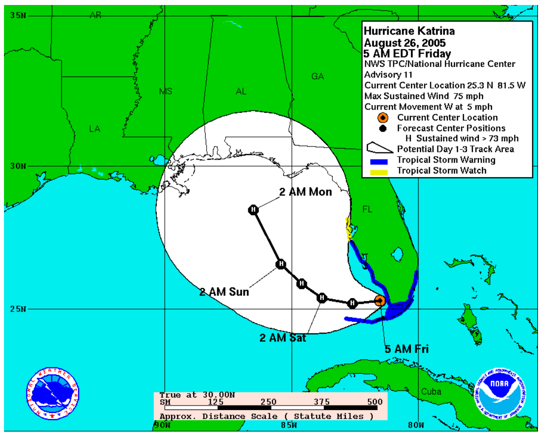
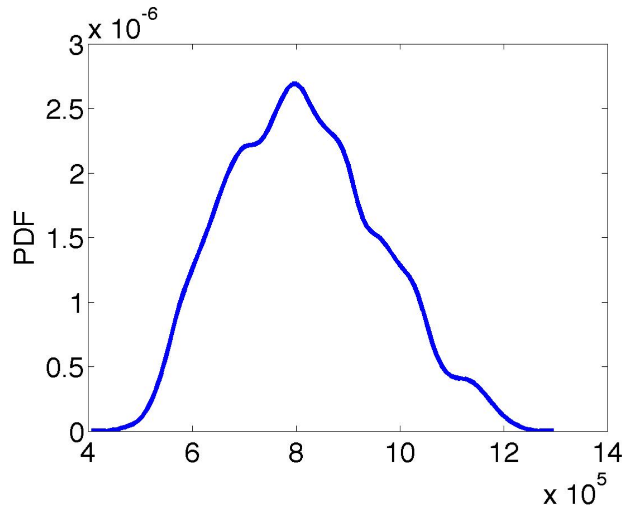
Uncertainty Quantification

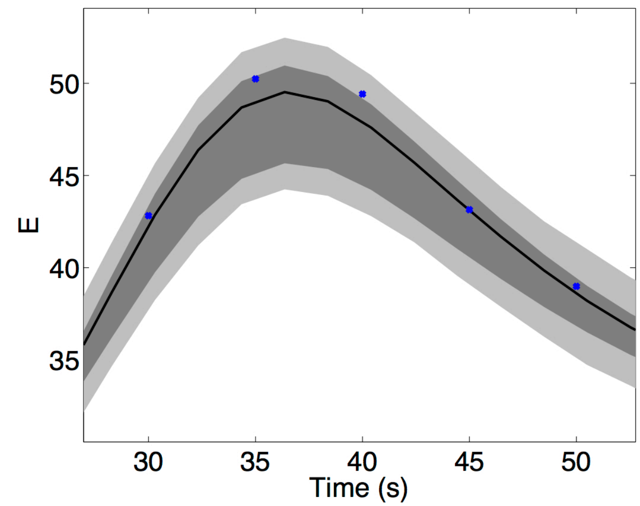
Ralph C. Smith

Department of Mathematics

North Carolina State University







Essentially, all models are wrong, but some are useful, George E.P. Box, Industrial Statistician.

Support: DOE Consortium for Advanced Simulation of LWR (CASL)

NNSA Consortium for Nonproliferation Enabling Capabilities (CNEC)

NSF Data-Enabled Science and Engineering of Atomic Structure (SEAS)

NSF Collaborative Research CDS&E

Air Force Office of Scientific Research (AFOSR)

Course Structure

Overview: 9:00 – 5:30

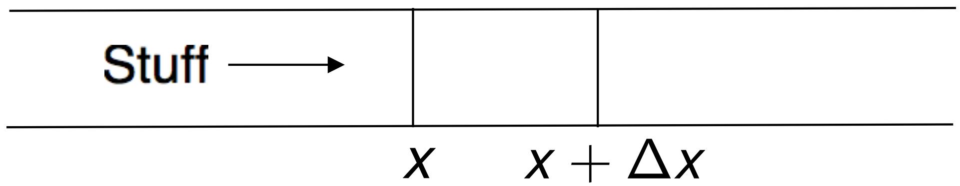
1. Introduction: Motivating examples
2. Overview of terminology and inverse problems
3. Bayesian inference
4. Forward uncertainty propagation
5. Sensitivity analysis and active subspaces
6. Surrogate model construction
7. Model discrepancy

Website:

- http://www4.ncsu.edu/~rsmith/DATAWORKS18/

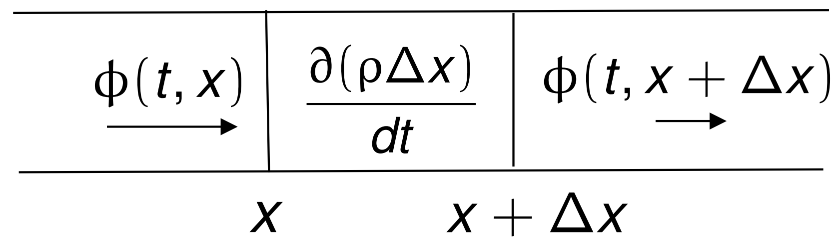
Modeling Strategy

General Strategy: Conservation of stuff



$$\frac{dStuff}{dt} = \text{Stuff in - Stuff out + Stuff created - Stuff destroyed}$$

Continuity Equation:



$$\frac{\partial(\rho \Delta x)}{\partial t} = \phi(t, x) - \phi(t, x + \Delta x)$$

$$\Rightarrow \lim_{\Delta x \to 0} \frac{\partial \rho}{\partial t} = \lim_{\Delta x \to 0} \frac{\phi(t, x) - \phi(t, x + \Delta x)}{\Delta x}$$

$$\Rightarrow \frac{\partial \rho}{\partial t} + \frac{\partial \phi}{\partial x} = 0$$

Density: $\rho(t, x)$ - Stuff per unit length or volume

Rate of Flow: $\phi(t, x)$ - Stuff per second

More Generally:

$$\Rightarrow \frac{\partial \rho}{\partial t} + \frac{\partial \phi}{\partial x} = \text{Sources - Sinks}$$

Example 1: Weather Models

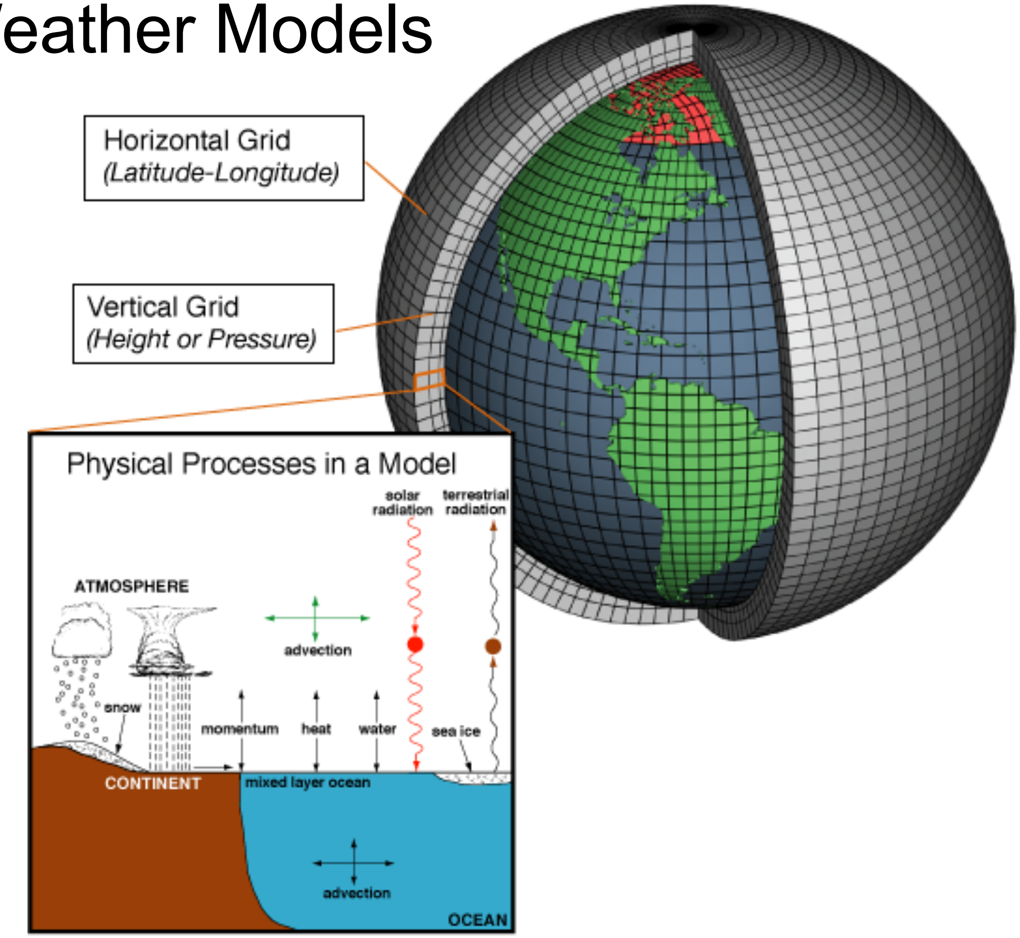
Challenges:

• Coupling between temperature, pressure gradients, precipitation, aerosol, etc.;

• Models and inputs contain uncertainties;

• Numerical grids necessarily larger than many phenomena; e.g., clouds

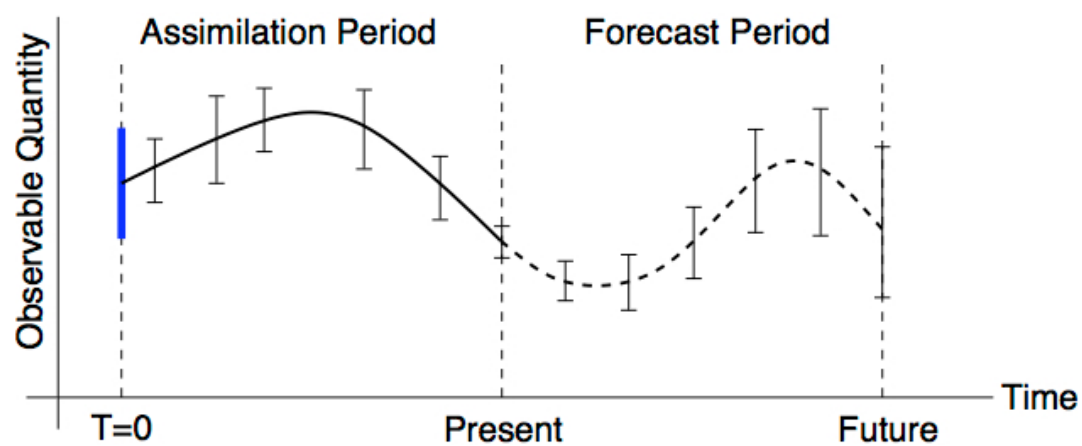
• Sensors positions may be uncertain; e.g., weather balloons, ocean buoys.



Goal:

• Assimilate data to quantify uncertain initial conditions and parameters;

• Make predictions with quantified uncertainties.



Equations of Atmospheric Physics

$$\boxed{\frac{\partial \rho}{\partial t} + \frac{\partial \phi}{\partial x} = \text{Sources - Sinks}}$$

Conservation Relations:

Mass $\quad \dfrac{\partial \rho}{\partial t} + \nabla \cdot (\rho v) = 0$

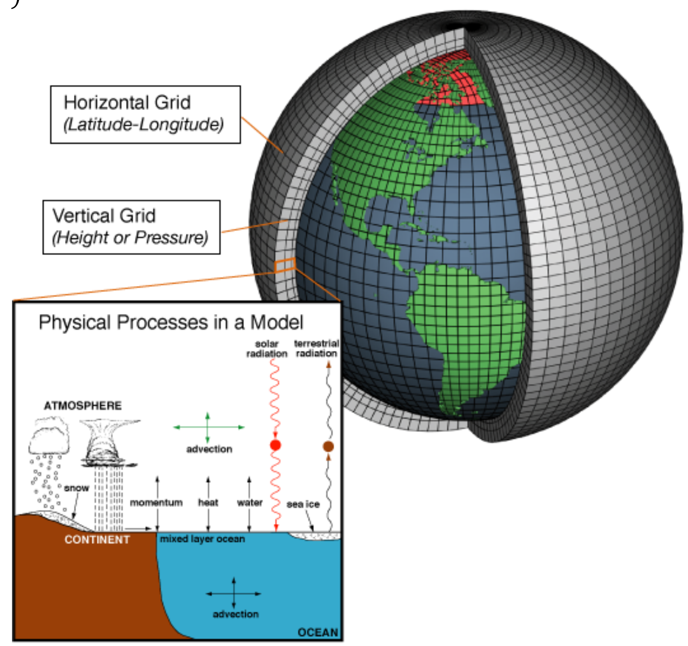
Momentum $\quad \dfrac{\partial v}{\partial t} = -v \cdot \nabla v - \dfrac{1}{\rho} \nabla p - g\hat{k} - 2\Omega \times v$

Energy $\quad \rho c_V \dfrac{\partial T}{\partial t} + p \nabla \cdot v = -\nabla \cdot F + \nabla \cdot (k \nabla T) + \rho \dot{q}(T, p, \rho)$

$$p = \rho R T$$

Water $\quad \dfrac{\partial m_j}{\partial t} = -v \cdot \nabla m_j + S_{m_j}(T, m_j, \chi_j, \rho)\ ,\ j = 1, 2, 3,$

Aerosol $\quad \dfrac{\partial \chi_j}{\partial t} = -v \cdot \nabla \chi_j + S_{\chi_j}(T, \chi_j, \rho)\ ,\ j = 1, \cdots, J,$



Constitutive Closure Relations: e.g.,

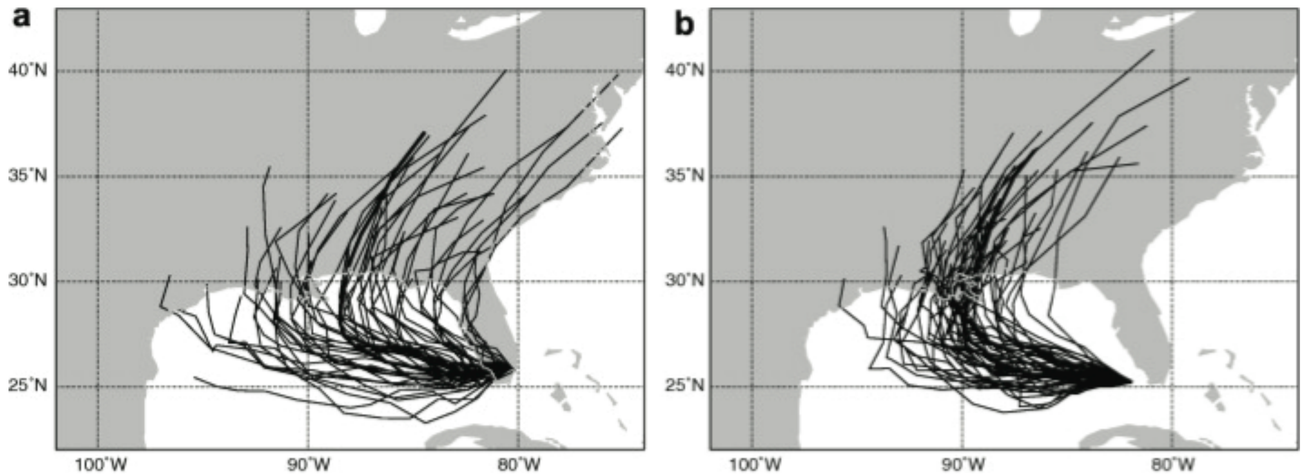
$$S_{m_2} = S_1 + S_2 + S_3 - S_4$$

where

$$S_1 = \bar{\rho}(m_2 - m_2^*)^2 \left[\underline{1.2 \times 10^{-4}} + \left(\underline{1.569 \times 10^{-12}} \frac{n_r}{d_0 (m_2 - m_2^*)} \right) \right]^{-1}$$

Ensemble Predictions

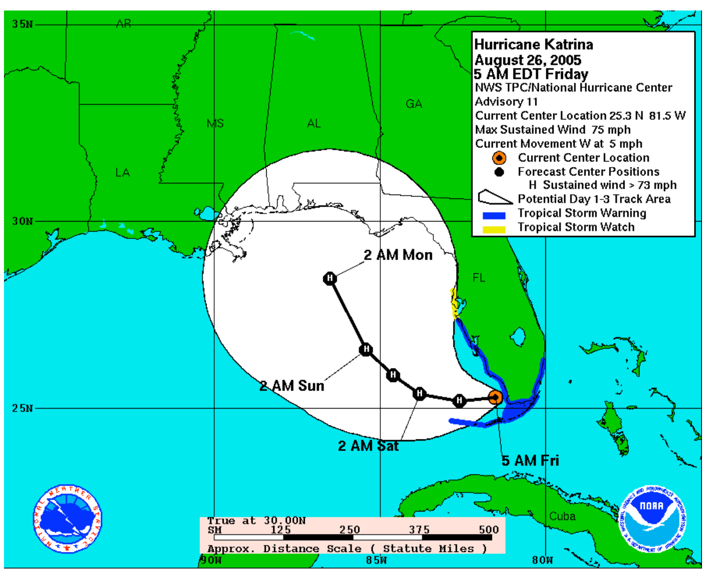
Ensemble Predictions:



00 UTC on August 26, 2005

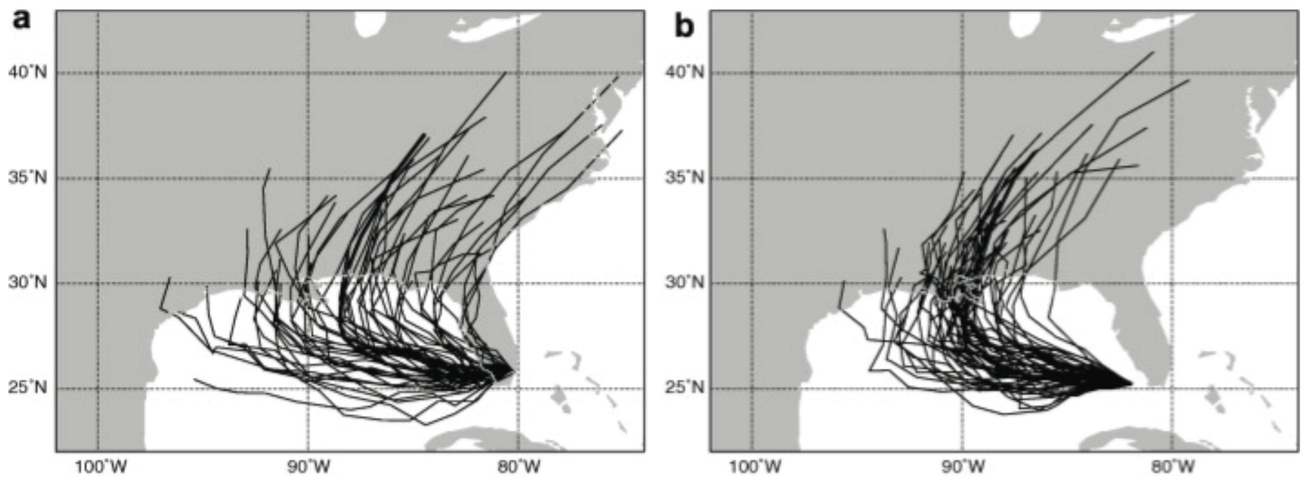
12 UTC on August 26, 2005

Cone of Uncertainty:



Ensemble Predictions

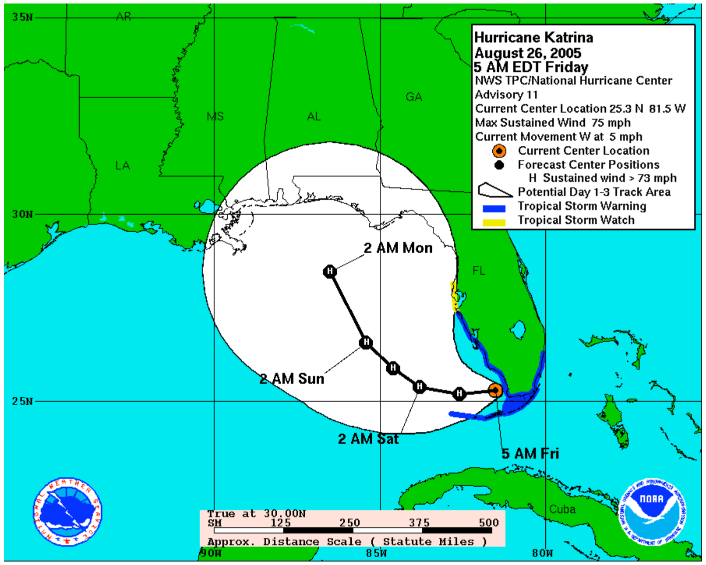
Ensemble Predictions:



00 UTC on August 26, 2005

12 UTC on August 26, 2005

Cone of Uncertainty:



General Questions:

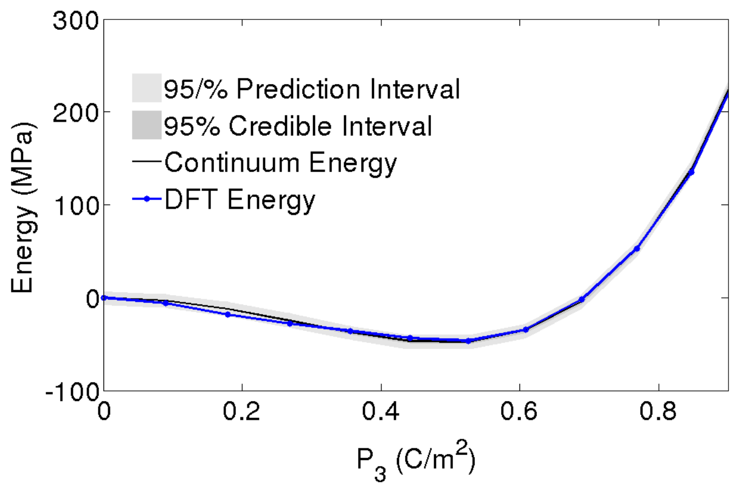
- What is expected rainfall on March 20?
- What are high and low temperatures?
- What is predicted average snow fall?
- Note: Quantities are statistical in nature.

Example 2: Quantum-Informed Continuum Models

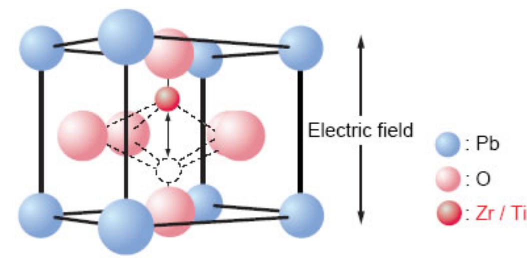
Objectives:

- Employ density function theory (DFT) to construct/calibrate continuum energy relations.
  - e.g., Helmholtz energy

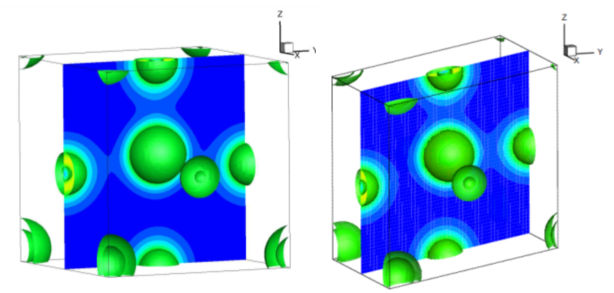
$$\psi(P) = \alpha_1 P^2 + \alpha_{11} P^4 + \alpha_{111} P^6$$



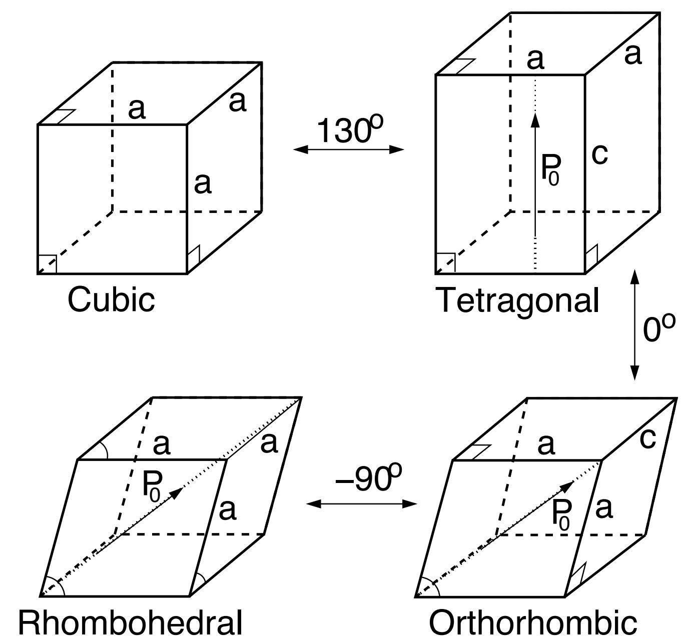
Helmholtz Energy



Lead Titanate Zirconate (PZT)



DFT Electronic Structure Simulation



UQ and SA Issues:

- Is 6th order term required to accurately characterize material behavior?

