Uncertainty Propagation

Setting:

- We assume that we have determined distributions for parameters
  - e.g., Bayesian inference, experiments

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

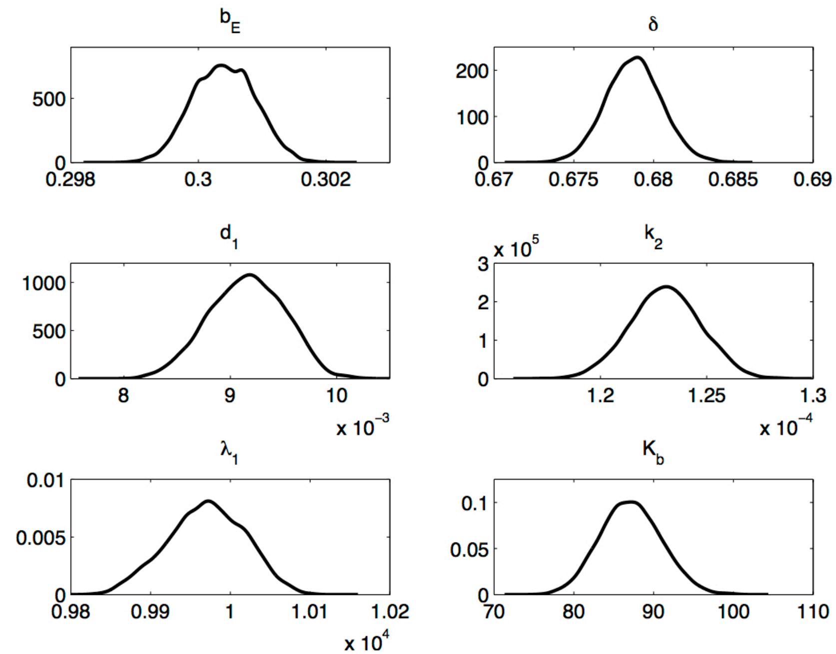
$$\dot{T}_2 = \lambda_2 - d_2 T_2 - (1 - f\varepsilon) k_2 V T_2$$

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

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

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

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



Goal: Construct statistics for QoI

- e.g., Expected viral load in HIV patient with appropriate uncertainty intervals
- Note: Often involves moderate to high-dimensional integration

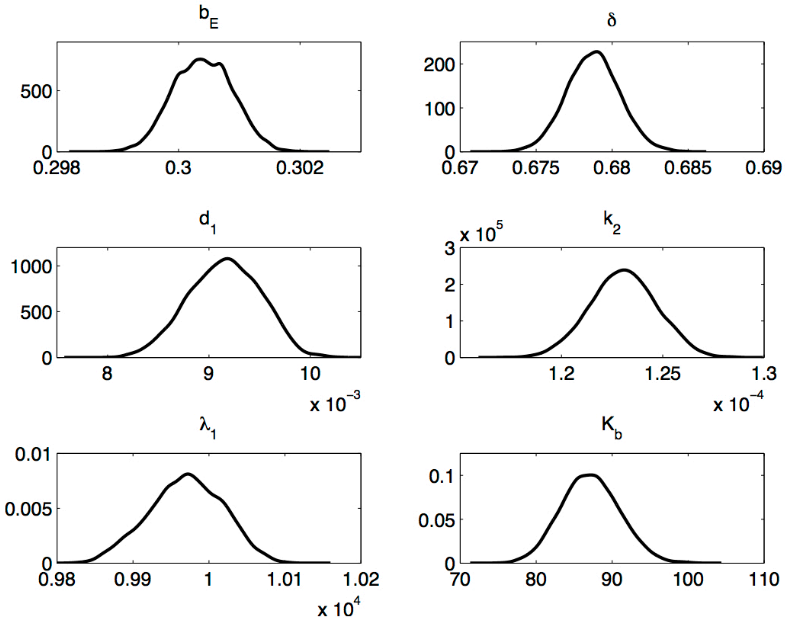
$$\mathbb{E}[V(t)] = \int_{\mathbb{R}^6} V(t,q)\rho(q)dq$$

Issues:

- How do we efficiently propagate input uncertainties through models? Surrogate models.
- How do we approximately integrate in moderate to high dimensions; e.g., p = 50-100?

