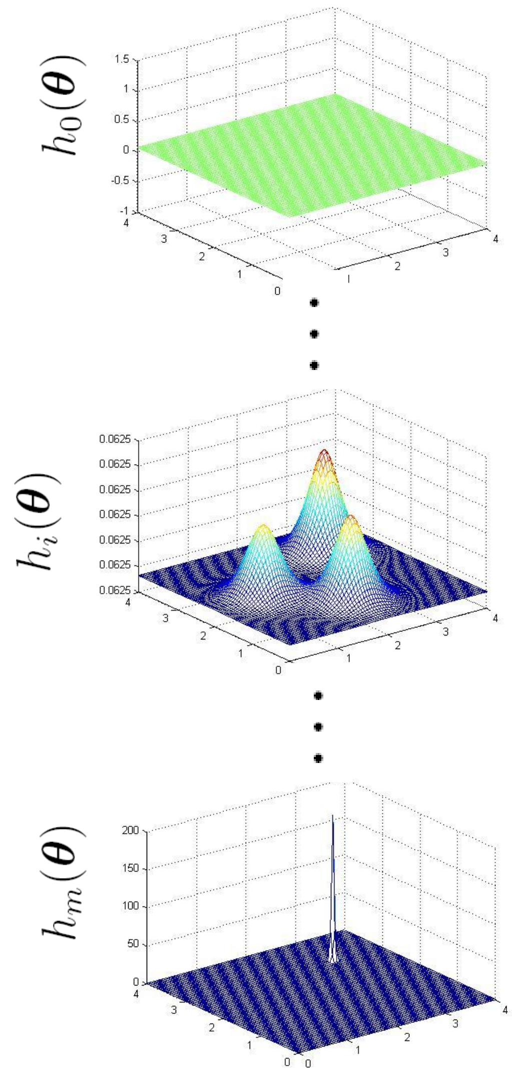
$$h_j(\boldsymbol{\theta}) = [\,p(D\,|\,\boldsymbol{\theta}, M)]^{1/T_j}\,\pi(\boldsymbol{\theta}\,|\,M), \qquad j = 0,\ldots,m$$

$$\infty = T_0 > T_1 > \ldots > T_m = 1$$

Start with samples distributed according to $\pi(\boldsymbol{\theta}\,|\,M)$

$$\left\{\boldsymbol{\theta}_1^{(0)}, \ldots, \boldsymbol{\theta}_N^{(0)}\right\}$$





Transitional Markov chain Monte Carlo

$$\left\{\boldsymbol{\theta}_1^{(i)},\ldots,\boldsymbol{\theta}_N^{(i)}\right\} \sim h_i(\boldsymbol{\theta})$$

1. Determine T_{i+1} for $h_{i+1}(\boldsymbol{\theta}) = h(\boldsymbol{\theta}; T_{i+1})$,

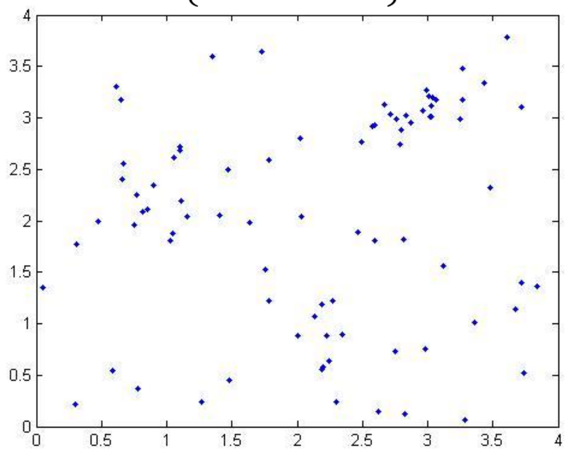
$$\mathrm{COV}\left\{ w\left(\boldsymbol{\theta}_k^{(i)}\right) = \frac{h_{i+1}\left(\boldsymbol{\theta}_k^{(i)}\right)}{h_i\left(\boldsymbol{\theta}_k^{(i)}\right)}, k = 1,\ldots,N \right\} = \mathrm{COV}_t$$