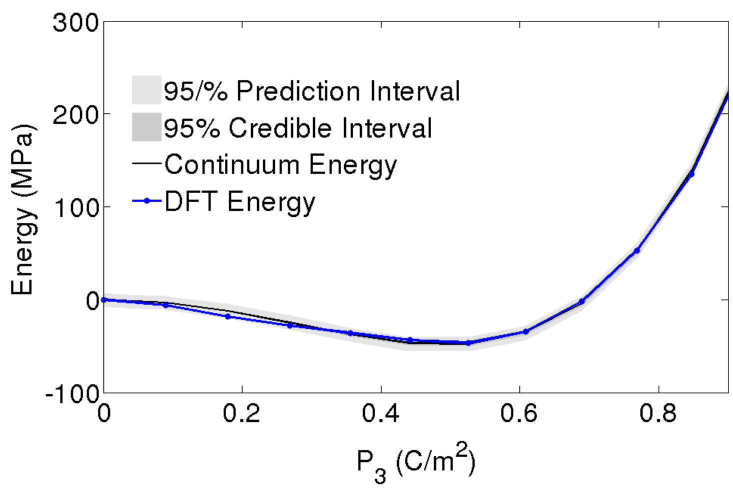
- **Note:** Determines molecular structure

Quantum-Informed Continuum Models

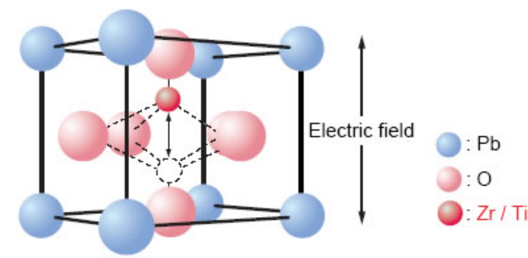
Objectives:

- Employ density function theory (DFT) to construct/calibrate continuum energy relations.
  - e.g., Helmholtz energy

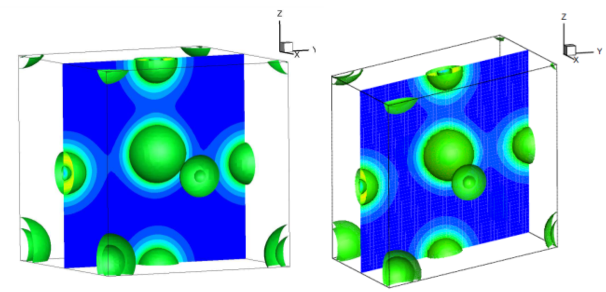
$$\psi(P) = \alpha_1 P^2 + \alpha_{11} P^4 + \alpha_{111} P^6$$



Helmholtz Energy



Lead Titanate Zirconate (PZT)



DFT Electronic Structure Simulation

UQ and SA Issues:

- Is 6^{th} order term required to accurately characterize material behavior?
- **Note:** Determines molecular structure

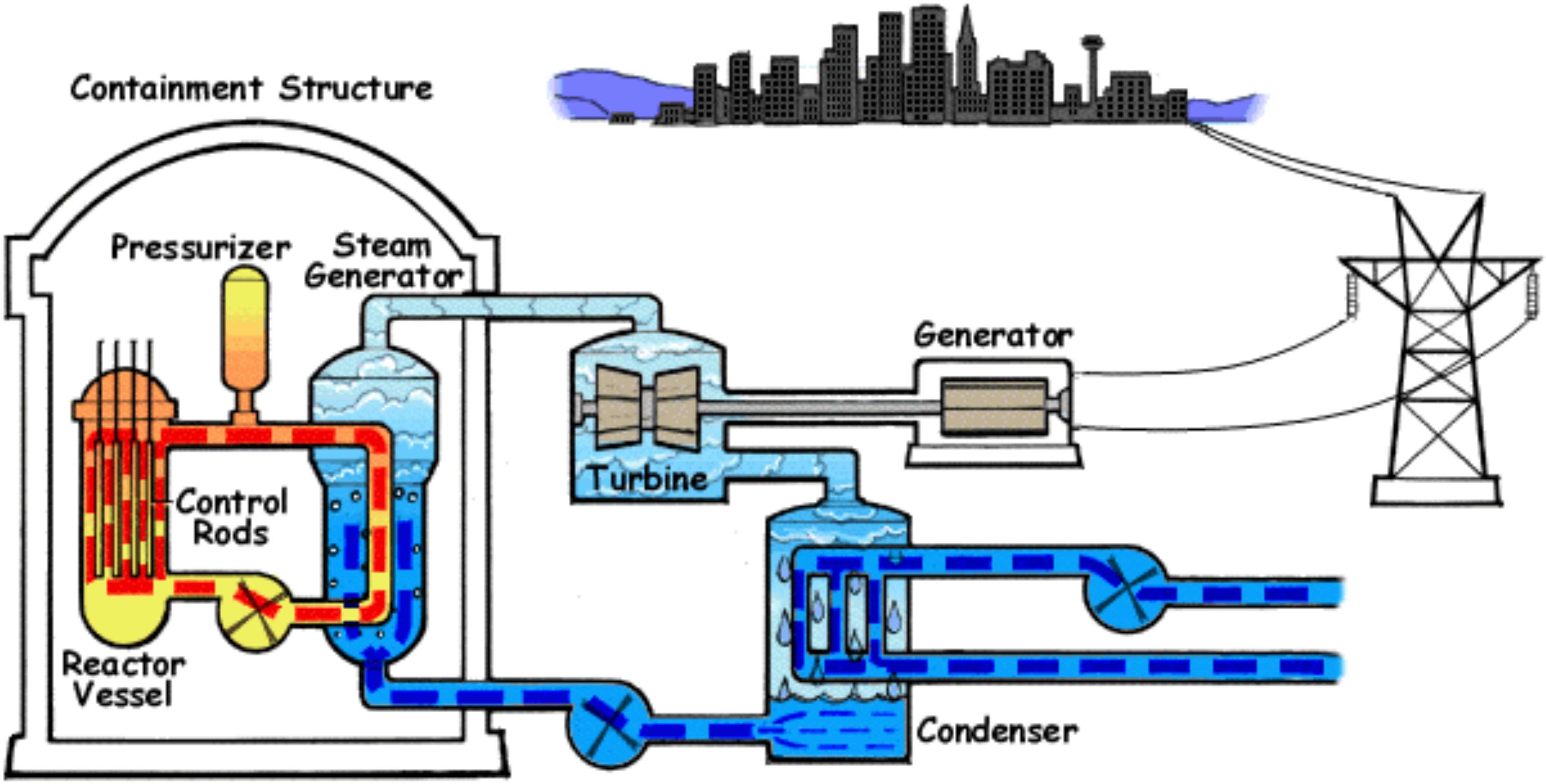
Broad Objective:

- Use UQ/SA to help bridge scales from quantum to system

Note:

- Linearly parameterized

Example 3: Pressurized Water Reactors (PWR)



Models:

- Involve neutron transport, thermal-hydraulics, chemistry.
- Inherently multi-scale, multi-physics.

CRUD Measurements: Consist of low resolution images at limited number of locations.

Example: Pressurized Water Reactors (PWR)

3-D Neutron Transport Equations:

$$\frac{1}{|v|}\frac{\partial \varphi}{\partial t} + \Omega \cdot \nabla \varphi + \Sigma_t(r,E)\varphi(r,E,\Omega,t)$$

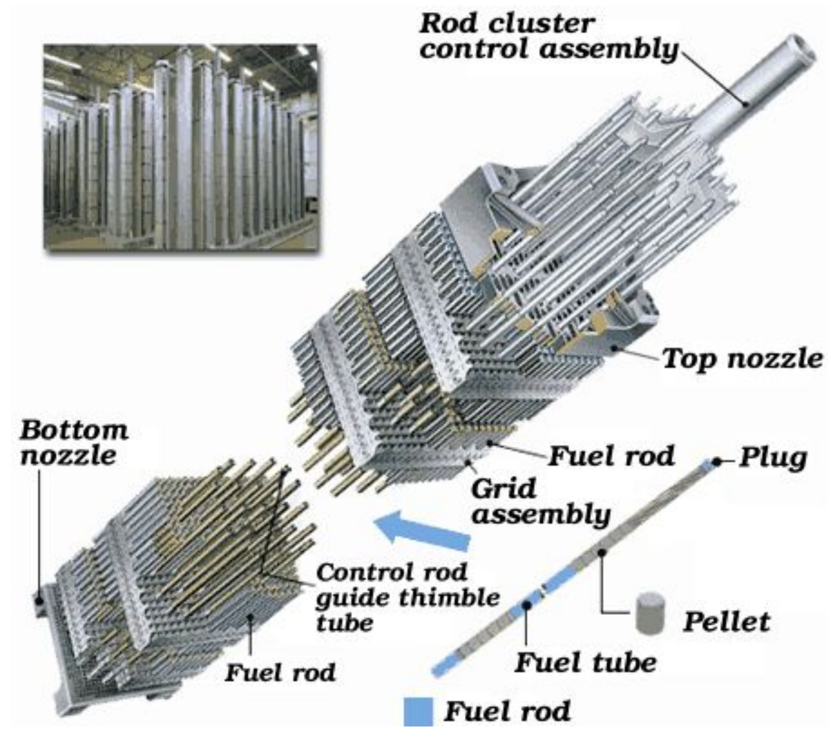
$$= \int_{4\pi} d\Omega' \int_0^\infty dE' \Sigma_s(E' \rightarrow E, \Omega' \rightarrow \Omega)\varphi(r,E',\Omega',t)$$

$$+ \frac{\chi(E)}{4\pi} \int_{4\pi} d\Omega' \int_0^\infty dE' \nu(E')\Sigma_f(E')\varphi(r,E',\Omega',t)$$

$$\boxed{\frac{\partial \rho}{\partial t} + \frac{\partial \phi}{\partial x} = \text{Sources - Sinks}}$$

Challenges:

- Very large number of inputs; e.g., 100,000; Active subspace construction critical.
- ORNL Code SCALE: Can take hours to run.
- Time-dependent surrogate models must accommodate PDE structure.
- Numerical errors often difficult to quantify.
- Predicting future requires extrapolatory or out-of-data predictions; one must address model discrepancy to construct validation intervals.



Example: Pressurized Water Reactors (PWR)

Thermo-Hydraulic Equations: Mass, momentum and energy balance for fluid

$$\frac{\partial}{\partial t}(\alpha_f \rho_f) + \nabla \cdot (\alpha_f \rho_f v_f) = -\Gamma$$

$$\alpha_f \rho_f \frac{\partial v_f}{\partial t} + \alpha_f \rho_f v_f \cdot \nabla v_f + \nabla \cdot \sigma_f^R + \alpha_f \nabla \cdot \sigma + \alpha_f \nabla p_f$$
$$= -F^R - F + \Gamma (v_f - v_g)/2 + \alpha_f \rho_f g$$

$$\frac{\partial}{\partial t}(\alpha_f \rho_f e_f) + \nabla \cdot (\alpha_f \rho_f e_f v_f + Th) = (T_g - T_f)H + T_f \Delta_f$$
$$-T_g(H - \alpha_g \nabla \cdot h) + h \cdot \nabla T - \Gamma[e_f + T_f(s^* - s_f)]$$
$$-p_f \left(\frac{\partial \alpha_f}{\partial t} + \nabla \cdot (\alpha_f v_f) + \frac{\Gamma}{\rho_f} \right)$$

$$\boxed{\frac{\partial \rho}{\partial t} + \frac{\partial \phi}{\partial x} = \text{Sources - Sinks}}$$

Notes:

- Similar relations for gas and bubbly phases
- Surrogate models must conserve mass, energy, and momentum

Challenges:

- Codes can have 15-30 closure relations and up to 75 parameters.
- Codes and closure relations often ”borrowed” from other physical phenomena; e.g., single phase fluids, airflow over a car (CFD code STAR-CCM+)
- Calibration necessary and closure relations can conflict.
- Inference of random fields requires high- (infinite-) dimensional theory.

Example: Pressurized Water Reactors (PWR)

Thermo-Hydraulic Equations: Mass, momentum and energy balance for fluid

$$\frac{\partial}{\partial t}(\alpha_f \rho_f) + \nabla \cdot (\alpha_f \rho_f v_f) = -\Gamma$$

$$\alpha_f \rho_f \frac{\partial v_f}{\partial t} + \alpha_f \rho_f v_f \cdot \nabla v_f + \nabla \cdot \sigma_f^R + \alpha_f \nabla \cdot \sigma + \alpha_f \nabla p_f$$
$$= -F^R - F + \Gamma(v_f - v_g)/2 + \alpha_f \rho_f g$$

$$\frac{\partial}{\partial t}(\alpha_f \rho_f e_f) + \nabla \cdot (\alpha_f \rho_f e_f v_f + Th) = (T_g - T_f)H + T_f \Delta_f$$
$$-T_g(H - \alpha_g \nabla \cdot h) + h \cdot \nabla T - \Gamma[e_f + T_f(s^* - s_f)]$$
$$-p_f\left(\frac{\partial \alpha_f}{\partial t} + \nabla \cdot (\alpha_f v_f) + \frac{\Gamma}{\rho_f}\right)$$

$$\boxed{\frac{\partial \rho}{\partial t} + \frac{\partial \phi}{\partial x} = \text{Sources - Sinks}}$$

Notes:

- Similar relations for gas and bubbly phases
- Surrogate models must conserve mass, energy, and momentum

Example: Shearon Harris outside Raleigh

UQ Questions:

- What is peak operating temperature?
- What is expected level of CRUD buildup?
- What is risk associated with operating regime?
- What is expected profit for new design?

Example 4: SIR Model for Disease Dynamics

SIR Model:

$$\frac{dS}{dt} = \delta N - \delta S - \gamma \underline{kIS} \quad , \; S(0) = S_0 \qquad \text{Susceptible}$$

$$\frac{dI}{dt} = \underline{\gamma k}IS - (r + \delta)I \quad , \; I(0) = I_0 \qquad \text{Infectious}$$

$$\frac{dR}{dt} = rI - \delta R \quad , \; R(0) = R_0 \qquad \text{Recovered}$$

Parameters:

- γ: Infection coefficient
- k: Interaction coefficient
- r: Recovery rate
- δ: Birth/death rate

Response:

$$y = \int_0^5 R(t, q)dt$$

Note: Parameters $q = [\gamma, k, r, \delta]$ not uniquely determined by data

Note: Presently employed cholera models have similar form; example this afternoon.

SIR Disease Example

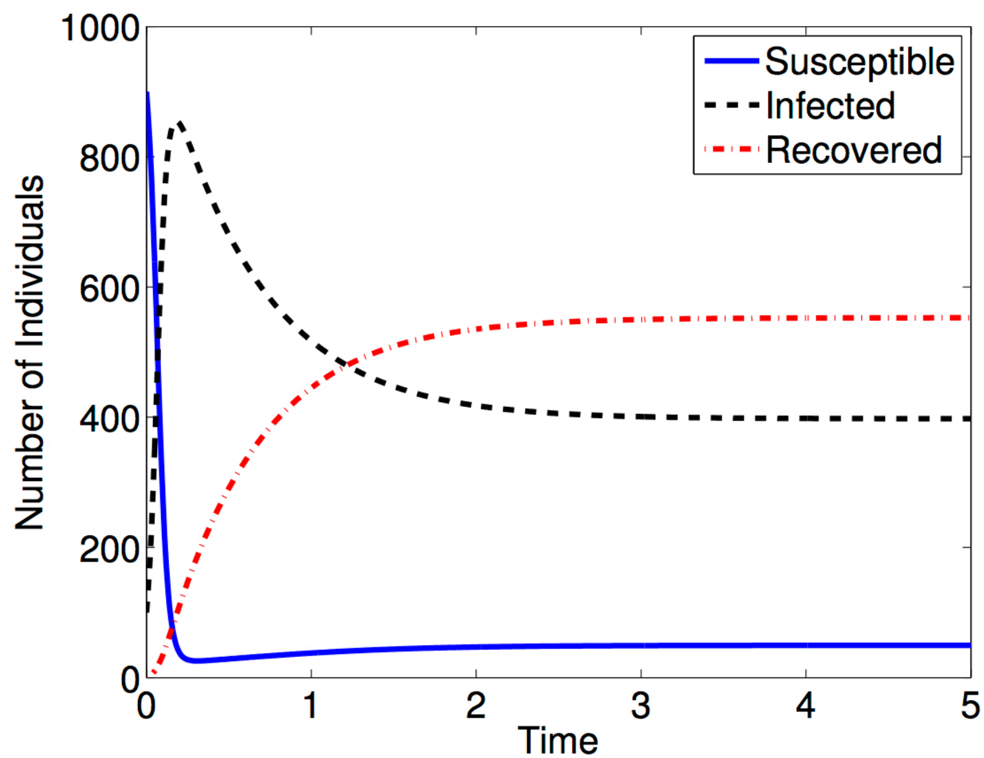
SIR Model:

$$\frac{dS}{dt} = \delta N - \delta S - \gamma \underline{kI}S \quad , \quad S(0) = S_0 \qquad \text{Susceptible}$$

$$\frac{dI}{dt} = \gamma \underline{kI}S - (r+\delta) I \quad , \quad I(0) = I_0 \qquad \text{Infectious}$$

$$\frac{dR}{dt} = rI - \delta R \quad , \quad R(0) = R_0 \qquad \text{Recovered}$$

Typical Realization:



SIR Disease Example

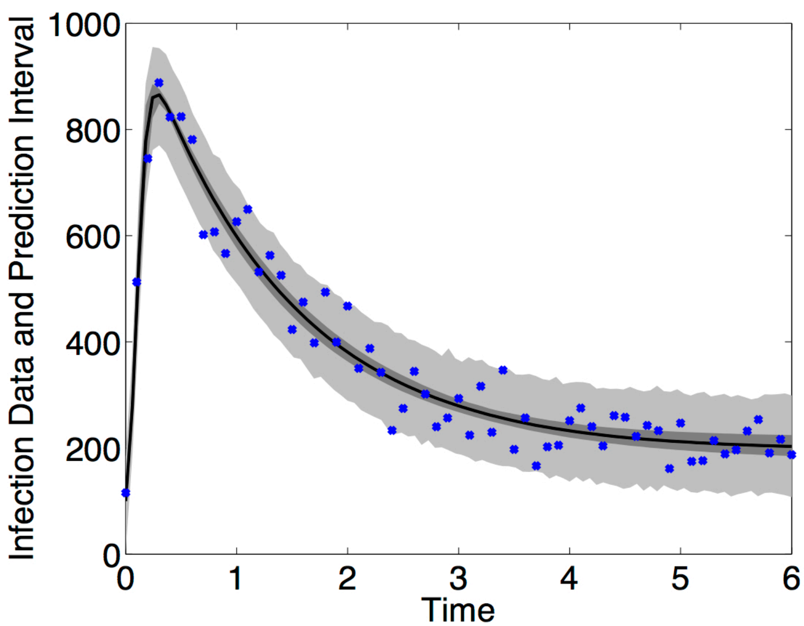
SIR Model:

$$\frac{dS}{dt} = \delta N - \delta S - \gamma k I S \quad , \quad S(0) = S_0$$

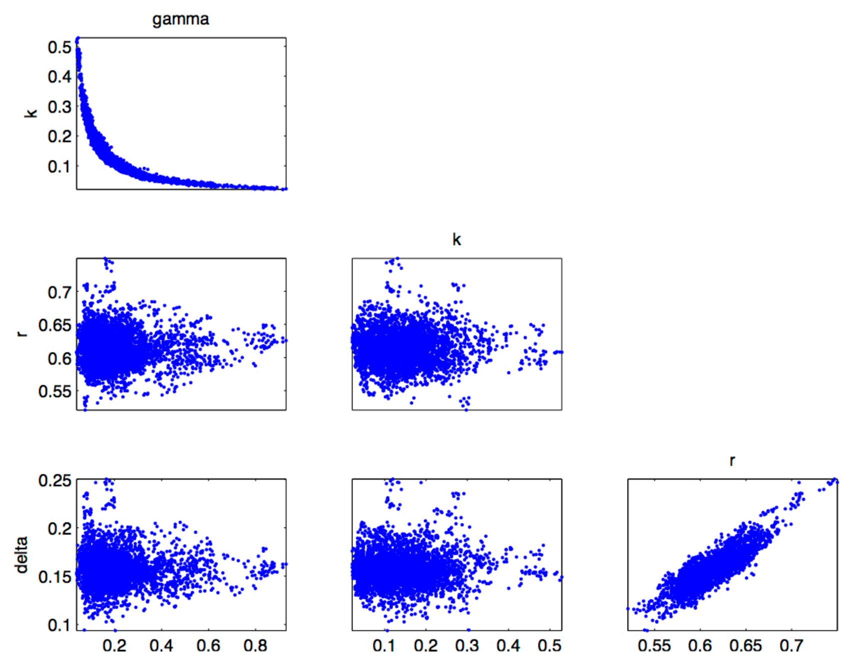
$$\frac{dI}{dt} = \gamma k I S - (r + \delta) I \quad , \quad I(0) = I_0$$

$$\frac{dR}{dt} = rI - \delta R \quad , \quad R(0) = R_0$$

UQ Goal: Predict I(t) with uncertainty intervals:



Problem: Cannot uniquely infer parameters



Solution:

- Active subspaces
- Identifiability analysis
- Sensitivity analysis
- Design of experiments

Example 5: HIV Model for Characterization and Control Regimes

HIV Model:

$$\dot{T}_1 = \lambda_1 - d_1 T_1 - (1-\varepsilon)k_1 V T_1$$

$$\dot{T}_2 = \lambda_2 - d_2 T_2 - (1-f\varepsilon)k_2 V T_2$$

$$\dot{T}_1^* = (1-\varepsilon)k_1 V T_1 - \delta T_1^* - m_1 E T_1^*$$

$$\dot{T}_2^* = (1-f\varepsilon)k_2 V T_2 - \delta T_2^* - m_2 E T_2^*$$

$$\dot{V} = N_T \delta (T_1^* + T_2^*) - cV - [(1-\varepsilon)\rho_1 k_1 T_1 + (1-f\varepsilon)\rho_2 k_2 T_2] V$$