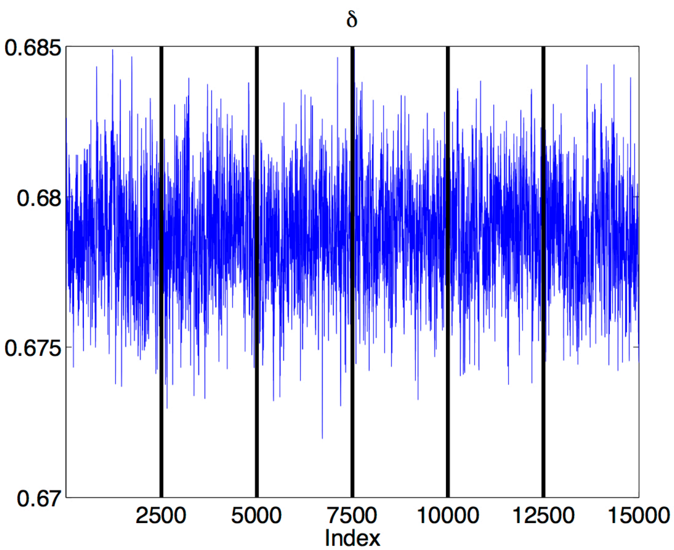
Propagation of Uncertainty – HIV Example

Parameter Densities:

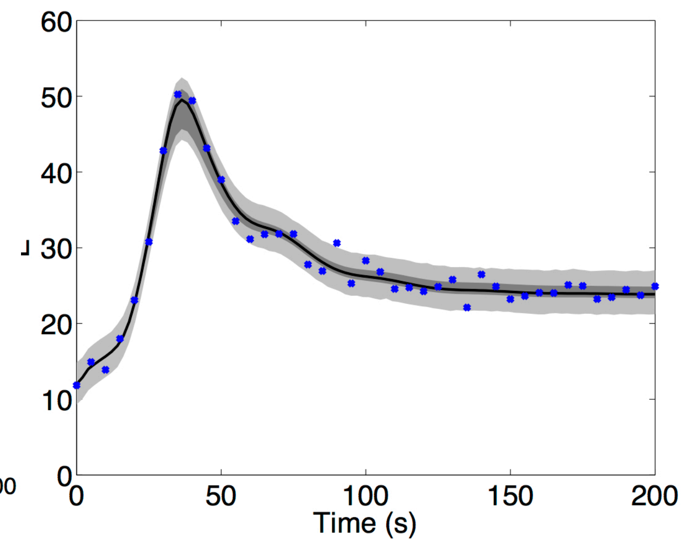


Techniques:

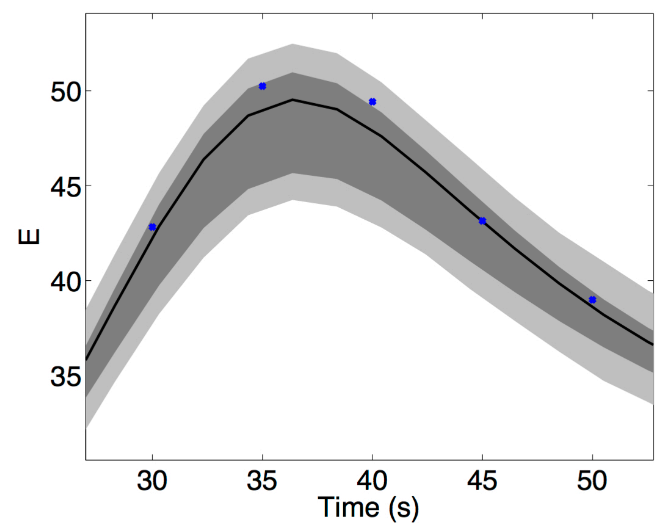
- Sample from parameter and observation error densities to construct mean response, credible intervals, and prediction intervals for QoI.
- Slow convergence rate $\mathcal{O}(1/\sqrt{M})$



Samples from Chain



Data, Credible Intervals and Prediction Intervals



Non-Gaussian Credible and Prediction Intervals

Forward Uncertainty Propagation: Sampling Methods

Strategy: Randomly sample from parameter and measurement error distributions and propagate through model to quantify response uncertainty.

Advantages:

- Applicable to nonlinear models.
- Parameters can be correlated and non-Gaussian.
- Straight-forward to apply and convergence rate is independent of number of parameters.
- Can directly incorporate both parameter and measurement uncertainties.