2. $\left\{\boldsymbol{\theta}_1^{(i)},\ldots,\boldsymbol{\theta}_N^{(i)}\right\} \sim \; p(k) = \dfrac{w\left(\boldsymbol{\theta}_k^{(i)}\right)}{\sum_{l=1}^{N} w\left(\boldsymbol{\theta}_l^{(i)}\right)}$

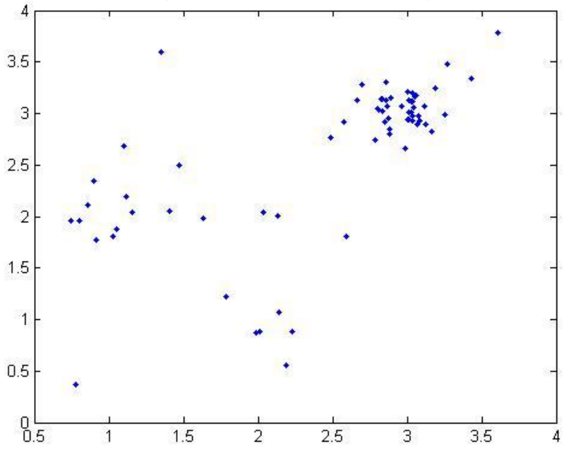
Use Markov chain Monte Carlo for samples selected repeatedly.

$$\rightsquigarrow \left\{\boldsymbol{\theta}_1^{(i+1)},\ldots,\boldsymbol{\theta}_N^{(i+1)}\right\} \sim h_{i+1}(\boldsymbol{\theta})$$

$\left\{\boldsymbol{\theta}_1^{(i)},\ldots,\boldsymbol{\theta}_N^{(i)}\right\}$



$\left\{\boldsymbol{\theta}_1^{(i+1)},\ldots,\boldsymbol{\theta}_N^{(i+1)}\right\}$



Concluding Remarks

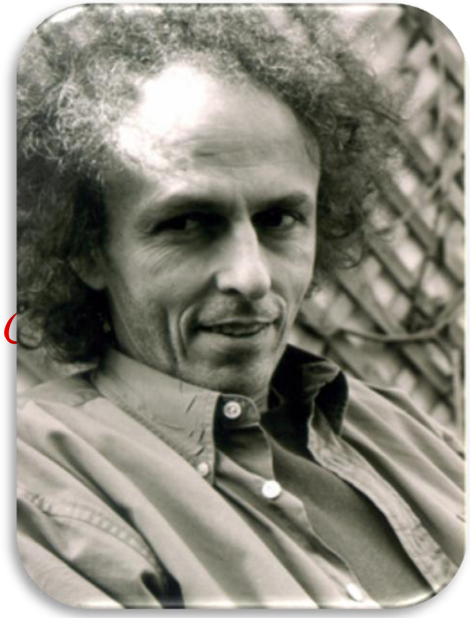
- A Bayesian Probabilistic Framework for Model Updating has been presented
- This framework allows for the explicit treatment of modeling errors, measurement noise, and non-uniqueness in the inverse problem
- Identifiability depends on prior information, the fidelity of the model class, the number of model parameters to be updated, as well as the amount and quality of measured data
- As a result, probability distributions of the updated model parameters are obtained. Shifts of such distributions can be used to infer damage
- Significant modeling error may pollute the results of the methodology; estimated severity of damage as well as location of damage may become unreliable
- The importance of good modeling (appropriate class of models and parameters to be updated) cannot be overemphasized
- Application-specific algorithms, asymptotic or simulation-based, need to be designed to make most efficient use of the particular data on hand
- The methodology can be extended to select of an optimal class of models among different such classes
- It can also be used to design an optimal sensor layout by minimizing the asymptotic estimate of the information entropy

Uncertainty is

the only certainty there is,

and knowing how to live with insec

is the only security.



----John Allen Paulos (Mathematics writer)