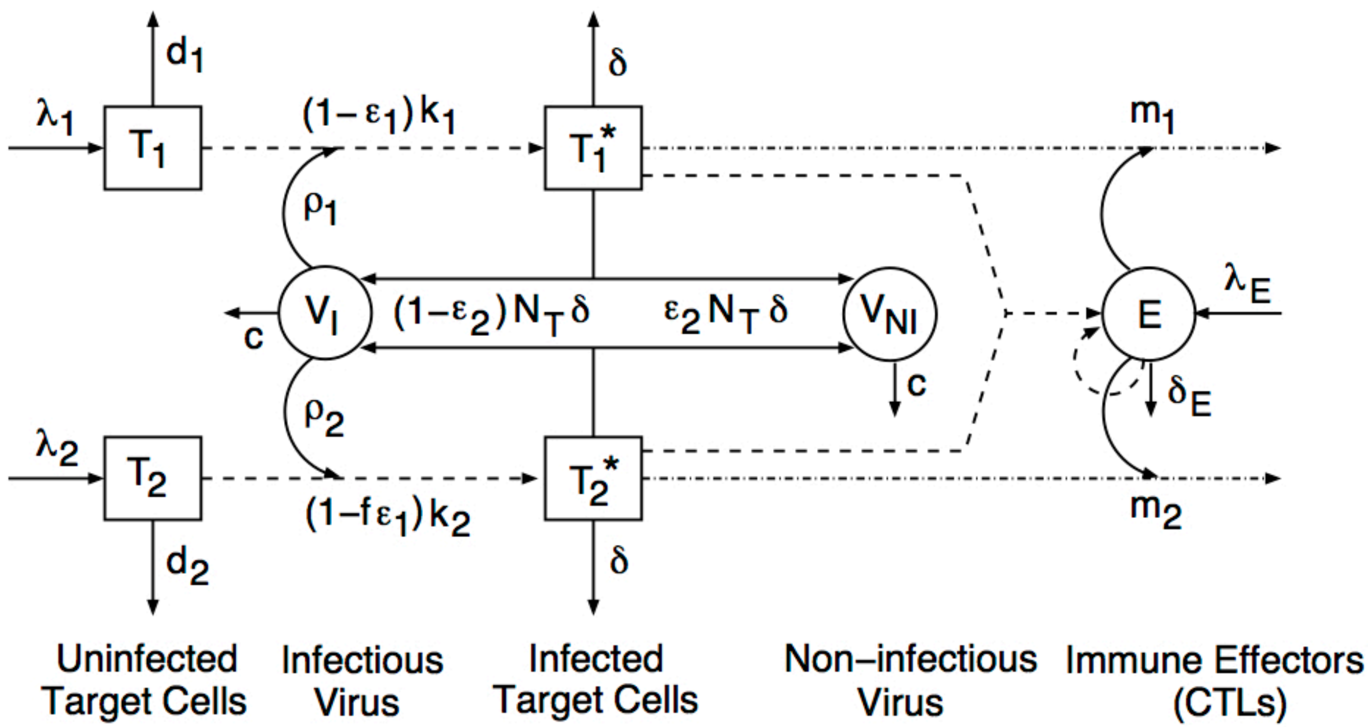
$$\dot{E} = \lambda_E + \frac{b_E(T_1^* + T_2^*)}{T_1^* + T_2^* + K_b} E - \frac{d_E(T_1^* + T_2^*)}{T_1^* + T_2^* + K_d} E - \delta_E E$$

Notes: 21 parameters

[Adams, Banks et al., 2005, 2007]

Notation: $\dot{E} \equiv \dfrac{dE}{dt}$

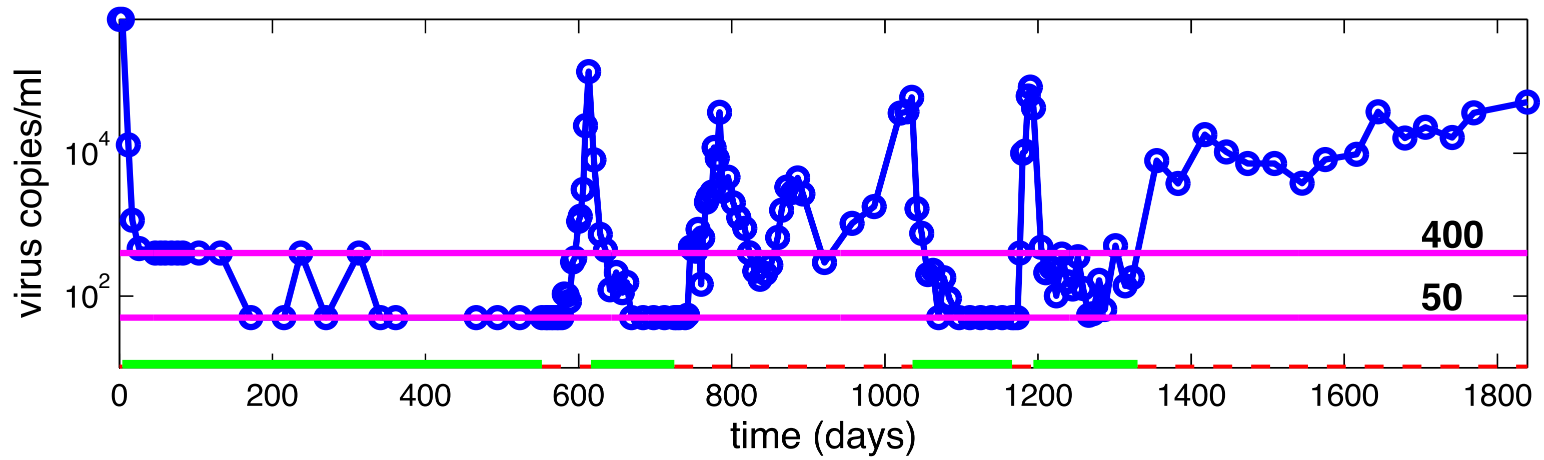
Compartments:



Example: HIV Model for Characterization and Treatment Regimes

HIV Model: Several sources of uncertainty including viral measurement techniques

Example: Upper and lower limits to assay sensitivity



UQ Questions:

- What are the uncertainties in parameters that cannot be directly measured?
- What is optimal treatment regime that is “safe” for patient?
- What is expected viral load? Issue: very often requires high-dimensional integration!
- e.g., $\mathbb{E}[V(t)] = \int_{\mathbb{R}^{21}} V(t,q)\rho(q)dq$

Experimental results are believed by everyone, except for the person who ran the experiment, source anonymous, quoted by Max Gunzburger, Florida State University.

2. Challenge: Terminology and Notation

Terminology:

- Inputs: Parameters, initial conditions, boundary conditions, exogenous forces; e.g., parameters in HIV models, initial conditions in weather models.

- Outputs or Responses: Quantities that we experimentally or numerically measure; e.g., viral load, outlet temperature in reactor.

- Quantities of Interest (QoI): Statistical quantity that we want to compute; e.g., average CRUD buildup, expected profit for a given design.

Input Notation: Can vary even within disciplines!

- Math Control Community: $q = [q_1, \ldots, q_p]$
- Math Reduced-Order Community: $p = [p_1, \ldots, p_q]$
- Statistics: $\theta = [\theta_1, \ldots, \theta_d]$
- Nuclear Engineering: $\alpha = [\alpha_1, \ldots, \alpha_k]$
- Active subspace community: $x = [x_1, \ldots, x_p]$

Note: Same variability in notation for outputs and quantities of interest

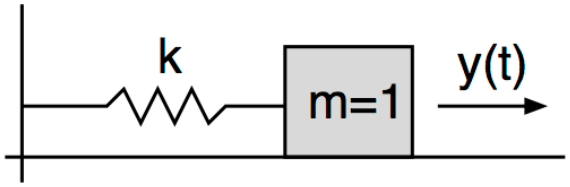
First Challenge: Terminology and Notation

Terminology:

- Linearly parameterized problems: e.g., portfolio model $y = c_1 q_1 + c_2 q_2$
  - Rare in applications except constitutive relations and image processing
- Nonlinearly parameterized problems: typical case
  - Differs from linear or nonlinear in state; e.g., spring model

$$\frac{d^2 y(t)}{dt^2} + k y(t) = 0$$

$$y(0) = y_0 \ , \ \frac{dy}{dt}(0) = 0$$



Notation: $\dot{y} \equiv \frac{dy}{dt} \ , \ \ddot{y} \equiv \frac{d^2 y}{dt^2}$

Inputs: $q = [k, y_0]$

Response: Displacement $y(t) = y_0 \cos(\sqrt{k} \cdot t)$

$$\Rightarrow \quad \begin{aligned} &\ddot{y}(t) + k y(t) = 0 \\ &y(0) = y_0 \ , \ \frac{dy}{dt}(0) = 0 \end{aligned}$$

Note:

- Linear state dependence
- Nonlinear parameter dependence

Uncertainty Quantification

I have always done uncertainty quantification. The difference now is that it is capitalized. Bill Browning, Applied Mathematics Incorporated.

Note: The field of "Uncertainty Quantification" has grown rapidly over the last 20 years. How is "Capital UQ" different from what statisticians do extremely well every day?

- E.g., When I proposed a course on "Uncertainty Quantification" in Mathematics, I had to carefully justify its existence to Statistics.
- Statistics students are now starting to take the course.

Uncertainty Quantification

I have always done uncertainty quantification. The difference now is that it is capitalized. Bill Browning, Applied Mathematics Incorporated.

Note: The field of "Uncertainty Quantification" has grown rapidly over the last 20 years. How is "Capital UQ" different from what statisticians do extremely well every day?

- E.g., When I proposed a course on "Uncertainty Quantification" in Mathematics, I had to carefully justify its existence to Statistics.
- Statistics students are now starting to take the course.

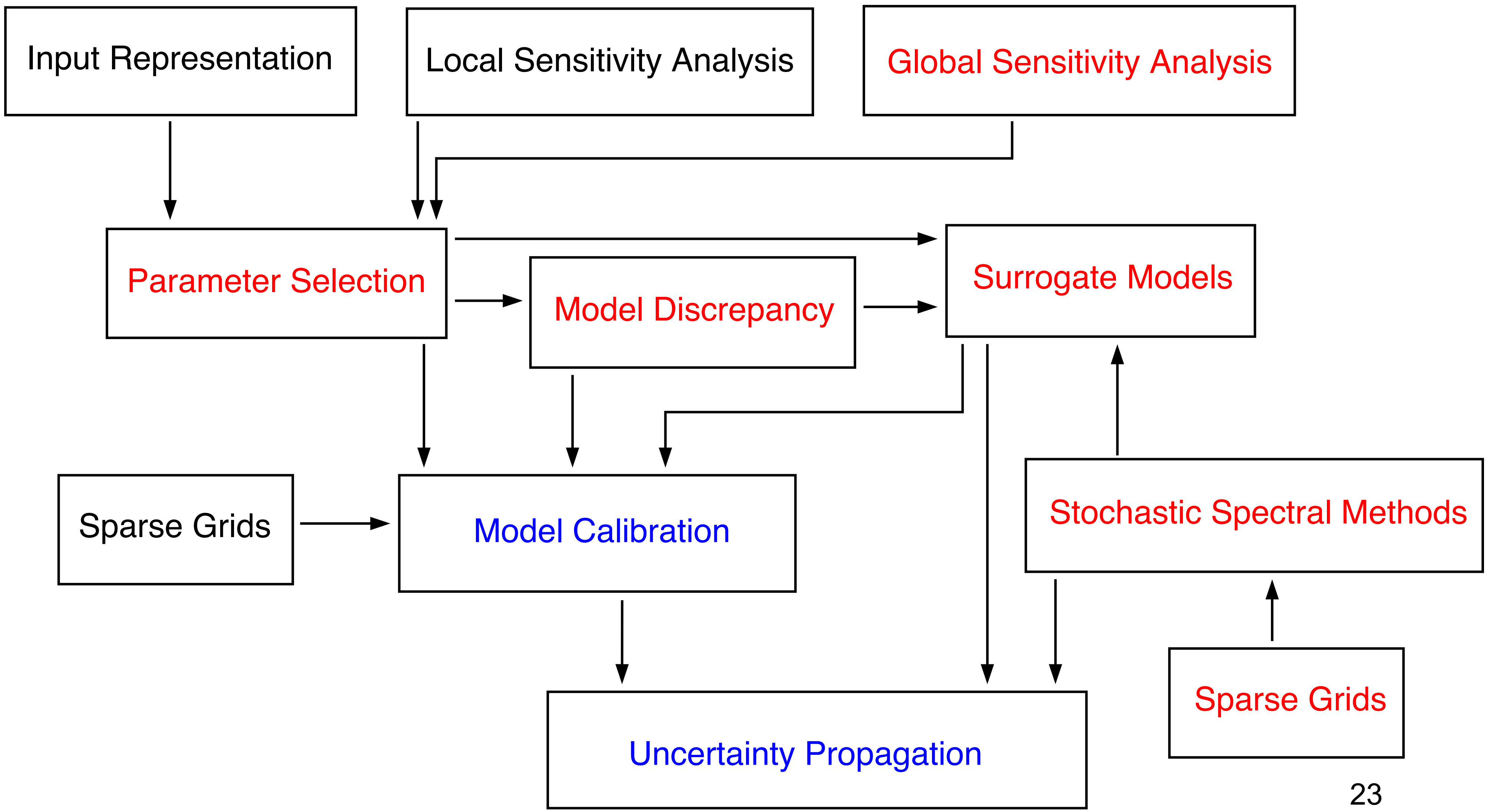
My Definition of "Capital UQ": The synergy between statistics, applied mathematics and domain sciences required to quantify uncertainties in inputs and QoI when models are too computationally complex to permit sole reliance on sampling-based methods."

- Involves orthogonal polynomial techniques, sparse grids, high-D (infinite-D) approximation theory, randomized linear algebra … and a lot of statistics!

No one trusts a model except the man who wrote it; everyone trusts an observation except the man who made it, Harlow Shapely.

Steps in Uncertainty Quantification

Note: Uncertainty quantification requires synergy between statistics, mathematics and application area.



Model Calibration

Sources of Uncertainty:

- Model
- Parameters
- Sensor measurements
- Initial conditions

Strategy:

- Quantify uncertainty in parameters
- Propagate uncertainty through model

Example: HIV model

$$\dot{T}_1 = \underline{\lambda_1} - \underline{d_1} T_1 - (1-\varepsilon) k_1 V T_1$$

$$\dot{T}_2 = \lambda_2 - d_2 T_2 - (1 - f\varepsilon) k_2 V T_2$$

$$\dot{T}_1^* = (1-\varepsilon) k_1 V T_1 - \underline{\delta} T_1^* - m_1 E T_1^*$$

$$\dot{T}_2^* = (1 - f\varepsilon) \underline{k_2} V T_2 - \underline{\delta} T_2^* - m_2 E T_2^*$$

$$\dot{V} = N_T \underline{\delta} (T_1^* + T_2^*) - cV - [(1-\varepsilon)\rho_1 k_1 T_1 + (1 - f\varepsilon)\rho_2 k_2 T_2] V$$

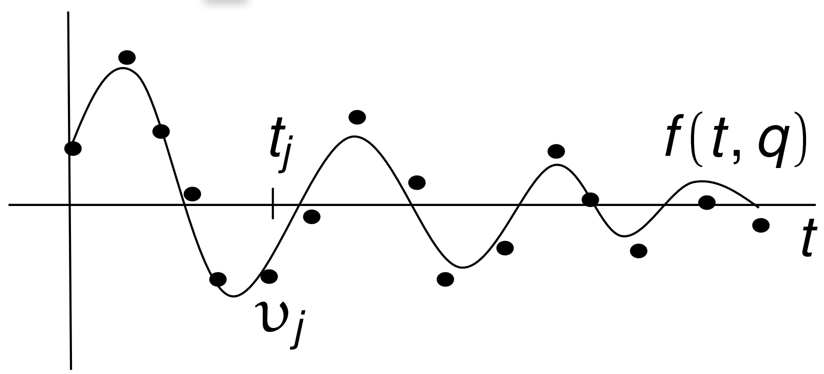
$$\dot{E} = \lambda_E + \frac{\underline{b_E}(T_1^* + T_2^*)}{T_1^* + T_2^* + \underline{K_b}} E - \frac{d_E (T_1^* + T_2^*)}{T_1^* + T_2^* + K_d} E - \delta_E E$$

Parameters: Reduced set

$$q = [b_E, \delta, d_1, k_2, \lambda_1, K_b]$$

Point Estimates: Ordinary least squares

$$q^0 = \arg\min_q \frac{1}{2} \sum_{j=1}^{N} [\upsilon_j - f(t_j, q)]^2$$



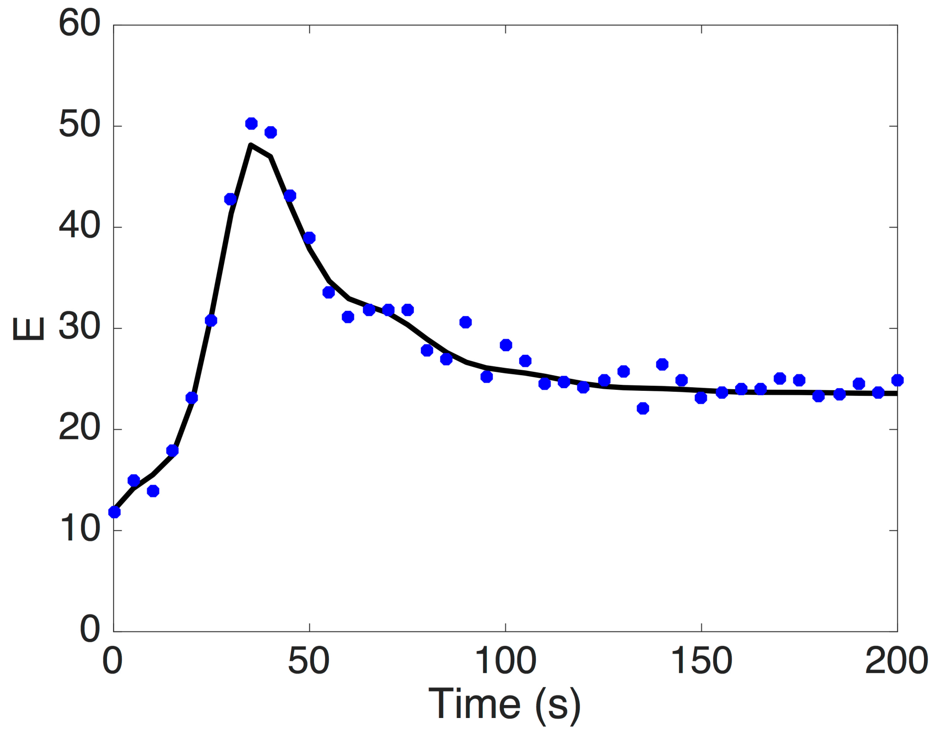
Note: Scaling critical since parameter values vary by 8 orders of magnitude.

Model Calibration and Predictions

Optimization Results:

b_E	δ	d_1	k_2	λ_1	K_b
0.30	0.68	9.1×10^{-3}	1.22×10^{-4}	9.95×10^{3}	88.5

Data and Prediction of Immune Effector Response E:



Note: Point estimates but no quantification of uncertainty in:

- Model
- Parameters
- Data

Goals:

- Replace point estimates with distributions.
- Construct credible and prediction intervals.
- Natural in a Bayesian framework

Objectives for Uncertainty Quantification

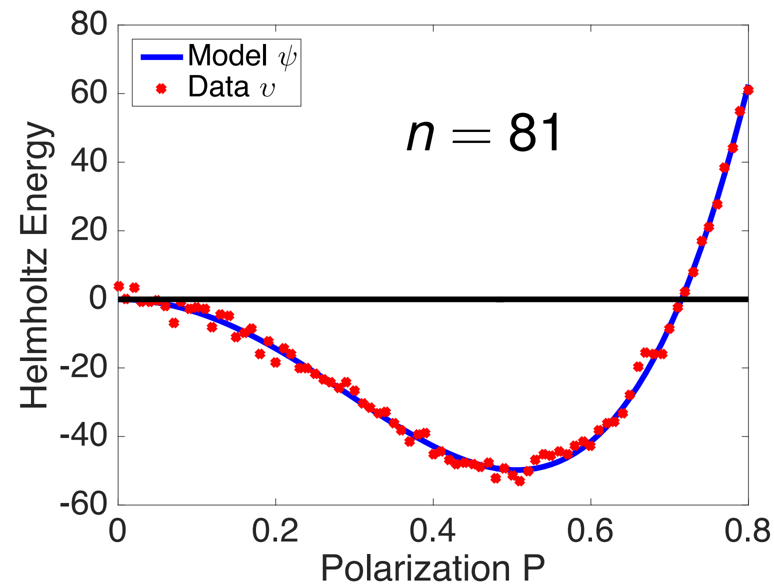
Example: Helmholtz energy $\psi(P, q) = \alpha_1 P^2 + \alpha_{11} P^4$, $q = [\alpha_1, \alpha_{11}]$

Statistical Model: Describes observation process

$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \quad , \; i = 1, \ldots, n$$

Common Assumption: $\varepsilon_i \sim N(0, \sigma^2)$

UQ Goals: Quantify parameter and response uncertainties



Strategy 1: Perform Experiments

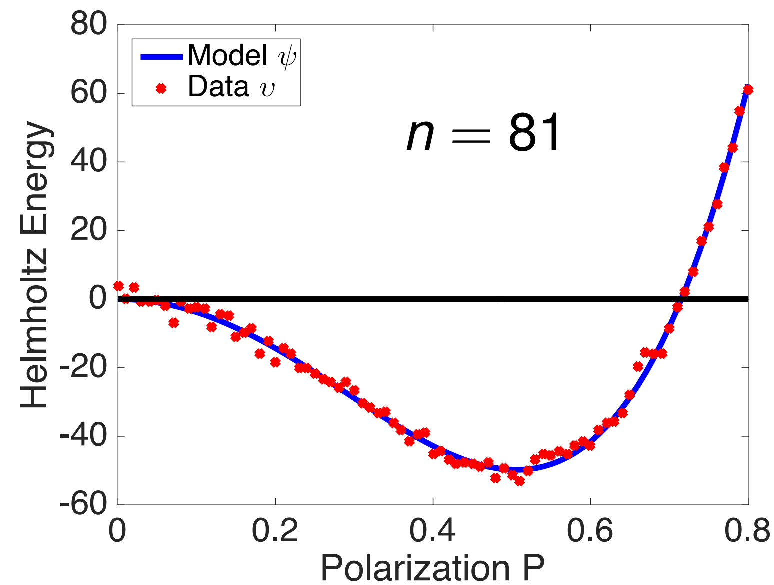
Example: Helmholtz energy $\psi(P, q) = \alpha_1 P^2 + \alpha_{11} P^4$, $q = [\alpha_1, \alpha_{11}]$

Statistical Model: Describes observation process

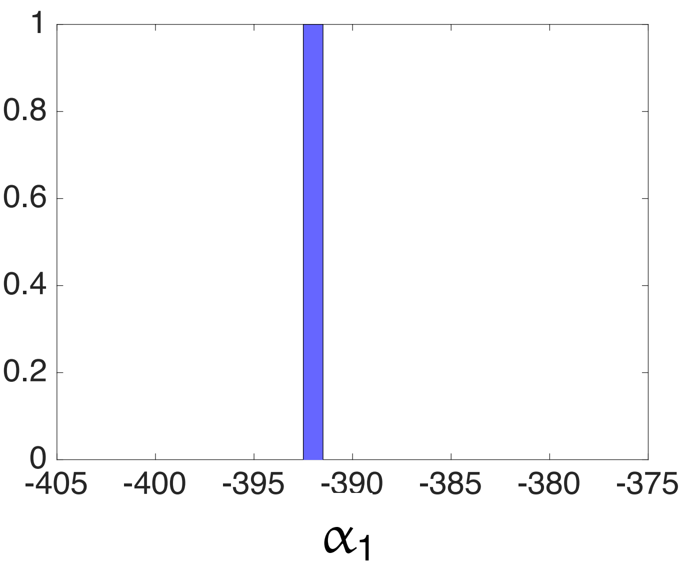
$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \quad , \; i = 1, \ldots, n$$