Disadvantages:

- Very slow convergence rate: $\mathcal{O}(1/\sqrt{M})$ where M is the number of samples.
- 100-fold more evaluations required to gain additional place of accuracy.

Notation: Consider p parameters $q = [\theta, \beta]$

Forward Uncertainty Propagation: Linear Models

Linear Models: Analytic mean and variance relations

Example: Helmholtz energy $q = \theta = [\alpha_1, \alpha_{11}]$

$$\begin{aligned} Y_i &= \psi(P_i, q) + \varepsilon_i \ , \ i = 1, ..., n \\ &= \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \varepsilon_i \end{aligned}$$

Model Statistics:

Let $\bar{\alpha}_1, \bar{\alpha}_{11}, \text{var}(\alpha_1), \text{var}(\alpha_{11})$ denote parameter means and variance. Then

$$\mathbb{E}[\psi(P_i, q)] = \bar{\alpha}_1 P_i^2 + \bar{\alpha}_{11} P_i^4$$

$$\begin{aligned} \text{var}[\psi(P_i, q)] &= \mathbb{E}[\psi^2(P_i, q)] - (\mathbb{E}[\psi(P_i, q)])^2 \\ &= \mathbb{E}[(\alpha_1 P_i^2 + \alpha_{11} P_i^4)^2] - [\bar{\alpha}_1 P_i^2 + \alpha_{11} P_i^4]^2 \\ &= P_i^2 \text{var}(\alpha_1) + P_i^8 \text{var}(\alpha_{11}) + 2 P_i^6 \text{cov}(\alpha_1, \alpha_{11}), \end{aligned}$$

Forward Uncertainty Propagation: Linear Models

Example: Helmholtz energy $q = \theta = [\alpha_1, \alpha_{11}]$

$$
\begin{aligned}
Y_i &= \psi(P_i, q) + \varepsilon_i \;,\; i = 1, \ldots, n \\
&= \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \varepsilon_i
\end{aligned}
$$

Model Statistics:

Let $\bar{\alpha}_1, \bar{\alpha}_{11}, \text{var}(\alpha_1), \text{var}(\alpha_{11})$ denote parameter means and variance. Then

$$
\mathbb{E}[\psi(P_i, q)] = \bar{\alpha}_1 P_i^2 + \bar{\alpha}_{11} P_i^4
$$

$$
\begin{aligned}
\text{var}[\psi(P_i, q)] &= \mathbb{E}[\psi^2(P_i, q)] - (\mathbb{E}[\psi(P_i, q)])^2 \\
&= \mathbb{E}[(\alpha_1 P_i^2 + \alpha_{11} P_i^4)^2] - [\bar{\alpha}_1 P_i^2 + \alpha_{11} P_i^4]^2 \\
&= P_i^2 \text{var}(\alpha_1) + P_i^8 \text{var}(\alpha_{11}) + 2P_i^6 \text{cov}(\alpha_1, \alpha_{11}),
\end{aligned}
$$

Response Statistics: Assume unbiased measurement errors which are uncorrelated from model response.

$$
\mathbb{E}[Y_i] = \bar{\alpha}_1 P_i^2 + \bar{\alpha}_{11} P_i^4
$$

$$
\text{var}[Y_i] = \text{var}[\psi(P_i, q)] + \text{var}(\varepsilon_i).
$$

Forward Uncertainty Propagation: Linear Models

General Linearly Parameterized Model:

$$Y = Xq + \varepsilon$$

Model Statistics:

$$\mathbb{E}[Xq] = X\bar{q}$$

$$\text{cov}[Xq] = XVX^T$$

Response Statistics:

$$\mathbb{E}[Y] = X\bar{q}$$

$$\text{cov}[Y] = XVX^T + V_{obs}$$

Note: Suppose $q_1, \ldots, q_p$ mutually independent and $q \sim \mathcal{N}(\bar{q}, V)$. Then

$$Xq \sim \mathcal{N}(X\bar{q}, XVX^T)$$

$$Y \sim \mathcal{N}(X\bar{q}, XVX^T + Vobs)$$

Forward Uncertainty Propagation: Linear Models

Example: Helmholtz energy

$$Y_i = \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \alpha_{111} P_i^6 + \varepsilon_i \; , \; i = 1, ..., n$$

with $q = \theta = [\alpha_1, \alpha_{11}, \alpha_{111}]$ and

$$X = \begin{bmatrix} P_1^2 & P_1^4 & P_1^6 \\ \vdots & \vdots & \\ P_n^2 & P_n^4 & P_n^6 \end{bmatrix}$$