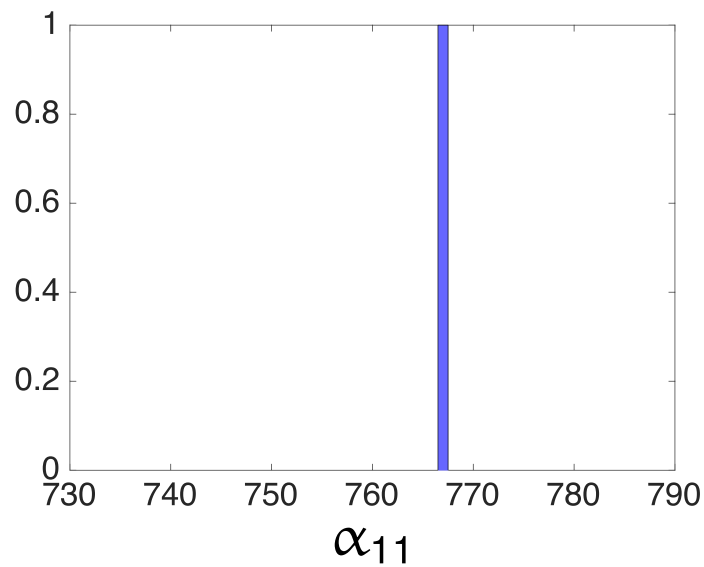
Common Assumption: $\varepsilon_i \sim N(0, \sigma^2)$

UQ Goals: Quantify parameter and response uncertainties



Strategy 1: Perform experiments; e.g., 1



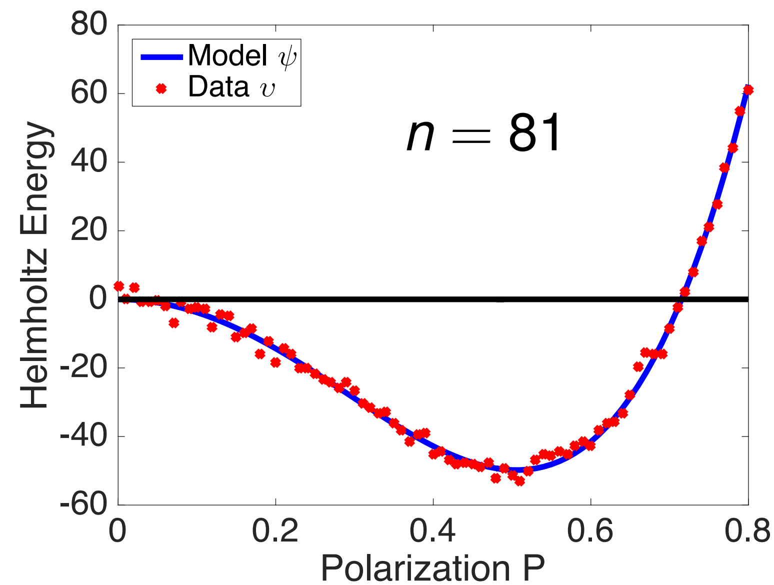


Strategy 1: Perform Experiments

Example: Helmholtz energy $\psi(P,q) = \alpha_1 P^2 + \alpha_{11} P^4$, $q = [\alpha_1, \alpha_{11}]$

Statistical Model: Describes observation process

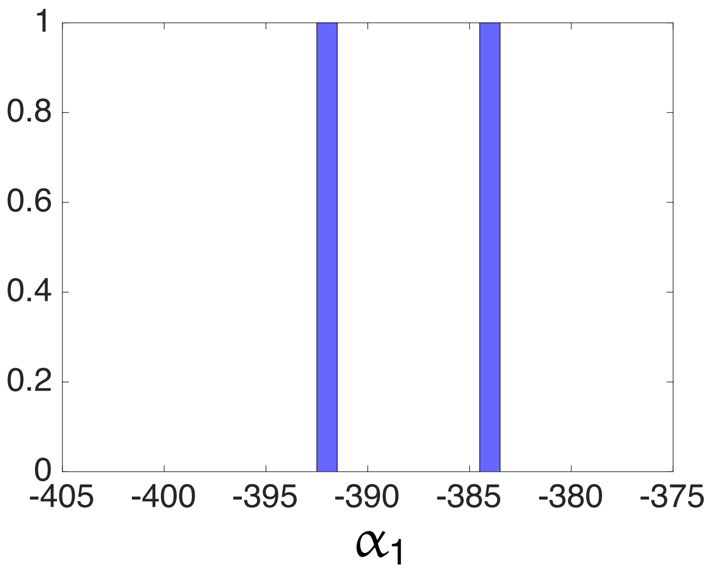
$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \quad , \; i = 1, \ldots, n$$

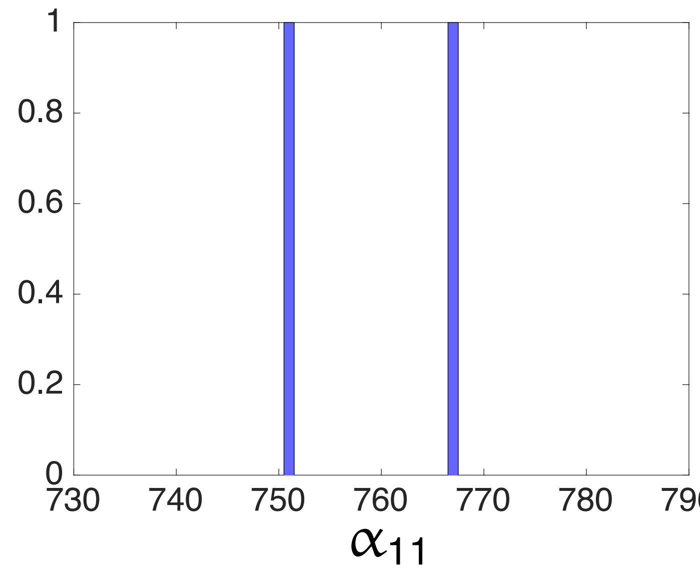


Common Assumption: $\varepsilon_i \sim N(0, \sigma^2)$

UQ Goals: Quantify parameter and response uncertainties

Strategy 1: Perform experiments; e.g., 2



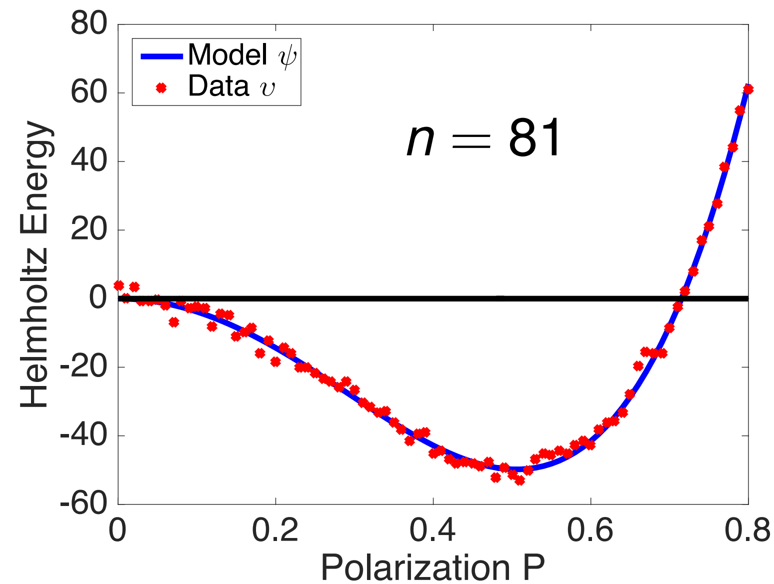


Strategy 1: Perform Experiments

Example: Helmholtz energy $\psi(P,q) = \alpha_1 P^2 + \alpha_{11} P^4$, $q = [\alpha_1, \alpha_{11}]$

Statistical Model: Describes observation process

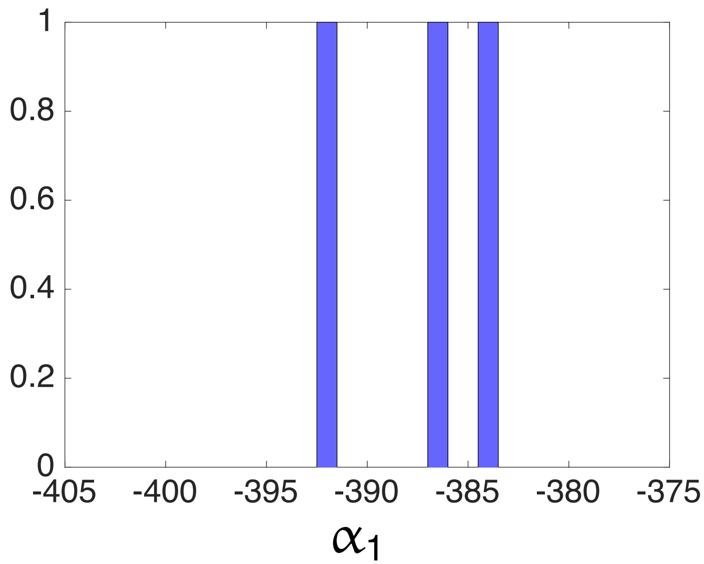
$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \quad , \; i = 1, \ldots, n$$

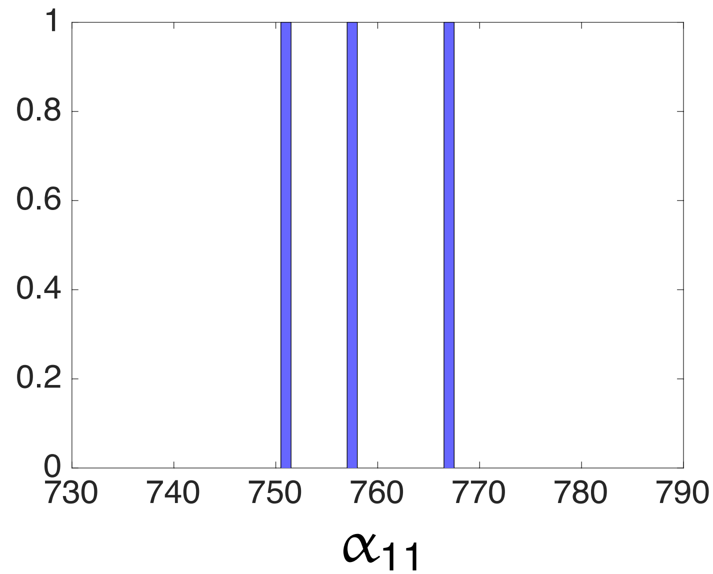


Common Assumption: $\varepsilon_i \sim N(0, \sigma^2)$

UQ Goals: Quantify parameter and response uncertainties

Strategy 1: Perform experiments; e.g., 3





Strategy 1: Perform Experiments

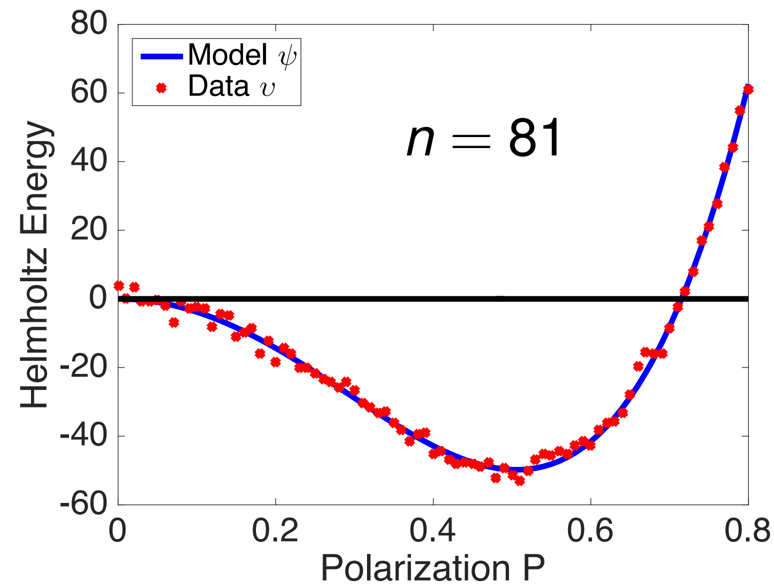
Example: Helmholtz energy $\psi(P,q) = \alpha_1 P^2 + \alpha_{11} P^4$, $q = [\alpha_1, \alpha_{11}]$

Statistical Model: Describes observation process

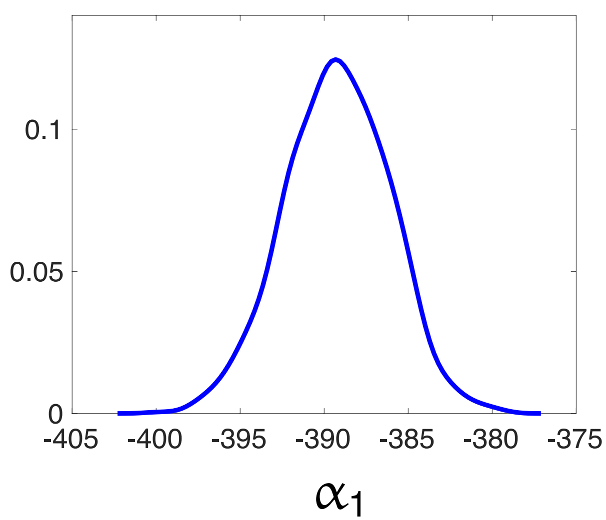
$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \quad , \; i = 1, \dots, n$$

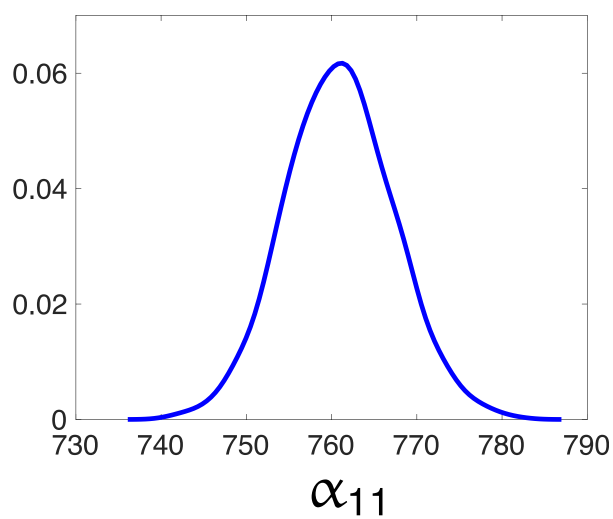
Common Assumption: $\varepsilon_i \sim N(0, \sigma^2)$

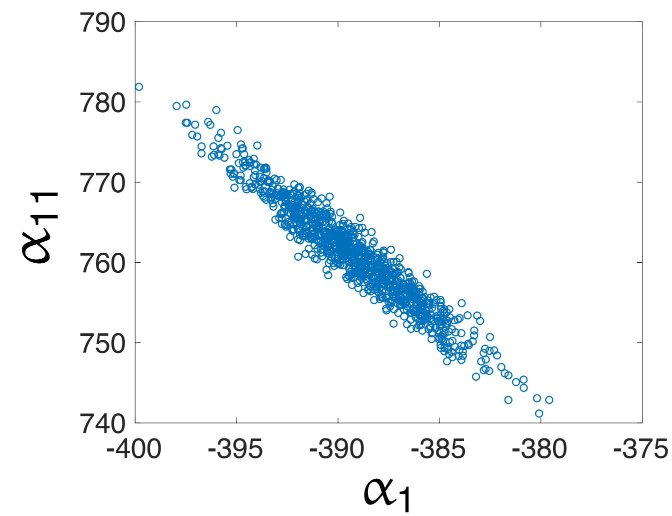
UQ Goals: Quantify parameter and response uncertainties



Strategy 1: Perform many experiments; e.g., 1000





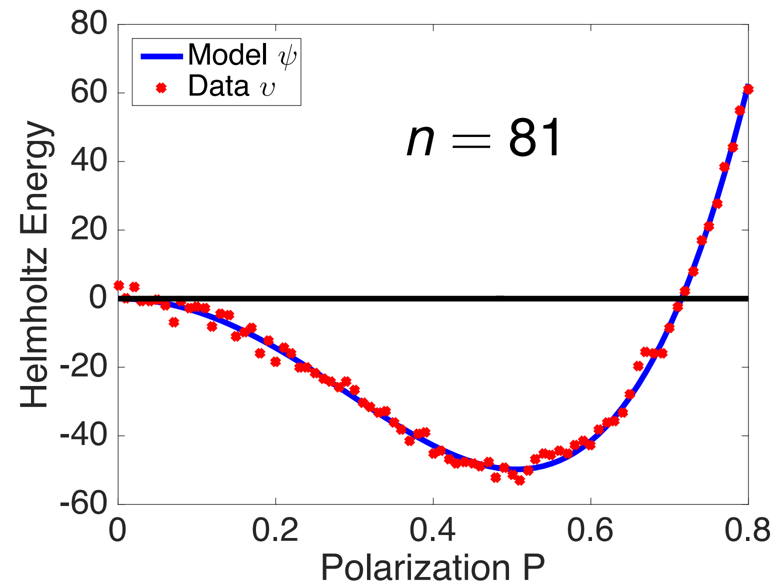


Strategy 1: Perform Experiments

Example: Helmholtz energy $\psi(P,q) = \alpha_1 P^2 + \alpha_{11} P^4$, $q = [\alpha_1, \alpha_{11}]$

Statistical Model: Describes observation process

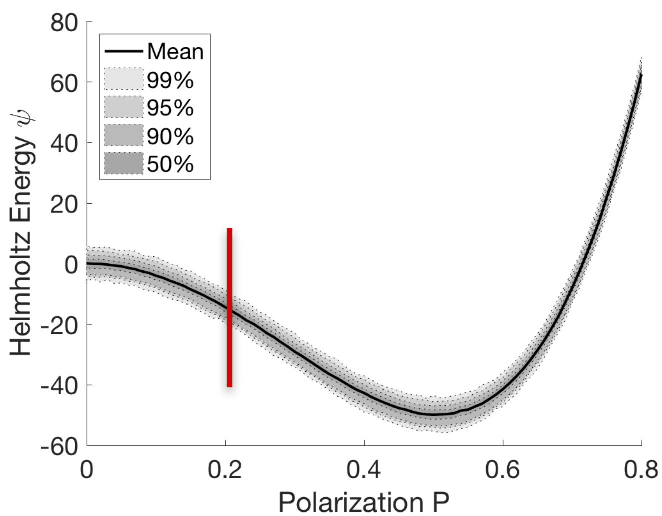
$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \quad , \; i = 1, \ldots, n$$

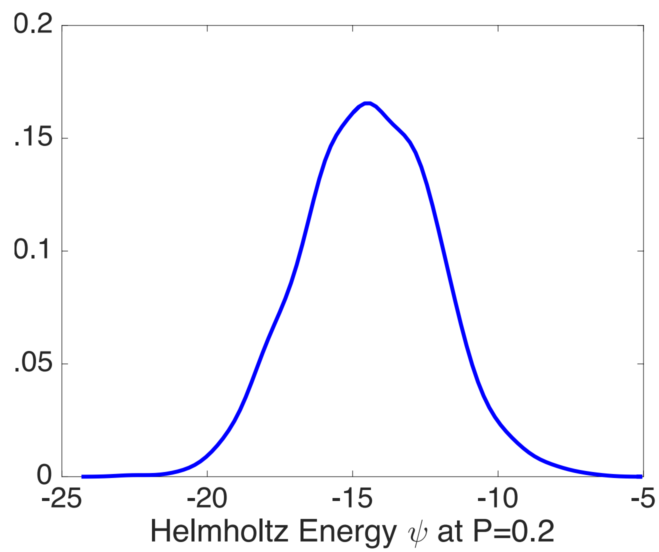


Common Assumption: $\varepsilon_i \sim N(0, \sigma^2)$

UQ Goals: Quantify parameter and response uncertainties

Strategy 1: Perform many experiments; e.g., 1000





Strategy 1: Perform Experiments

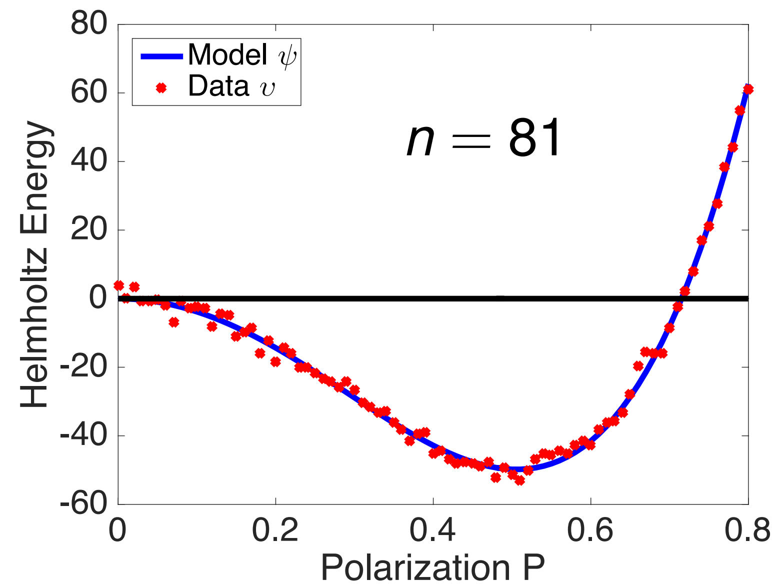
Example: Helmholtz energy $\psi(P,q) = \alpha_1 P^2 + \alpha_{11} P^4$, $q = [\alpha_1, \alpha_{11}]$

Statistical Model: Describes observation process

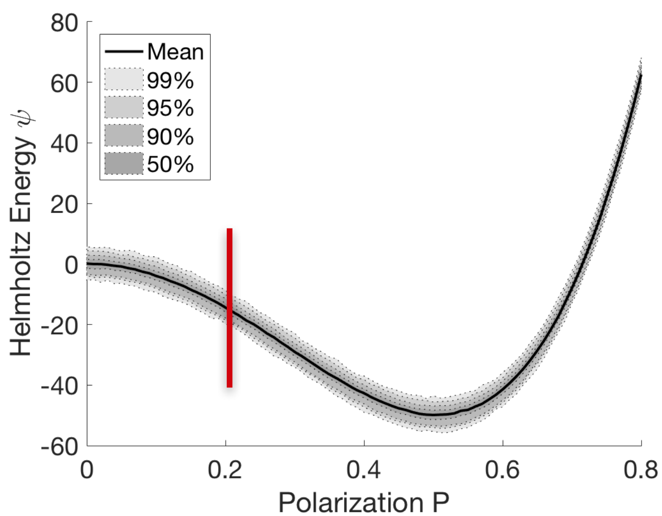
$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \quad , \; i = 1, \ldots, n$$

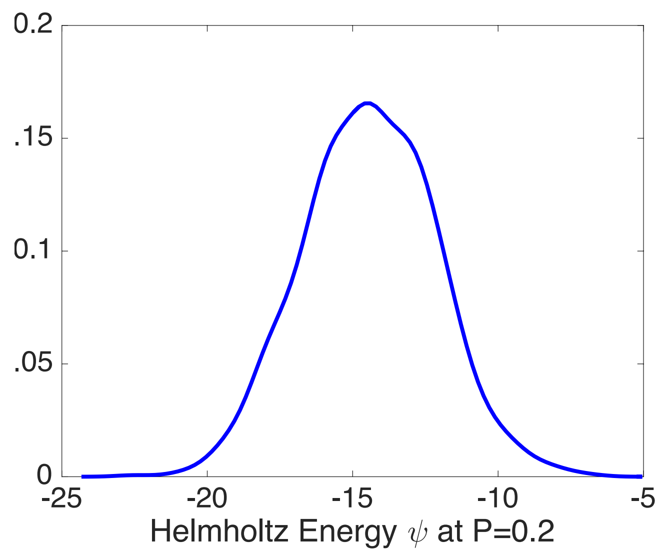
Common Assumption: $\varepsilon_i \sim N(0, \sigma^2)$

UQ Goals: Quantify parameter and response uncertainties



Strategy 1: Perform many experiments; e.g., 1000





Problem: Often cannot perform required number of experiments or high-fidelity simulations.

Solution: Statistical inference

3. Statistical Inference

Goal: The goal in statistical inference is to make conclusions about a phenomenon based on observed data.

Frequentist: Observations made in the past are analyzed with a specified model. Result is regarded as confidence about state of real world.

• Probabilities defined as frequencies with which an event occurs if experiment is repeated several times.

• Parameter Estimation:

o Relies on estimators derived from different data sets and a specific sampling distribution.

o Parameters may be unknown but are fixed and deterministic.