Bayesian inference yields

$$\bar{q} = \begin{bmatrix} -357.38 \\ 437.90 \\ 774.39 \end{bmatrix} \; , \; V = \begin{bmatrix} 0.0118 & -0.1011 & 0.1996 \\ -0.1011 & 0.9459 & -1.9649 \\ 0.1996 & -1.9649 & 4.2206 \end{bmatrix} \times 10^5$$

for $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$ with $\sigma^2 = 5.5$.

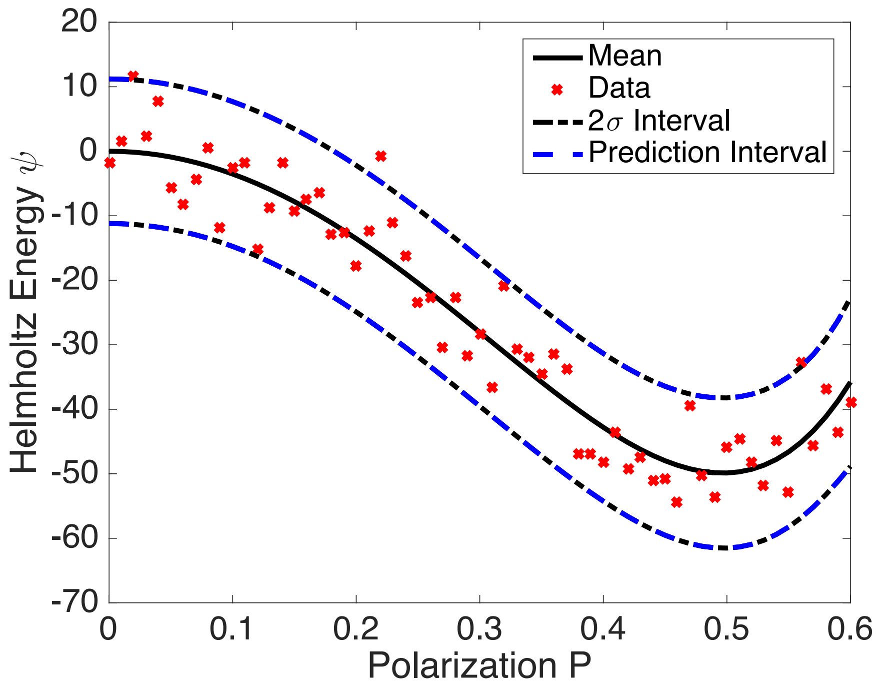
Forward Uncertainty Propagation: Linear Models