Bayesian: Interpretation of probability is subjective and can be updated with new data.

• Parameter Estimation: Parameters are considered to be random variables having associated densities.

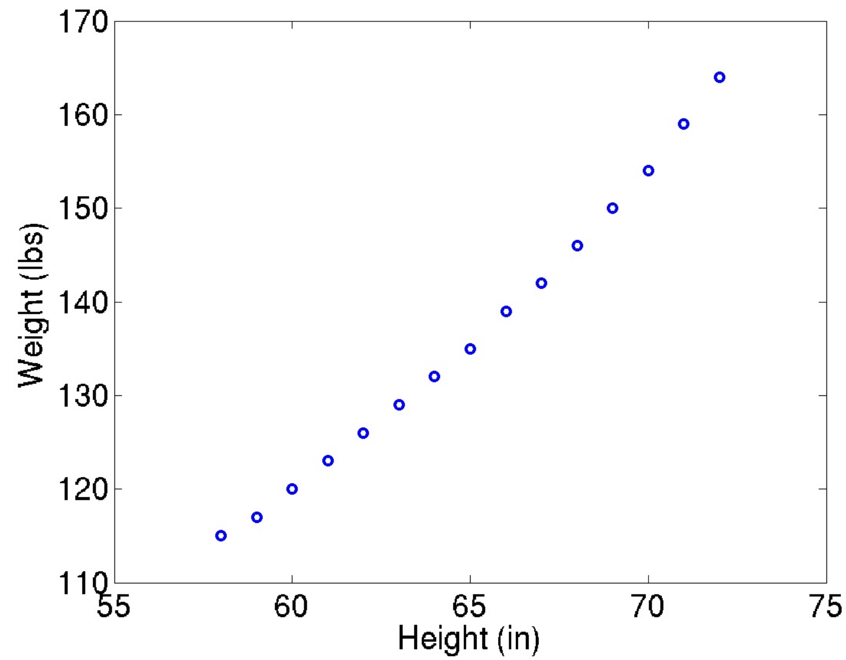
Frequentist Techniques for Model Calibration

Example: Consider the height-weight data from the *1975 World Almanac and Book of Facts*

Height (in)	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72
Weight (lbs)	115	117	120	123	126	129	132	135	139	142	146	150	154	159	164

Consider the model

$$\Upsilon_i = q_1 + q_2(x_i/12) + q_3(x_i/12)^2 + \varepsilon_i$$



Linear Regression

Consider

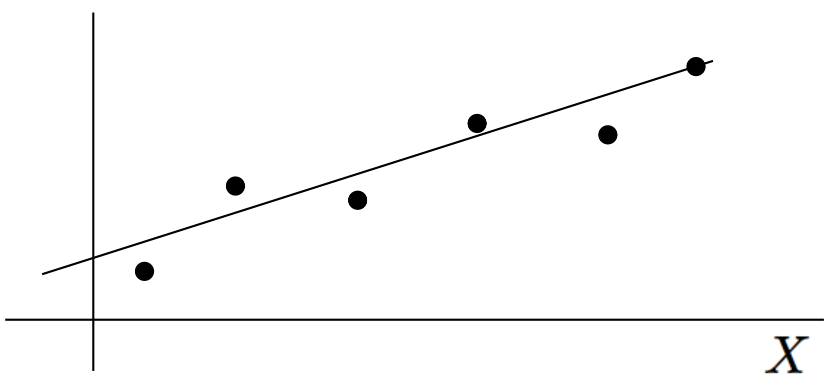
$$\Upsilon = X q_0 + \varepsilon$$

where

$$\Upsilon = \begin{bmatrix} \Upsilon_1 \\ \vdots \\ \Upsilon_n \end{bmatrix}, \; X = \begin{bmatrix} X_{11} & \cdots & X_{1p} \\ \vdots & & \vdots \\ X_{n1} & \cdots & X_{np} \end{bmatrix}, \; q_0 = \begin{bmatrix} q_1 \\ \vdots \\ q_p \end{bmatrix}, \; \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Observations | Design Matrix | Unknown Parameters | Errors

Example: $\Upsilon_i = (q_0 + q_1 X_i) + \varepsilon_i \;, \; i = 1, \cdots, n$



Linear Regression

Statistical Model:

$$\Upsilon = Xq_0 + \varepsilon$$

Assumptions:

(i) $\mathbb{E}(\varepsilon_i) = 0$

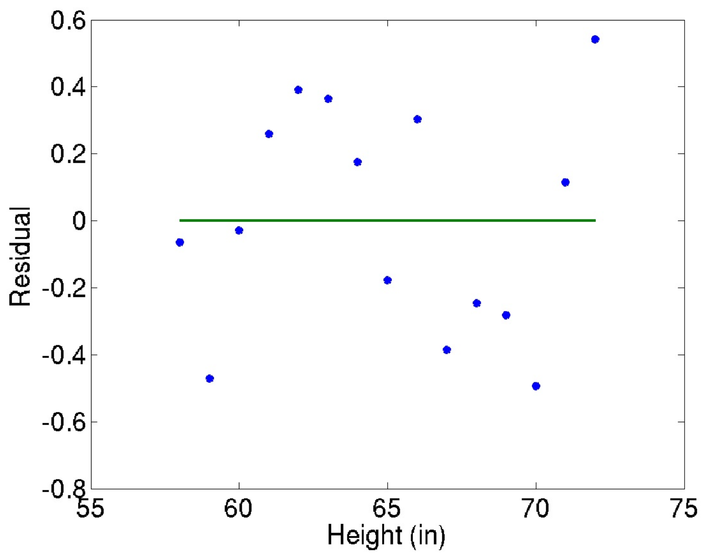
(ii) ε_i iid (independent and identically distributed)

$$\Rightarrow \quad \text{var}(\varepsilon_i) = \sigma_0^2$$

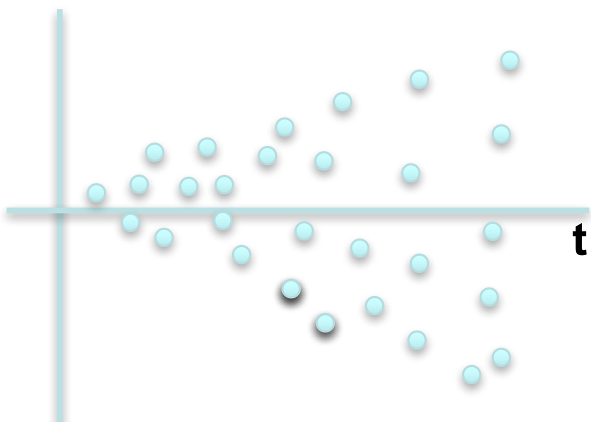
$$\mathbb{E}[(\varepsilon_i - \mathbb{E}(\varepsilon_i))(\varepsilon_j - \mathbb{E}(\varepsilon_j))] = \text{cov}(\varepsilon_i, \varepsilon_j) = 0 \text{ for } i \neq j$$

Examples:

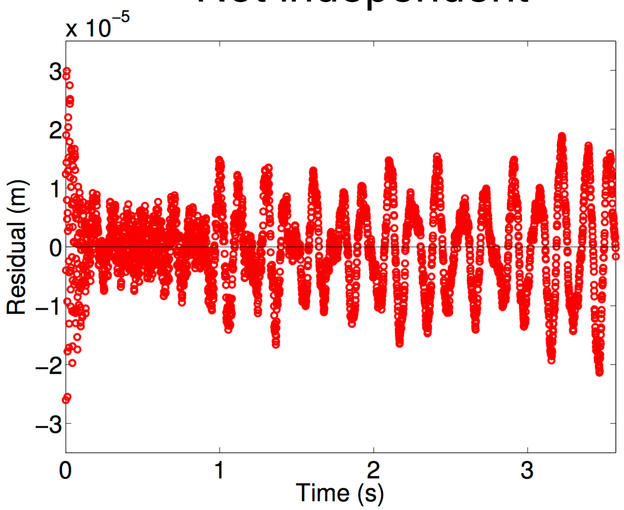
iid errors



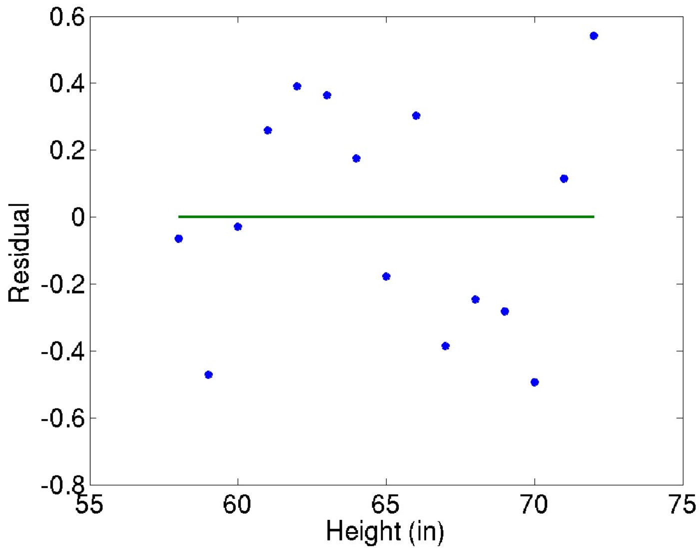
Not identically distributed



Not independent



Linear Regression



Statistical Model:

$$\Upsilon = Xq_0 + \varepsilon$$

Assumptions:

(i) $\mathbb{E}(\varepsilon_i) = 0$

(ii) ε_i iid (independent and identically distributed)

$$\Rightarrow \quad \text{var}(\varepsilon_i) = \sigma_0^2$$

$$\mathbb{E}[(\varepsilon_i - \mathbb{E}(\varepsilon_i))(\varepsilon_j - \mathbb{E}(\varepsilon_j))] = \text{cov}(\varepsilon_i, \varepsilon_j) = 0 \text{ for } i \neq j$$

Goals:

(1) Construct a 'good' estimator $\hat{q}$ for q.

(2) Construct an estimator $\hat{\sigma}^2$ for σ_0^2.

Terminology:

- Estimator: Random variable having associated sampling distributions
- Estimate: Realization so real number

Least Squares Problem

Minimize

$$\mathcal{J}(q) = (\Upsilon - Xq)^T(\Upsilon - Xq)$$

Note:

$$\nabla_q \mathcal{J} = 2[\nabla_q(\Upsilon - Xq)^T][\Upsilon - Xq] = 0$$

where

$$\nabla_q(\Upsilon - Xq)^T = -\nabla_q q^T X^T = -X^T$$

Least Squares Estimator: $\hat{q}_{OLS} = (X^T X)^{-1} X^T \Upsilon$

Least Squares Estimate: $q_{OLS} = (X^T X)^{-1} X^T \upsilon$

Parameter Estimator Properties

Estimator Mean:

$$\begin{aligned} \mathbb{E}(\hat{q}) &= \mathbb{E}\left[(X^TX)^{-1}X^T\Upsilon\right] \\ &= (X^TX)^{-1}X^T\mathbb{E}(\Upsilon) \\ &= q_0 \end{aligned}$$

$$\Upsilon = Xq_0 + \varepsilon$$

Estimator Covariance: Let $A = (X^TX)^{-1}X^T$

$$\begin{aligned} V(\hat{q}) = &\ \mathbb{E}[(\hat{q} - q_0)(\hat{q} - q_0)^T] \\ = &\ \mathbb{E}[(q_0 + A\varepsilon - q_0)(q_0 + A\varepsilon - q_0)^T]\ ,\ \text{since } \hat{q} = A\Upsilon = A(Xq_0 + \varepsilon) \\ = &\ A\mathbb{E}(\varepsilon\varepsilon^T)A^T \\ = &\ \sigma_0^2(X^TX)^{-1} \end{aligned}$$

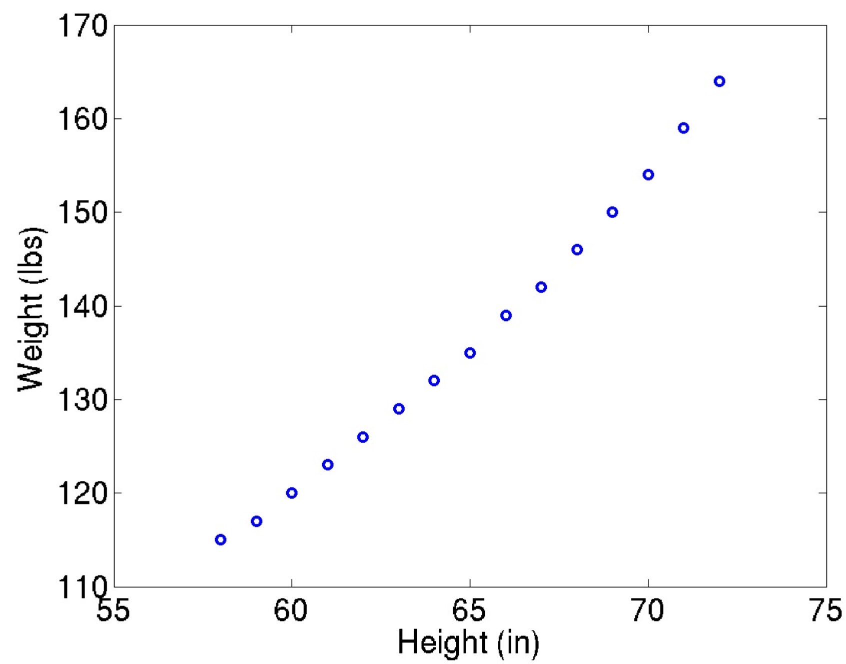
Example

Example: Consider the height-weight data from the *1975 World Almanac and Book of Facts*

Height (in)	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72
Weight (lbs)	115	117	120	123	126	129	132	135	139	142	146	150	154	159	164

Consider the model

$$\Upsilon_i = q_1 + q_2(x_i/12) + q_3(x_i/12)^2 + \varepsilon_i$$



Example

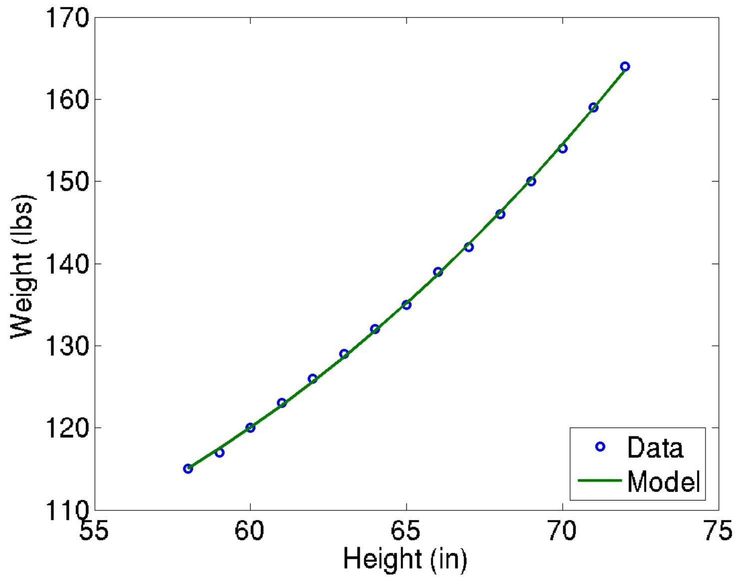
Here

$$X = \begin{bmatrix} 1 & 4.83 & 112.91 \\ 1 & 4.92 & 118.85 \\ 1 & 5.00 & 125.00 \\ 1 & 5.08 & 131.35 \\ 1 & 5.17 & 137.92 \\ 1 & 5.25 & 144.70 \\ 1 & 5.33 & 151.70 \\ 1 & 5.42 & 158.93 \\ 1 & 5.50 & 166.38 \\ 1 & 5.58 & 174.05 \\ 1 & 5.67 & 181.96 \\ 1 & 5.75 & 190.11 \\ 1 & 5.83 & 198.50 \\ 1 & 5.92 & 207.12 \\ 1 & 6.00 & 216.00 \end{bmatrix}$$

Least Square Estimate

$$(X^T X)q = X^T v$$

$$\Rightarrow \begin{aligned} q_1 &= 261.88 \\ q_2 &= -88.18 \\ q_3 &= 11.96 \end{aligned}$$



Variance Estimate:

$$\sigma^2 = 0.15$$

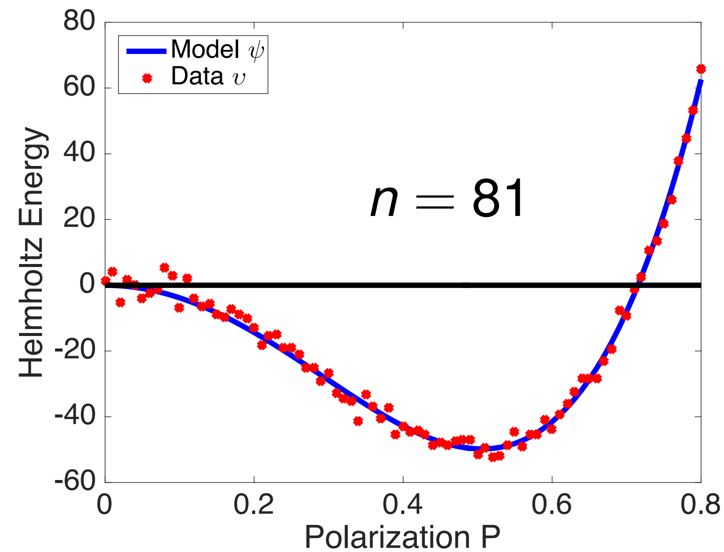
Parameter Covariance Estimate:

$$V = \begin{bmatrix} 634.88 & -235.04 & 21.66 \\ -235.04 & 87.09 & -8.03 \\ 21.66 & -8.03 & 0.74 \end{bmatrix}$$

Note: This yields variances and standard deviations for parameter estimates

$$\begin{aligned} q_1 &= 261.88 \pm 50.39 \\ q_2 &= -88.18 \pm 18.66 \\ q_3 &= 11.96 \pm 1.72 \end{aligned} \quad \Rightarrow \quad \begin{aligned} q_1 &\in [211.48, 312.27] \\ q_2 &\in [-106.84, -69.51] \\ q_3 &\in [10.24, 13.68]. \end{aligned}$$

Polarization Example



Statistical Model: For i = 1, …, n

$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \longleftarrow \varepsilon_i \sim N(0, \sigma^2)$$

$$= \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \varepsilon_i$$

$$\Rightarrow \left[\upsilon_i\right] = \left[P_i^2 \;\; P_i^4\right] \begin{bmatrix} \alpha_1 \\ \alpha_{11} \end{bmatrix} + \left[\varepsilon_i\right]$$

$$\Rightarrow \upsilon = Xq + \varepsilon$$

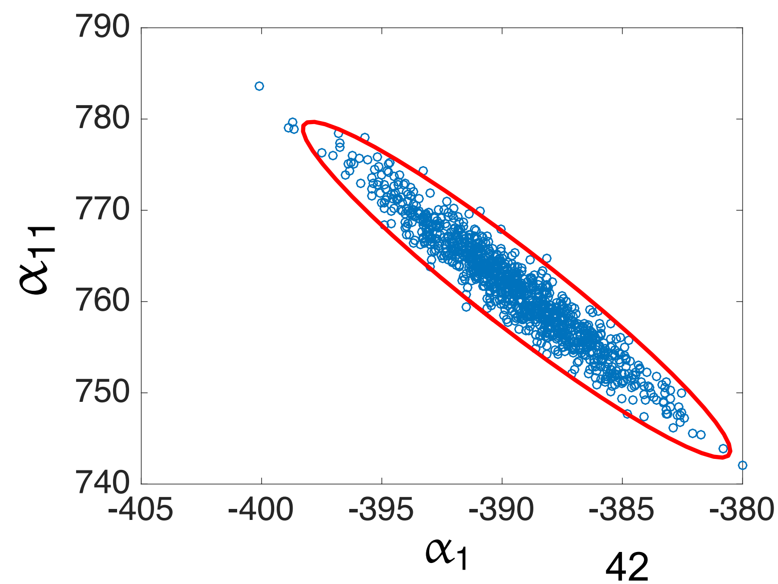
Statistical Quantities:

$$q = (X^T X)^{-1} X^T \upsilon$$

$$\mathrm{var}(\alpha_1)$$

$$V = \underline{\sigma^2}(X^T X)^{-1} = \begin{bmatrix} 8.8 & -17.4 \\ -17.4 & 37.6 \end{bmatrix}$$

$$\mathrm{cov}(\alpha_1, \alpha_{11}) \qquad \mathrm{var}(\alpha_{11})$$



Note: Covariance matrix incorporates “geometry”

Goal: Employ Bayesian inference for UQ

Statistical Inference

Goal: The goal in statistical inference is to make conclusions about a phenomenon based on observed data.

Frequentist: Observations made in the past are analyzed with a specified model. Result is regarded as confidence about state of real world.

• Probabilities defined as frequencies with which an event occurs if experiment is repeated several times.

• Parameter Estimation:

o Relies on estimators derived from different data sets and a specific sampling distribution.

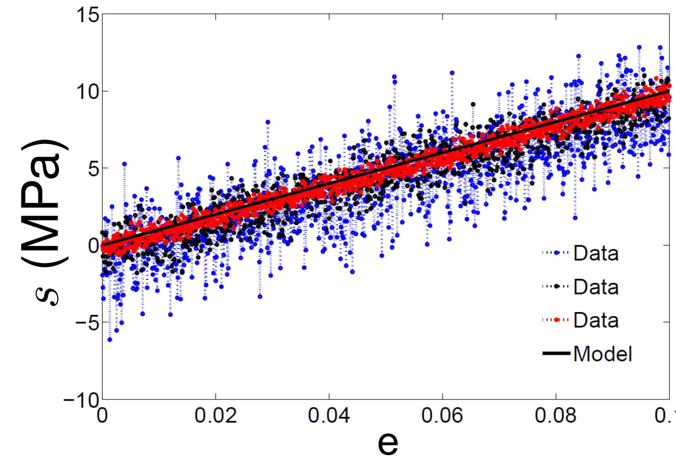
o Parameters may be unknown but are fixed and deterministic.

Bayesian: Interpretation of probability is subjective and can be updated with new data.

• Parameter Estimation: Parameters are considered to be random variables having associated densities.

Bayesian Inference: More General Model

Example: Displacement-force relation (Hooke's Law)

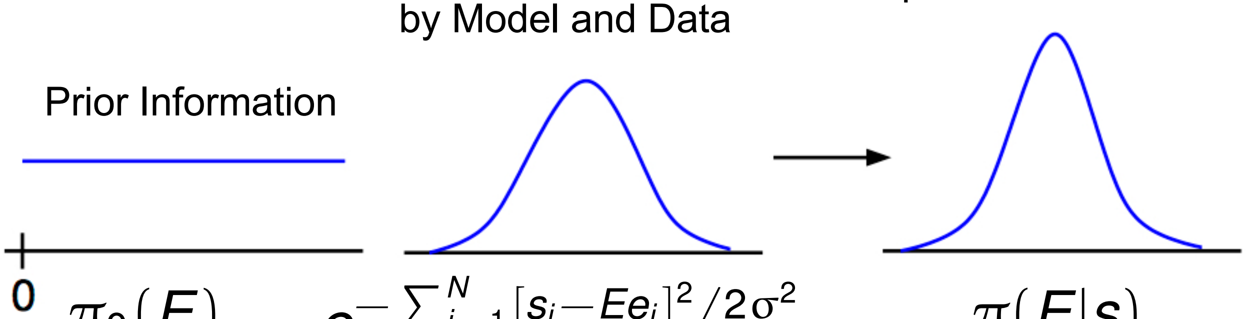


$$s_i = Ee_i + \varepsilon_i \ , \ i = 1, \ldots, N$$

$$\varepsilon_i \sim N(0, \sigma^2)$$

Parameter: Stiffness E

Strategy: Use model fit to data to update prior information



$$\pi_0(E) \qquad e^{-\sum_{i=1}^N [s_i - Ee_i]^2/2\sigma^2} \qquad \pi(E|s)$$

Data Model

Non-normalized Bayes' Relation:

$$\pi(E|s) = e^{-\sum_{i=1}^N [s_i - Ee_i]^2/2\sigma^2} \pi_0(E)$$

Bayesian Inference

Bayes' Relation: Specifies posterior in terms of likelihood and prior

Likelihood: $e^{-\sum_{i=1}^{N}[s_i - Ee_i]^2/2\sigma^2}$, $q = E$

$\upsilon = [s_1, \ldots, s_N]$

Posterior Distribution

$$\boxed{\pi(q|\upsilon) = \frac{\pi(\upsilon|q)\pi_0(q)}{\int_{\mathbb{R}^p}\pi(\upsilon|q)\pi_0(q)dq}}$$

Prior Distribution

Normalization Constant

- Prior Distribution: Quantifies prior knowledge of parameter values
- Likelihood: Probability of observing a data given set of parameter values.
- Posterior Distribution: Conditional distribution of parameters given observed data.

Problem: Can require high-dimensional integration

• e.g., HIV Model: $p = 6 - 23$!

• Solution: Sampling-based Markov Chain Monte Carlo (MCMC) algorithms.

• Metropolis algorithms first used by nuclear physicists during Manhattan Project in 1940's to understand particle movement underlying first atomic bomb.

Bayesian Model Calibration

Bayes' Relation:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Bayesian Model Calibration:

- Parameters assumed to be random variables

$$\pi(q|\upsilon) = \frac{\pi(\upsilon|q)\pi_0(q)}{\int_{\mathbb{R}^p} \pi(\upsilon|q)\pi_0(q)dq}$$

Example: Coin Flip

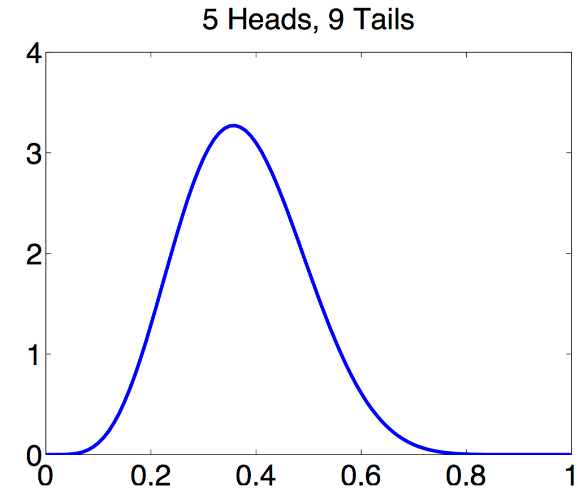
$$\Upsilon_i(\omega) = \begin{cases} 0 & , \quad \omega = T \\ 1 & , \quad \omega = H \end{cases}$$

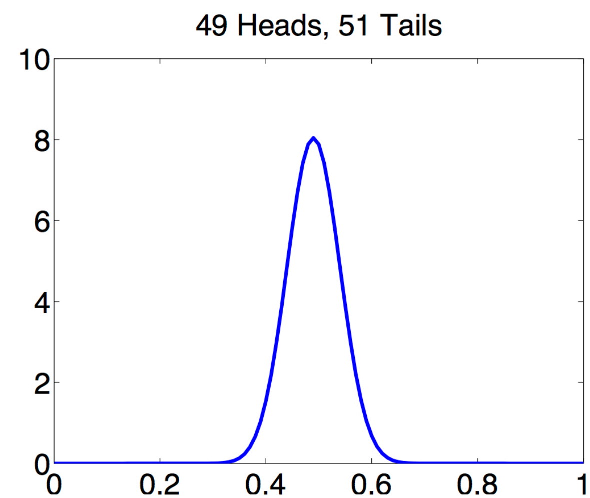
Likelihood:

$$\begin{aligned} \pi(\upsilon|q) &= \prod_{i=1}^{N} q^{\upsilon_i}(1-q)^{1-\upsilon_i} \\ &= q^{N_1}(1-q)^{N_0} \end{aligned}$$

Posterior with Noninformative Prior: $\pi_0(q) = 1$

$$\pi(q|\upsilon) = \frac{q^{N_1}(1-q)^{N_0}}{\int_0^1 q^{N_1}(1-q)^{N_0}dq} = \frac{(N+1)!}{N_0!N_1!} q^{N_1}(1-q)^{N_0}$$



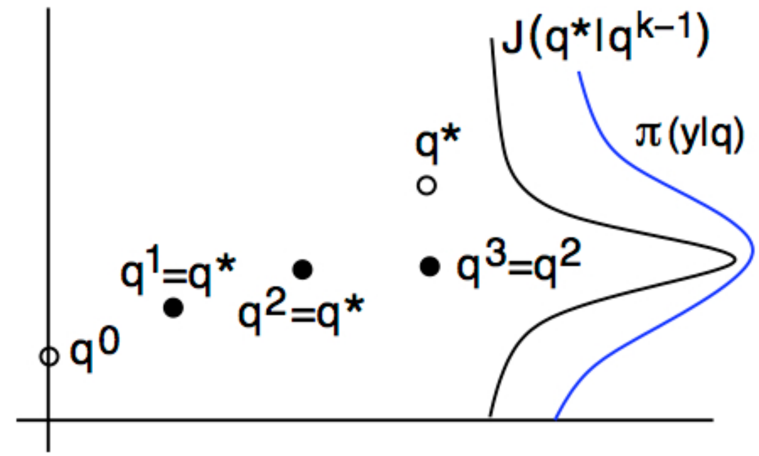


Bayesian Model Calibration

Bayesian Model Calibration:

• Parameters considered to be random variables with associated densities.

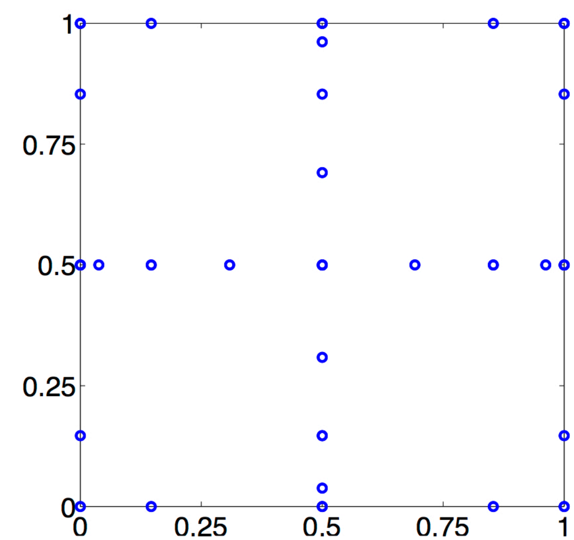
$$\pi(q|\upsilon) = \frac{\pi(\upsilon|q)\pi_0(q)}{\int_{\mathbb{R}^p} \pi(\upsilon|q)\pi_0(q)dq}$$



Problem:

•Often requires high dimensional integration;

- e.g., p = 23 for HIV model
- p = hundreds to thousands for some models



Strategies:

• Sampling methods

• Sparse grid quadrature techniques

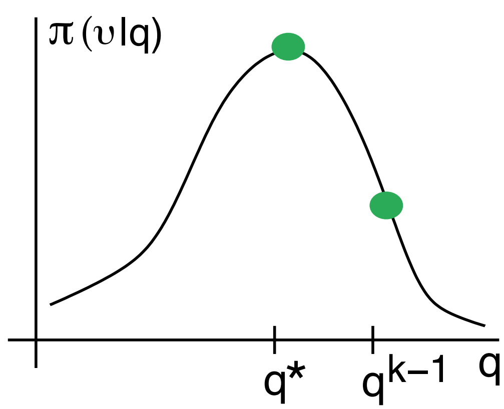
Markov Chain Monte Carlo Methods

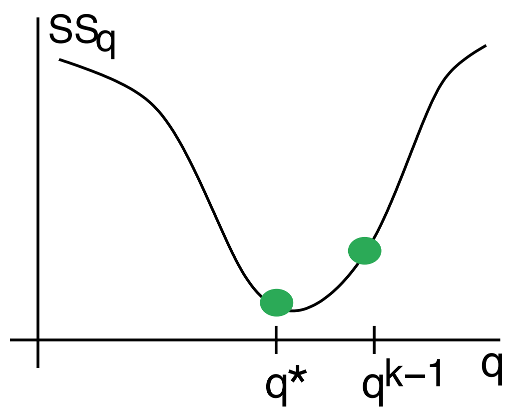
Strategy:

- Sample values from proposal distribution $J(q^*|q^{k-1})$ that reflects geometry of posterior distribution
- Compute $r(q^*|q^{k-1}) = \frac{\pi(\upsilon|q^*)\pi_0(q^*)}{\pi(\upsilon|q^{k-1})\pi_0(q^{k-1})}$
  - If $r \geqslant 1$, accept with probability $\alpha = 1$
  - If $r < 1$, accept with probability $\alpha = r$

Intuition: Consider flat prior $\pi_0(q) = 1$ and Gaussian observation model

$$\pi(\upsilon|q) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-SS_q/2\sigma^2} \qquad SS_q = \sum_{i=1}^{N} [\upsilon_i - f(t_i, q)]^2$$





Delayed Rejection Adaptive Metropolis (DRAM)

Algorithm: [Haario et al., 2006] – MATLAB, Python, R

Example: Helmholtz energy

$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \longleftarrow \varepsilon_i \sim N(0, \sigma^2)$$
$$= \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \varepsilon_i$$

Delayed Rejection Adaptive Metropolis (DRAM)

Algorithm: [Haario et al., 2006] – MATLAB, Python, R

1. Determine $q^0 = \arg\min_q \sum_{i=1}^{N} [\upsilon_i - \psi(P_i, q)]^2]$

Example: Helmholtz energy

$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \longleftarrow \varepsilon_i \sim N(0, \sigma^2)$$
$$= \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \varepsilon_i$$

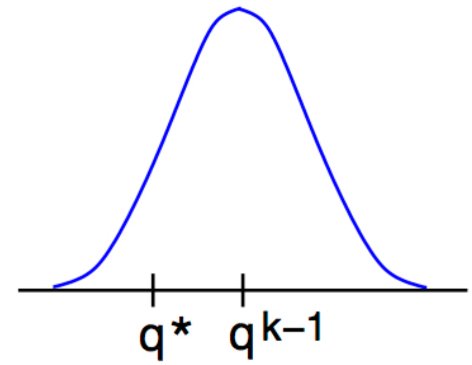
Delayed Rejection Adaptive Metropolis (DRAM)

Algorithm: [Haario et al., 2006] – MATLAB, Python, R

1. Determine $q^0 = \arg\min_q \sum_{i=1}^{N} [\upsilon_i - \psi(P_i, q)]^2]$

2. For $k = 1, \ldots, M$

   (a) Construct candidate $q^* \sim N(q^{k-1}, \underline{V})$

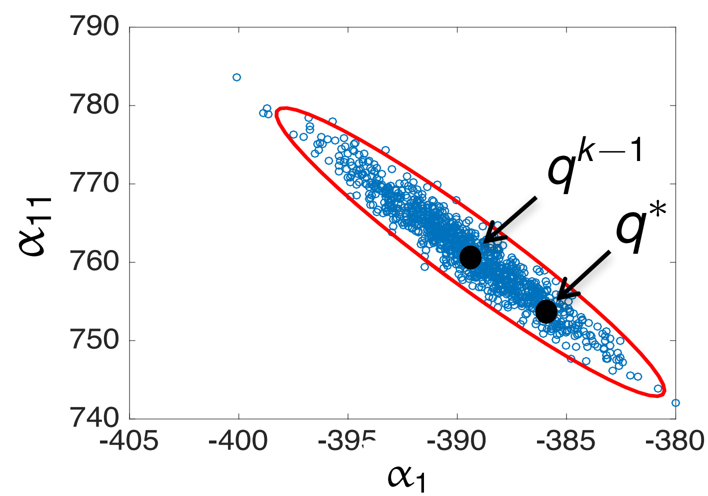


Example: Helmholtz energy

$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \longleftarrow \varepsilon_i \sim N(0, \sigma^2)$$

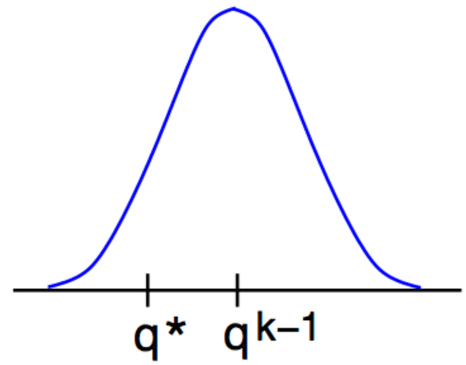
$$= \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \varepsilon_i$$

Recall: Covariance V incorporates geometry



Delayed Rejection Adaptive Metropolis (DRAM)

Algorithm: [Haario et al., 2006] – MATLAB, Python, R



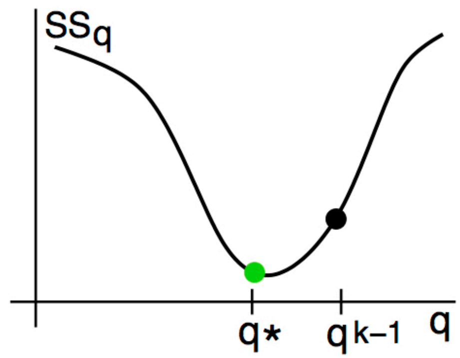
1. Determine $q^0 = \arg\min_q \sum_{i=1}^{N} [\upsilon_i - \psi(P_i, q)]^2]$
2. For $k = 1, \ldots, M$

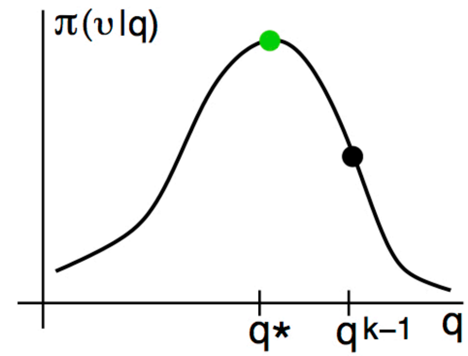
   (a) Construct candidate $q^* \sim N(q^{k-1}, V)$

   (b) Compute likelihood

$$SS_{q^*} = \sum_{i=1}^{N} \upsilon_i - \psi(P_i, q^*)]^2$$

$$\pi(\upsilon|q) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-SS_q/2\sigma^2}$$





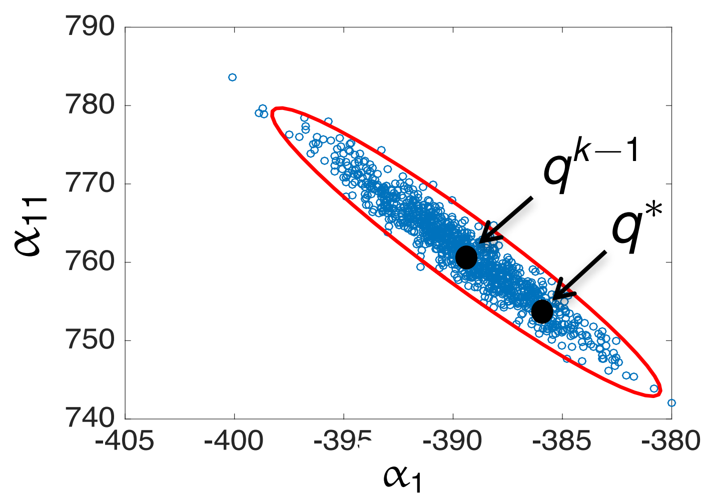
   (c) Accept q^* with probability dictated by likelihood

Example: Helmholtz energy

$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \longleftarrow \varepsilon_i \sim N(0, \sigma^2)$$

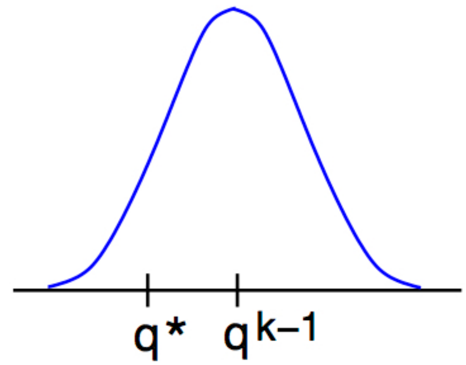
$$= \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \varepsilon_i$$

Recall: Covariance V incorporates geometry



Delayed Rejection Adaptive Metropolis (DRAM)

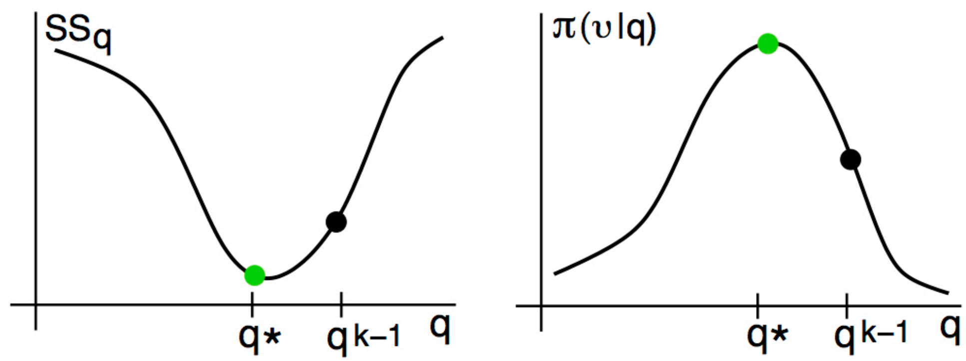
Algorithm: [Haario et al., 2006] – MATLAB, Python, R



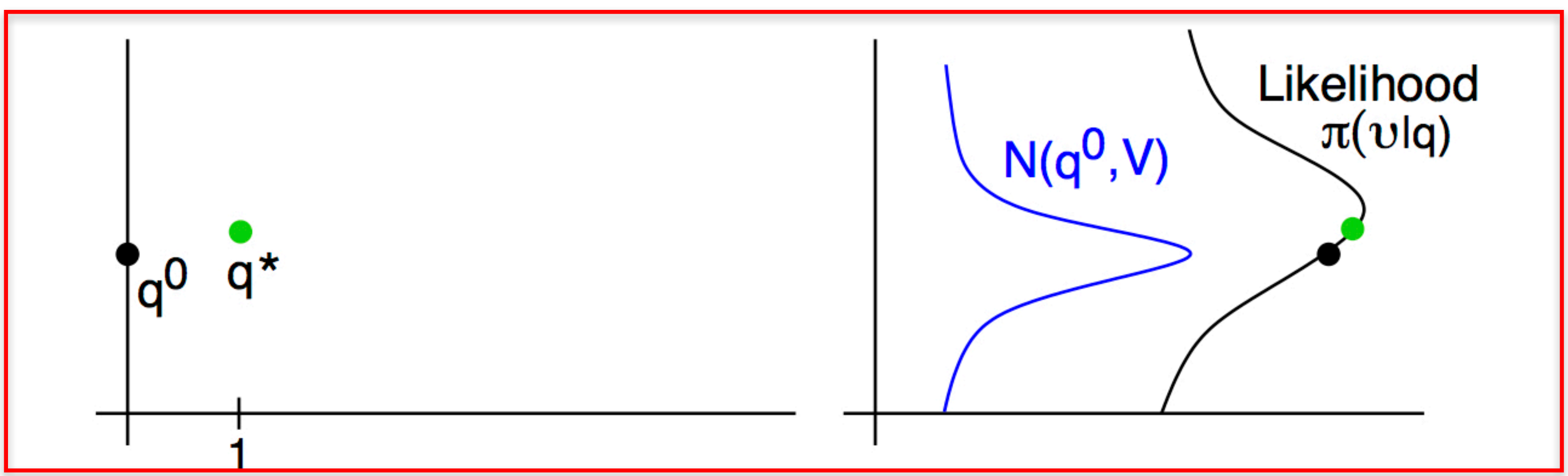
1. Determine $q^0 = \arg\min_q \sum_{i=1}^{N} [\upsilon_i - \psi(P_i, q)]^2]$
2. For $k = 1, \ldots, M$
   (a) Construct candidate $q^* \sim N(q^{k-1}, V)$
   (b) Compute likelihood

$$SS_{q^*} = \sum_{i=1}^{N} \upsilon_i - \psi(P_i, q^*)]^2$$

$$\pi(\upsilon|q) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-SS_q/2\sigma^2}$$

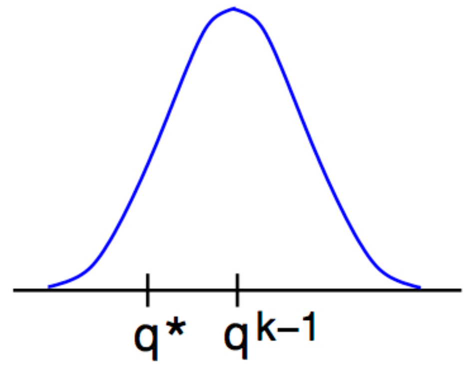


   (c) Accept q^* with probability dictated by likelihood



Delayed Rejection Adaptive Metropolis (DRAM)

Algorithm: [Haario et al., 2006] – MATLAB, Python, R



1. Determine $q^0 = \arg\min_q \sum_{i=1}^{N} [\upsilon_i - \psi(P_i, q)]^2]$

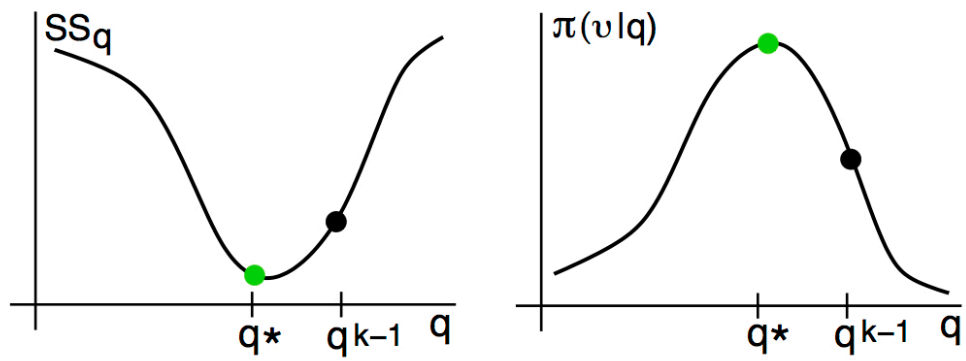
2. For $k = 1, \ldots, M$

   (a) Construct candidate $q^* \sim N(q^{k-1}, V)$

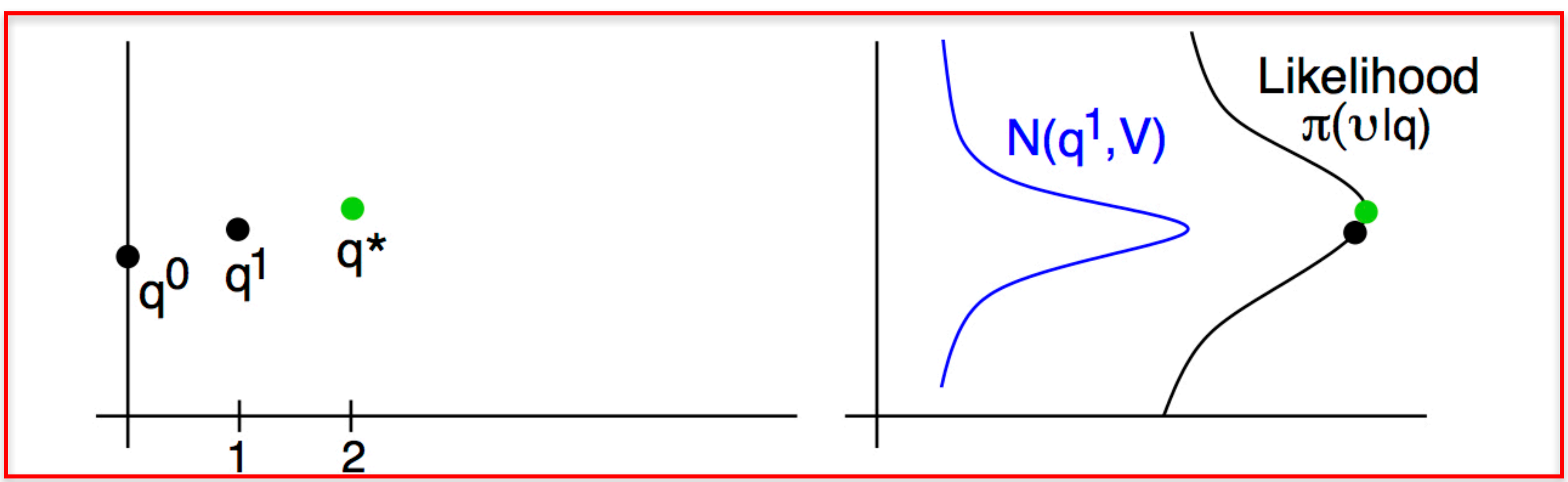
   (b) Compute likelihood

$$SS_{q^*} = \sum_{i=1}^{N} \upsilon_i - \psi(P_i, q^*)]^2$$

$$\pi(\upsilon|q) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-SS_q/2\sigma^2}$$

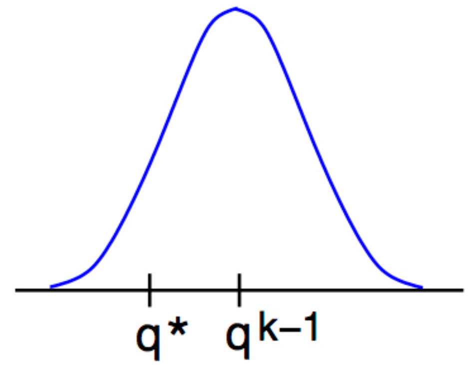


   (c) Accept q^* with probability dictated by likelihood



Delayed Rejection Adaptive Metropolis (DRAM)

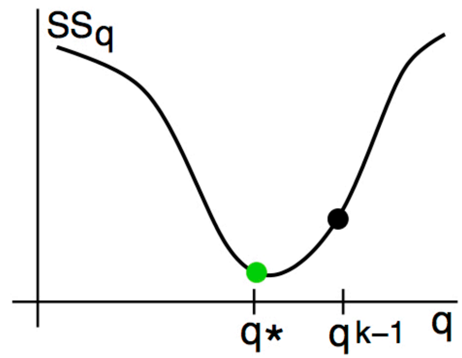
Algorithm: [Haario et al., 2006] – MATLAB, Python, R

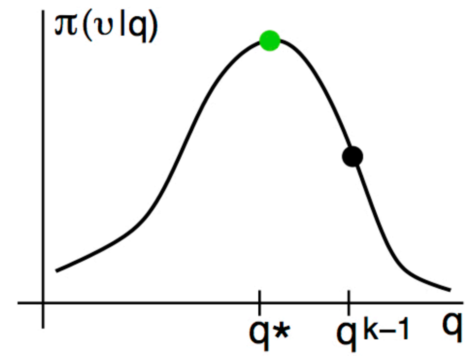


1. Determine $q^0 = \arg\min_q \sum_{i=1}^{N} [\upsilon_i - \psi(P_i, q)]^2]$
2. For $k = 1, \ldots, M$
   (a) Construct candidate $q^* \sim N(q^{k-1}, V)$
   (b) Compute likelihood

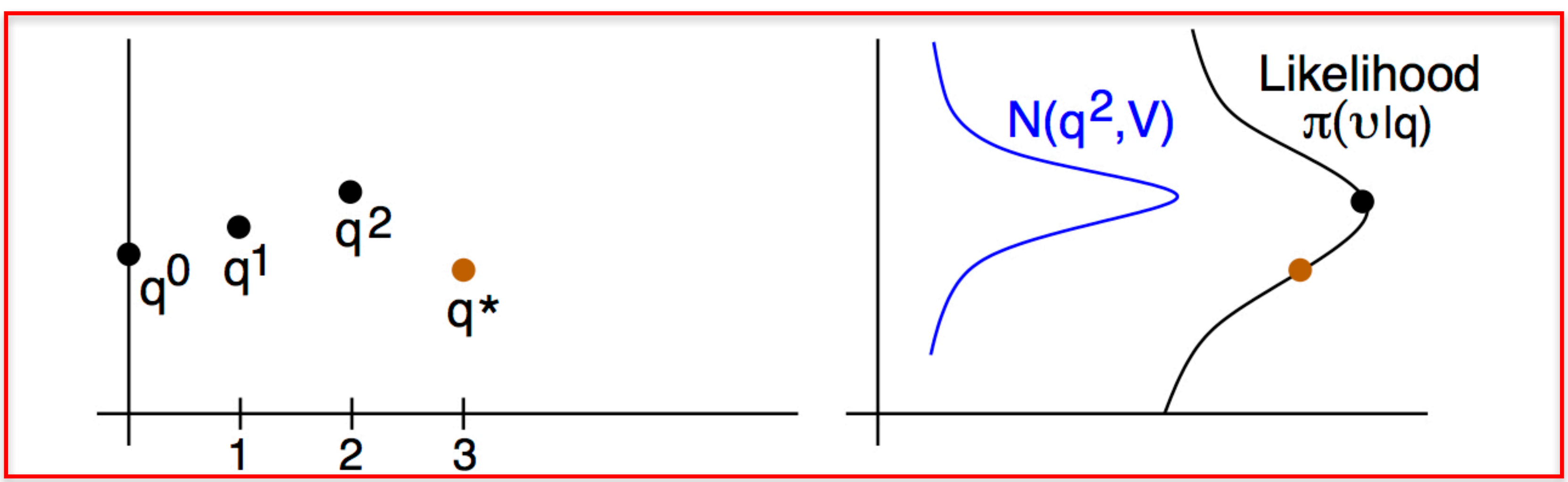
$$SS_{q^*} = \sum_{i=1}^{N} \upsilon_i - \psi(P_i, q^*)]^2$$

$$\pi(\upsilon|q) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-SS_q/2\sigma^2}$$





   (c) Accept q^* with probability dictated by likelihood



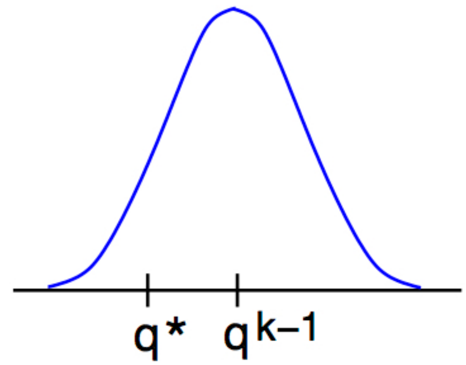
Delayed Rejection Adaptive Metropolis (DRAM)

Algorithm: [Haario et al., 2006] – MATLAB, Python, R

1. Determine $q^0 = \arg\min_q \sum_{i=1}^{N} [\upsilon_i - \psi(P_i, q)]^2]$

2. For $k = 1, \ldots, M$

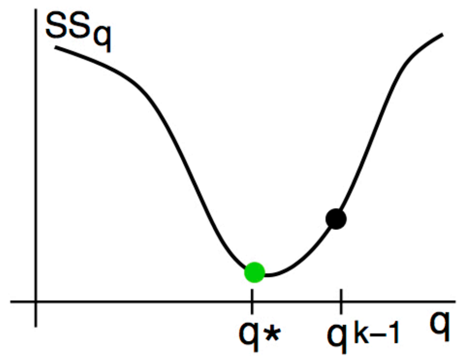
   (a) Construct candidate $q^* \sim N(q^{k-1}, V)$

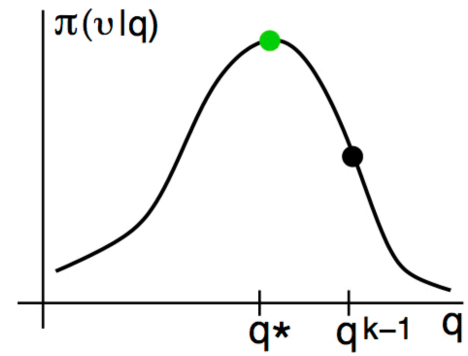


   (b) Compute likelihood

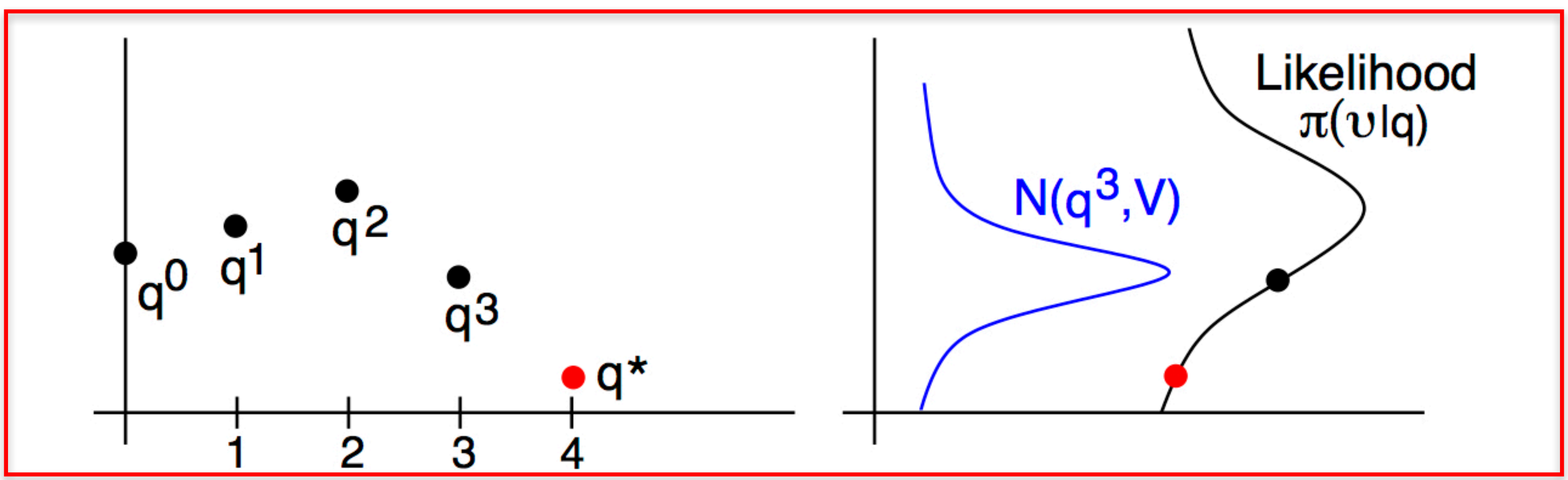
$$SS_{q^*} = \sum_{i=1}^{N} \upsilon_i - \psi(P_i, q^*)]^2$$

$$\pi(\upsilon|q) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-SS_q/2\sigma^2}$$



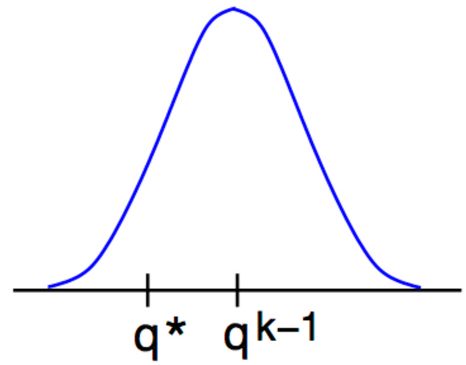


   (c) Accept q^* with probability dictated by likelihood



Delayed Rejection Adaptive Metropolis (DRAM)

Algorithm: [Haario et al., 2006] – MATLAB, Python, R



1. Determine $q^0 = \arg\min_q \sum_{i=1}^{N} [\upsilon_i - \psi(P_i, q)]^2]$

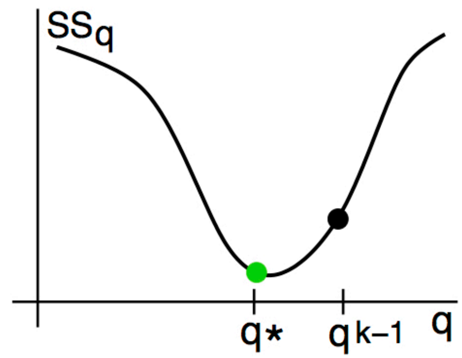
2. For $k = 1, \ldots, M$

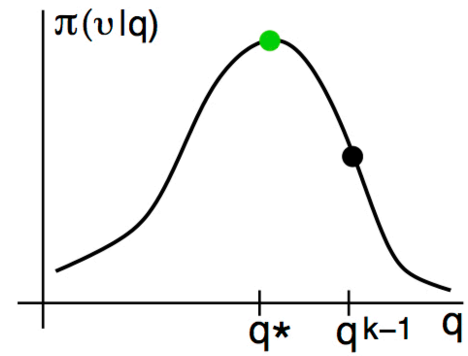
   (a) Construct candidate $q^* \sim N(q^{k-1}, V)$

   (b) Compute likelihood

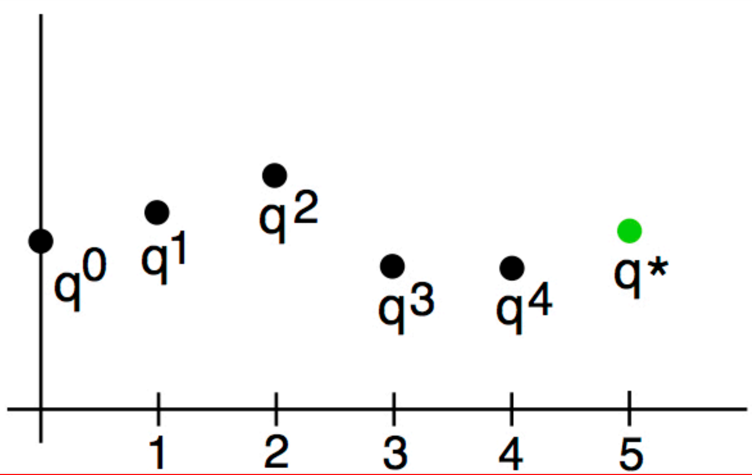
$$SS_{q^*} = \sum_{i=1}^{N} \upsilon_i - \psi(P_i, q^*)]^2$$

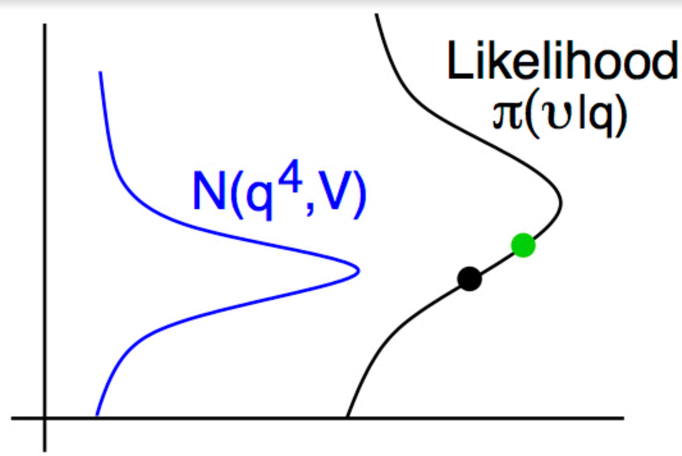
$$\pi(\upsilon|q) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-SS_q/2\sigma^2}$$





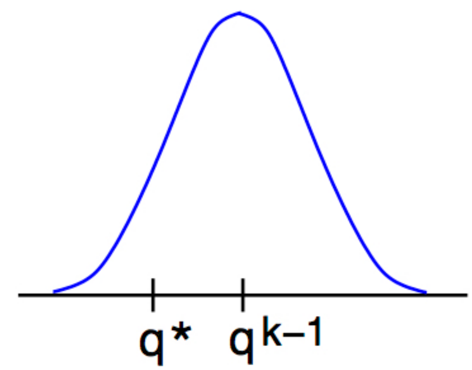
   (c) Accept q^* with probability dictated by likelihood





Delayed Rejection Adaptive Metropolis (DRAM)

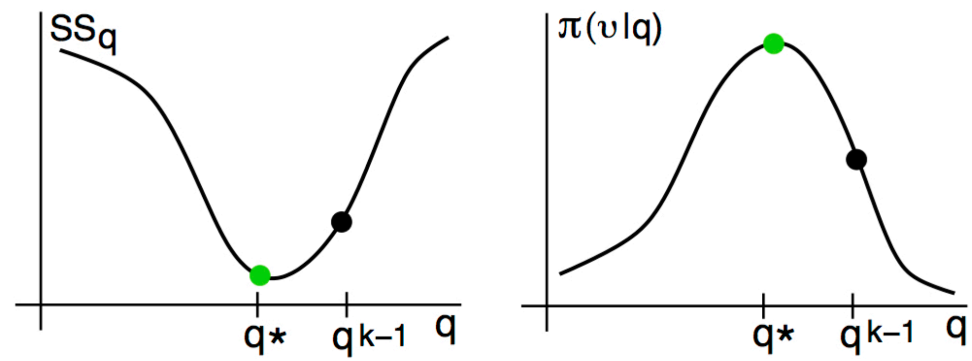
Algorithm: [Haario et al., 2006] – MATLAB, Python, R



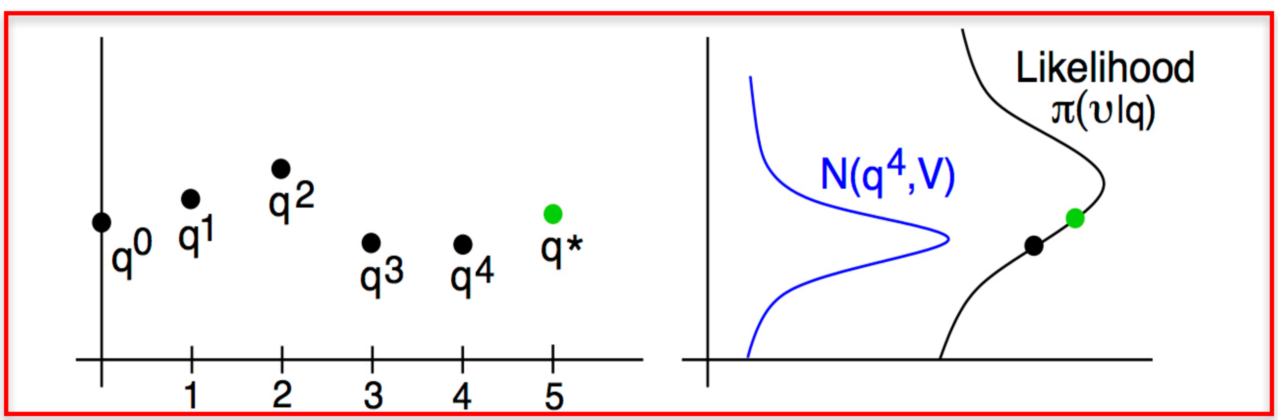
1. Determine $q^0 = \arg\min_q \sum_{i=1}^{N} [\upsilon_i - \psi(P_i, q)]^2]$
2. For $k = 1, \ldots, M$
   (a) Construct candidate $q^* \sim N(q^{k-1}, V)$
   (b) Compute likelihood

$$SS_{q^*} = \sum_{i=1}^{N} \upsilon_i - \psi(P_i, q^*)]^2$$

$$\pi(\upsilon|q) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-SS_q/2\sigma^2}$$



   (c) Accept q^* with probability dictated by likelihood



Note:

- Delayed Rejection: Shrink proposal: γV
- Adaptive Metropolis: Update proposal as samples are accepted

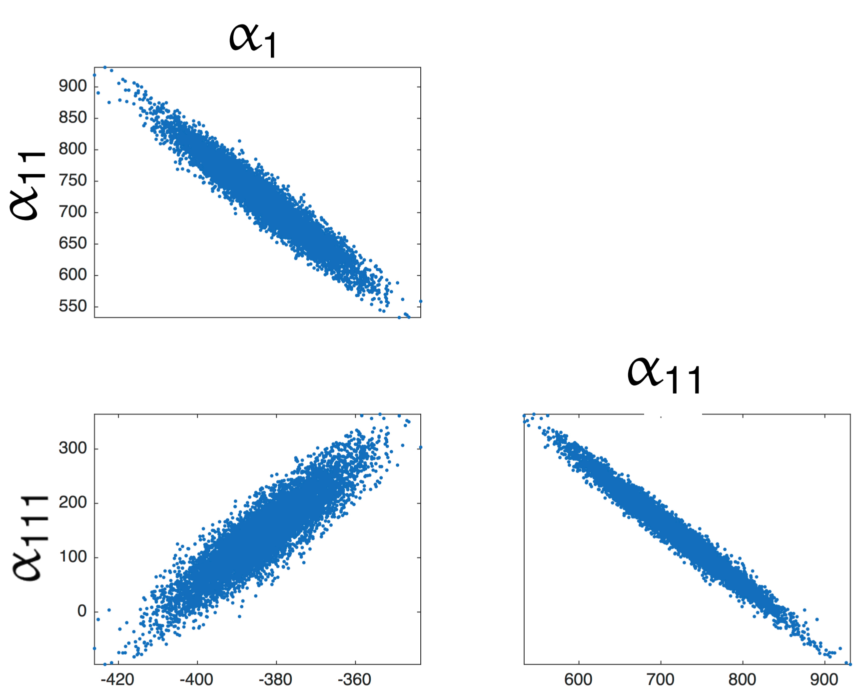
Delayed Rejection Adaptive Metropolis (DRAM)

Example: Helmholtz energy with 3 parameters

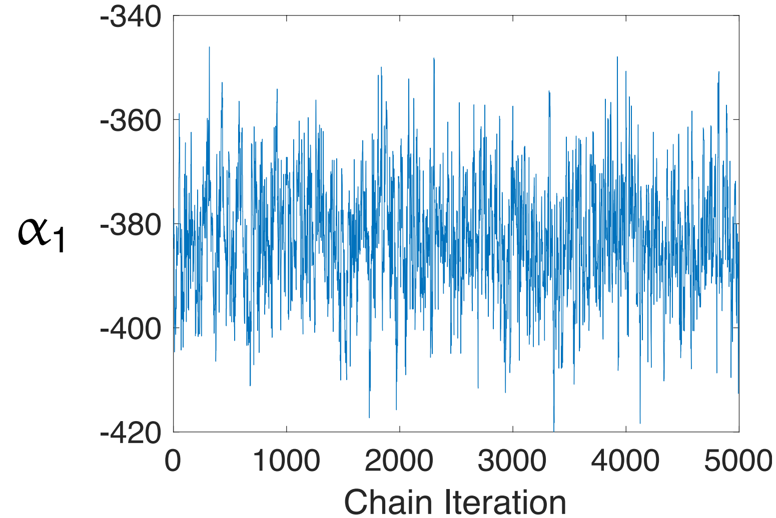
$$\psi(P,q) = \underline{\alpha_1} P^2 + \underline{\alpha_{11}} P^4 + \underline{\alpha_{111}} P^6$$

Note: Similar results for α_{11} and α_{111}

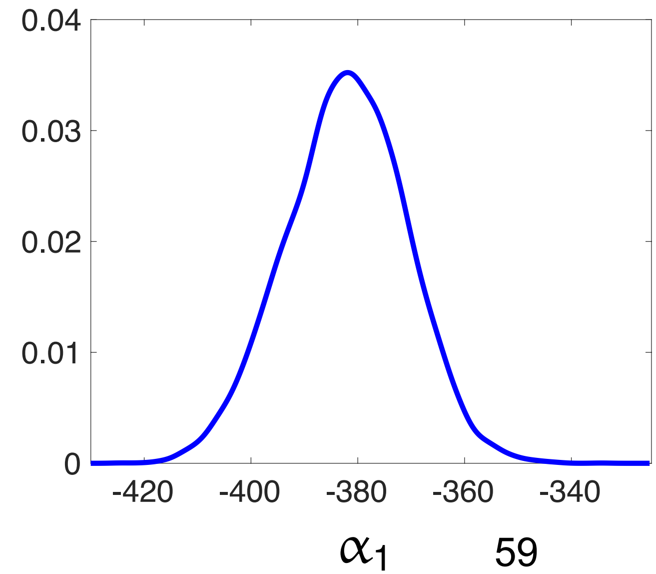
Pairwise Plots: Quantify correlation



Chain for α_1 with 5000 samples



Marginal density for α_1



Bayesian Model Calibration – HIV Example

Model:

$$\dot{T}_1 = \lambda_1 - d_1 T_1 - (1-\varepsilon)k_1 V T_1$$

$$\dot{T}_2 = \lambda_2 - d_2 T_2 - (1-f\varepsilon)k_2 V T_2$$

$$\dot{T}_1^* = (1-\varepsilon)k_1 V T_1 - \delta T_1^* - m_1 E T_1^*$$

$$\dot{T}_2^* = (1-f\varepsilon)k_2 V T_2 - \delta T_2^* - m_2 E T_2^*$$

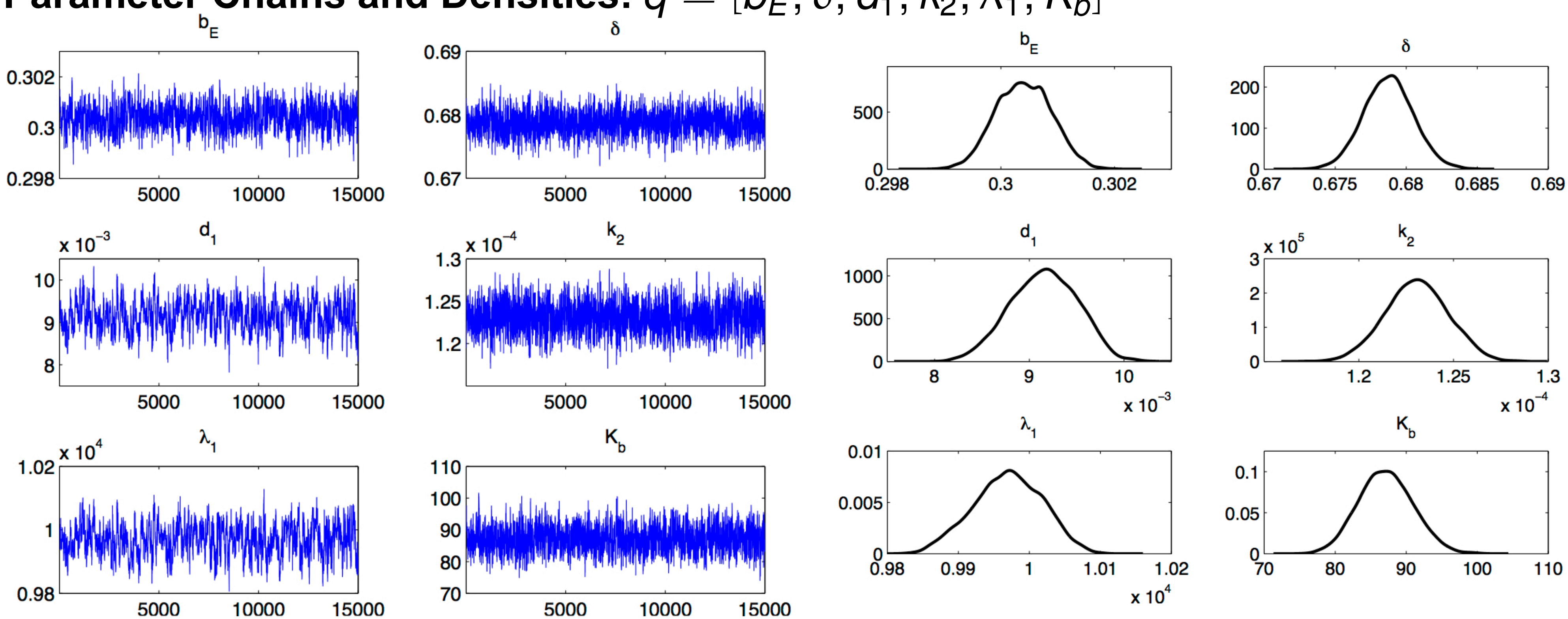
$$\dot{V} = N_T \delta(T_1^* + T_2^*) - cV - [(1-\varepsilon)\rho_1 k_1 T_1 + (1-f\varepsilon)\rho_2 k_2 T_2]V$$

$$\dot{E} = \lambda_E + \frac{b_E(T_1^* + T_2^*)}{T_1^* + T_2^* + K_b}E - \frac{d_E(T_1^* + T_2^*)}{T_1^* + T_2^* + K_d}E - \delta_E E$$

Verification: Why do we trust results??

- Compare results from different algorithms; e.g., DRAM and Gibbs

Parameter Chains and Densities: $q = [b_E, \delta, d_1, k_2, \lambda_1, K_b]$



Bayesian Inference: Advantages and Disadvantages

Advantages:

• Advantageous over frequentist inference when data is limited.

• Directly provides parameter densities, which can subsequently be propagated to construct response uncertainties.

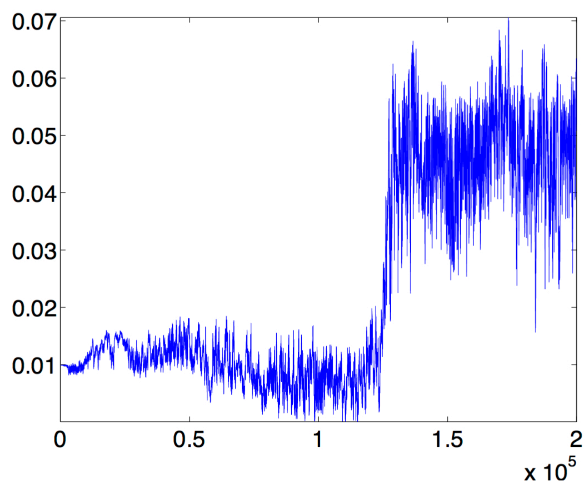
• Can be used to infer non-identifiable parameters if priors are tight.

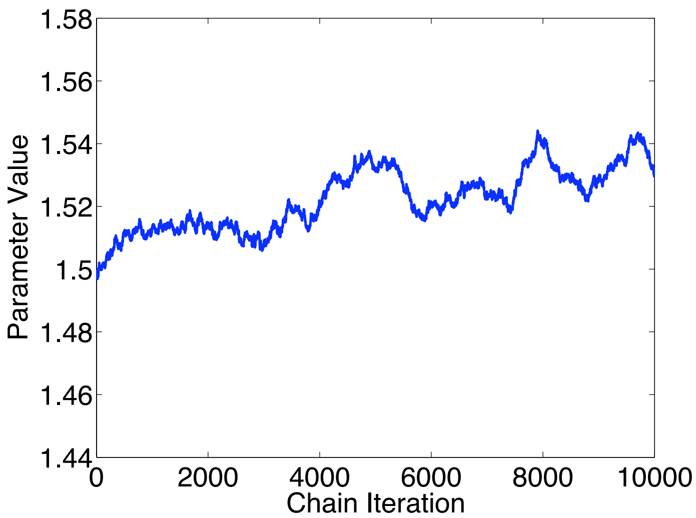
• Provides natural framework for experimental design.

Disadvantages:

• More computationally intense than frequentist inference.

• Can be difficult to confirm that chains have burned-in or converged.





Delayed Rejection Adaptive Metropolis (DRAM)

Websites:

- http://www4.ncsu.edu/~rsmith/UQ_TIA/CHAPTER8/index_chapter8.html
- http://helios.fmi.fi/~lainema/mcmc/

Delayed Rejection Adaptive Metropolis (DRAM)

We fit the Monod model

$$y = \theta_1 \frac{1}{\theta_2 + 1} + \epsilon \quad , \quad \epsilon \sim N(0, I\sigma^2)$$

to observations

x (mg / L COD):	28	55	83	110	138	225	375
y (1 / h):	0.053	0.060	0.112	0.105	0.099	0.122	0.125

First clear some variables from possible previous runs.

```
clear data model options
```

Next, create a data structure for the observations and control variables. Typically one could make a structure data that contains fields xdata and ydata.

```
data.xdata = [28    55    83    110   138   225   375]';   % x (mg / L COD)
data.ydata = [0.053 0.060 0.112 0.105 0.099 0.122 0.125]'; % y (1 / h)
```

Construct model

```
modelfun = @(x,theta) theta(1)*x./(theta(2)+x);
ssfun = @(theta,data) sum((data.ydata-modelfun(data.xdata,theta)).^2);
model.ssfun = ssfun;
model.sigma2 = 0.01^2;
```

Delayed Rejection Adaptive Metropolis (DRAM)

Input parameters

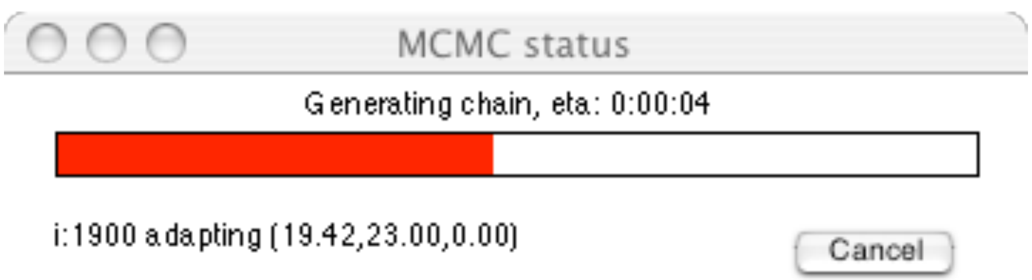
```
params = {
    {'theta1', tmin(1), 0}
    {'theta2', tmin(2), 0} };
```

and set options

```
options.nsimu = 4000;
options.updatesigma = 1;
options.qcov = tcov;
```

Run code

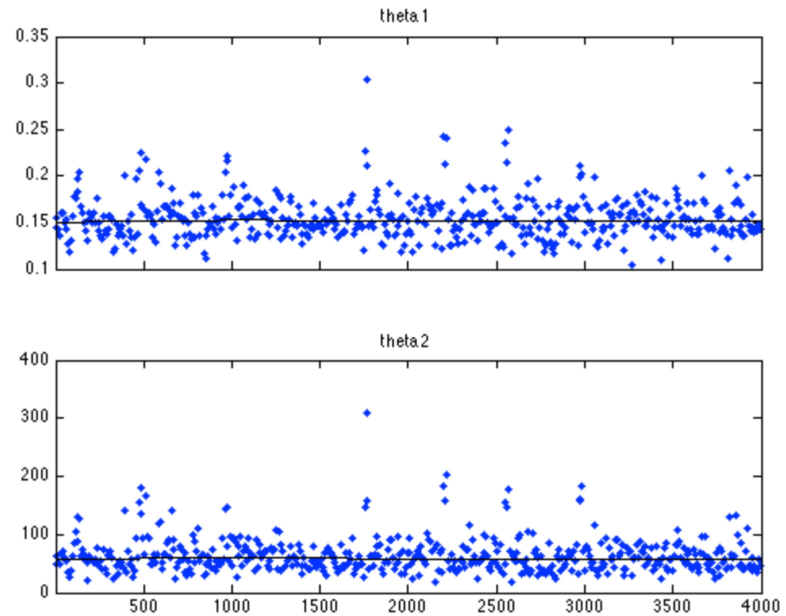
```
[res,chain,s2chain] = mcmcrun(model,data,params,options);
```

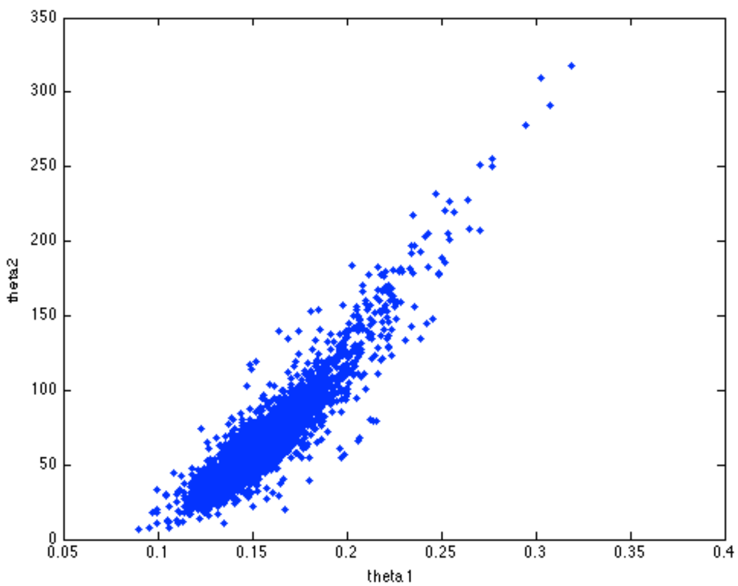


Delayed Rejection Adaptive Metropolis (DRAM)

Plot results

```
figure(2); clf
mcmcplot(chain,[],res,'chainpanel');
figure(3); clf
mcmcplot(chain,[],res,'pairs');
```





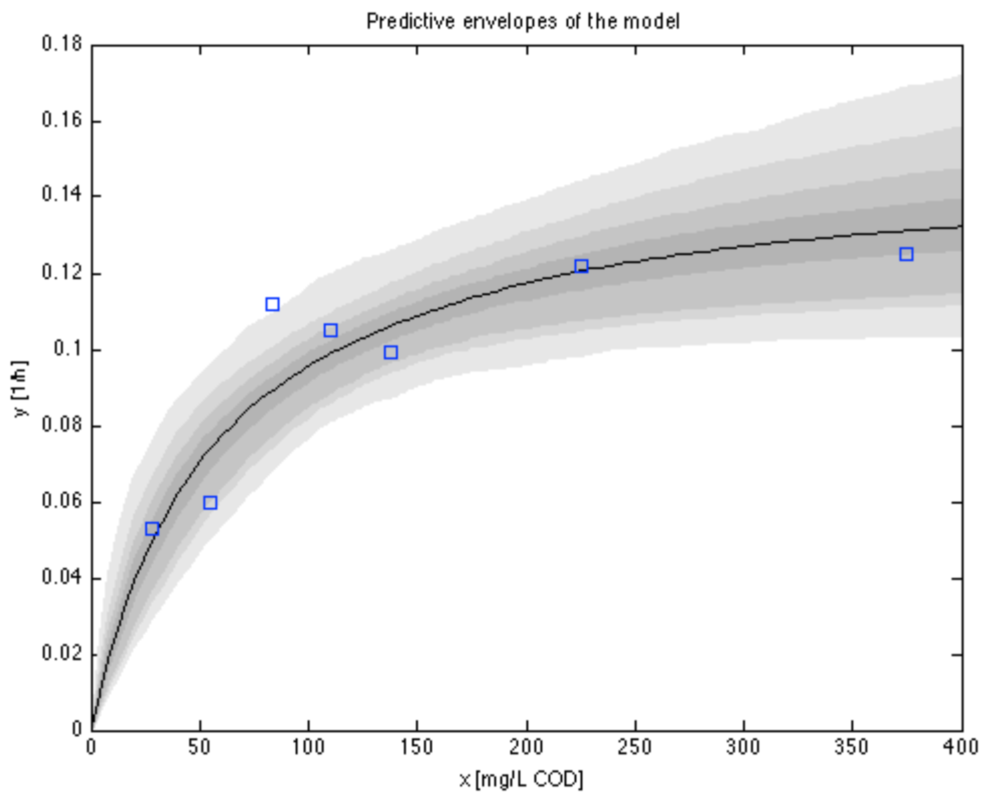
Examples:

- Several available in MCMC_EXAMPLES
- ODE solver illustrated in algae example

Delayed Rejection Adaptive Metropolis (DRAM)

Construct credible and prediction intervals

```
figure(5); clf
out = mcmcpred(res,chain,[],x,modelfun);
mcmcpredplot(out);
hold on
plot(data.xdata,data.ydata,'s');  % add data points to the plot
xlabel('x [mg/L COD]');
ylabel('y [1/h]');
hold off
title('Predictive envelopes of the model')
```



DRAM for SIR Example

SIR Model:

$$\frac{dS}{dt} = \delta N - \delta S - \underline{\gamma k} IS \quad , \; S(0) = S_0 \qquad \text{Susceptible}$$

$$\frac{dI}{dt} = \underline{\gamma k} IS - (r + \delta)I \quad , \; I(0) = I_0 \qquad \text{Infectious}$$

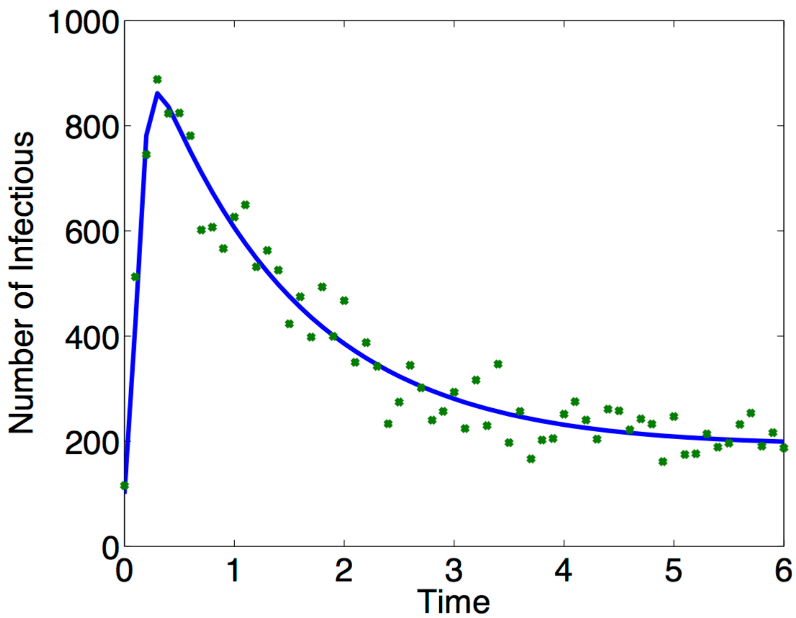
$$\frac{dR}{dt} = rI - \delta R \quad , \; R(0) = R_0 \qquad \text{Recovered}$$

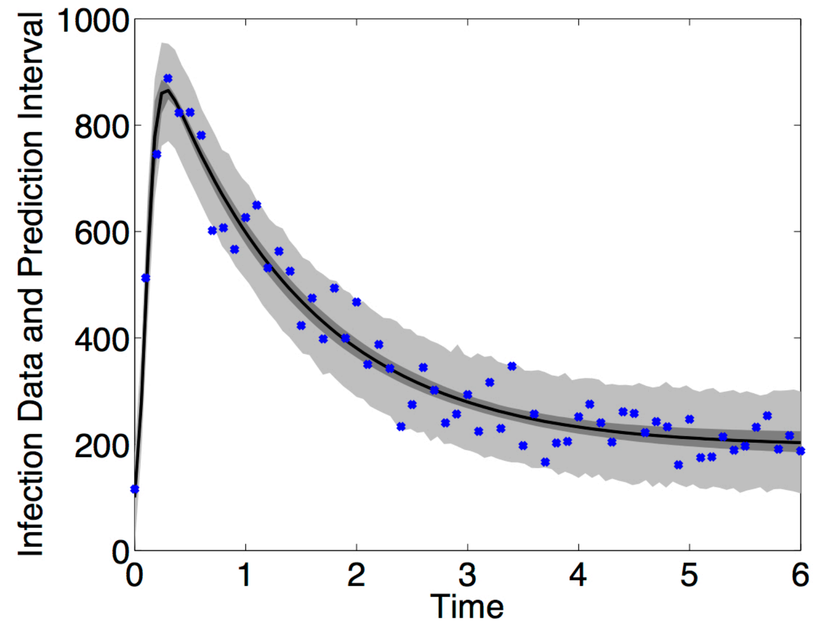
Note: Parameter set $q = [\gamma, k, r, \delta]$ is not identifiable

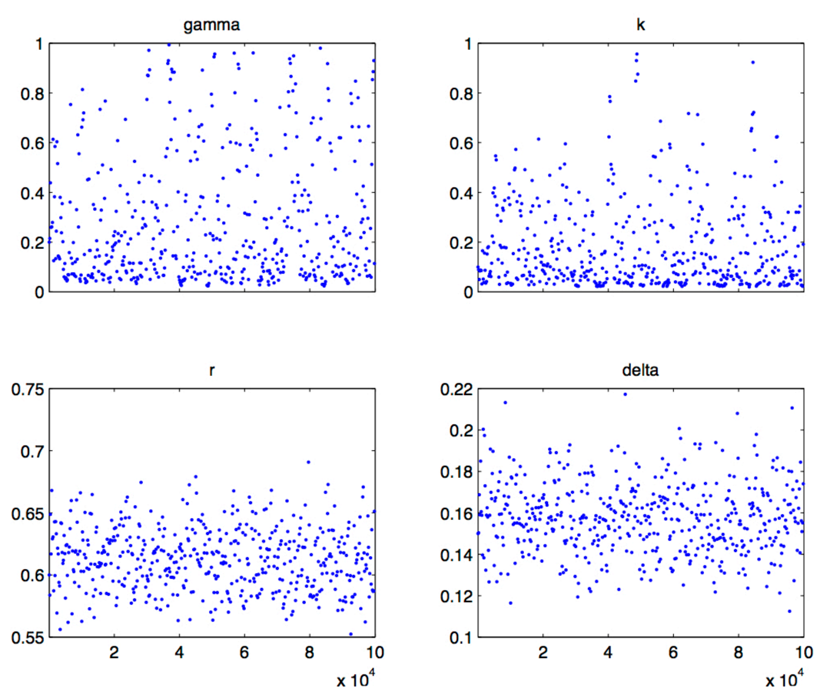
Website

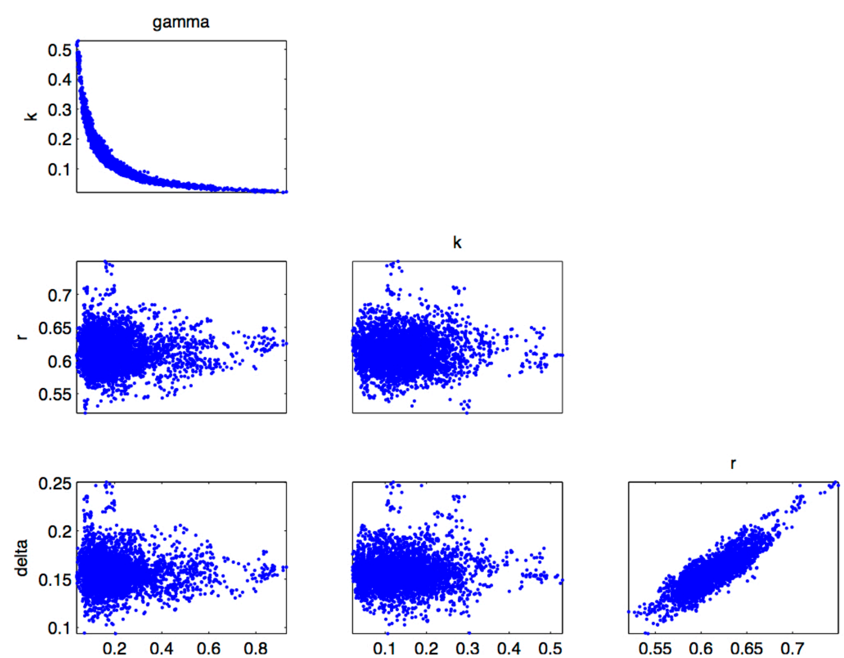
- http://helios.fmi.fi/~lainema/mcmc/
- http://www4.ncsu.edu/~rsmith/

DRAM for SIR Example: Results









SIR Example

3 Parameter SIR Model:

$$\frac{dS}{dt} = \delta N - \delta S - \gamma IS \quad , \; S(0) = S_0 \qquad \text{Susceptible}$$

$$\frac{dI}{dt} = \gamma IS - (r+\delta)I \quad , \; I(0) = I_0 \qquad \text{Infectious}$$

$$\frac{dR}{dt} = rI - \delta R \quad , \; R(0) = R_0 \qquad \text{Recovered}$$

Note:

- Run the posted 4 parameter code and experiment with the chain length.
- Now run the 3 parameter model and compare your results.

Website:

- http://www4.ncsu.edu/~rsmith/DATAWORKS18/