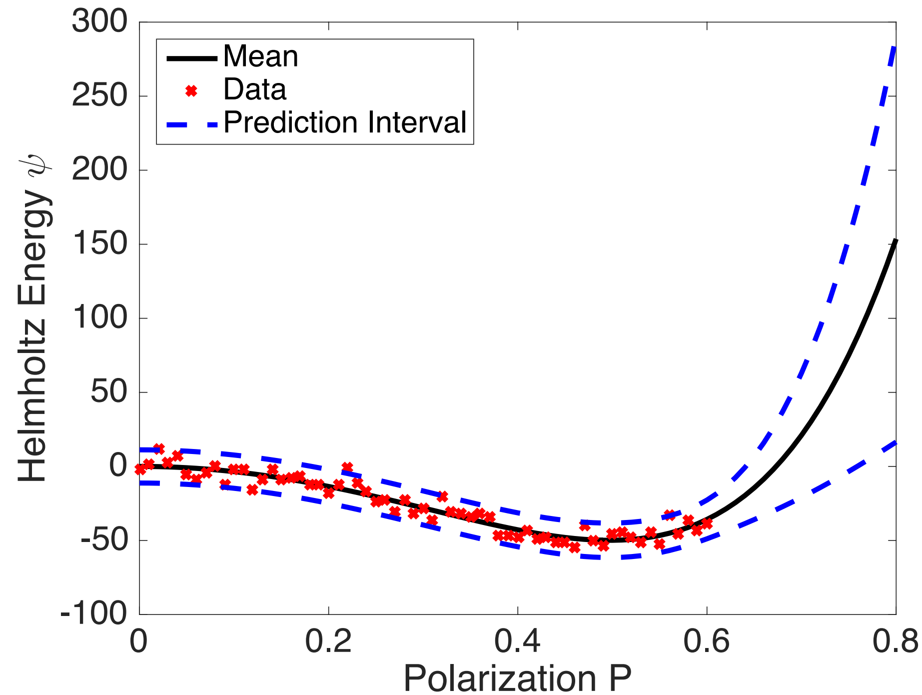
Example: Helmholtz energy

$$Y_i = \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \alpha_{111} P_i^6 + \varepsilon_i \; , \; i = 1, \ldots, n$$

with $q = \theta = [\alpha_1, \alpha_{11}, \alpha_{111}]$ and $P_i \in [0, 6]$

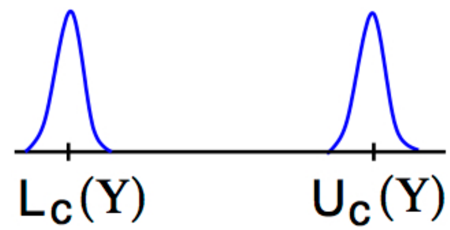
Intervals: We will discussion prediction intervals next.





Confidence, Credible and Prediction Intervals

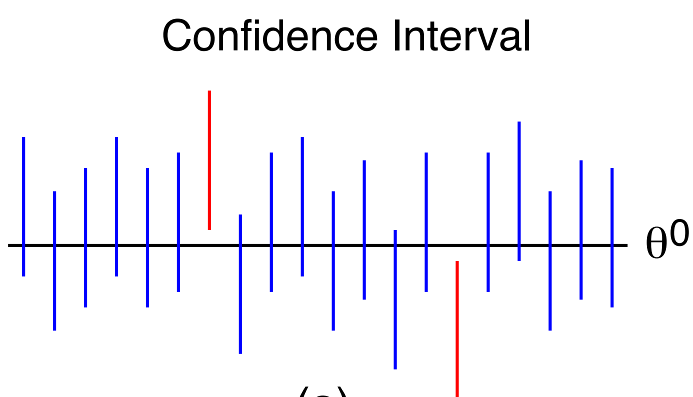
Data: $Y = [Y_1, \ldots, Y_n]$ of iid random observations



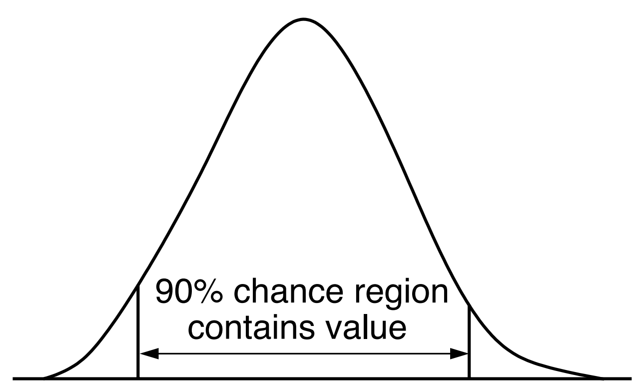
Confidence Interval (Frequentist): A $100 \times (1-\alpha)\%$ confidence interval for a fixed, unknown parameter θ^0 is a random interval $[L_c(Y), U_c(Y)]$ having probability at least $1-\alpha$ of covering θ^0 under joint distribution of Y.

Credible Interval (Bayesian): A $100 \times (1-\alpha)$ % credible interval is that having probability at least $1-\alpha$ of containing θ.

Strategy: Sample out of parameter density $\rho(\theta)$



(a)



(b)

Confidence and Prediction Intervals for Responses

Linear Model:

$$Y = X\theta + \varepsilon \quad , \quad \varepsilon_i \overset{iid}{\sim} \mathcal{N}(0, \sigma_0^2)$$

Goal: Predict Y_{x^*} at $x^* = [x_1^*, \dots, x_p^*]$ not among data in X used to infer θ

Two Cases:

(i) Mean response $\mu_{x^*} = \mathbb{E}(Y_{x^*})$

(ii) New prediction Y_{x^*}

Case (i): Consider unbiased point estimator

$$\widehat{Y}_{x^*} = x^* \hat{\theta}$$

Reference: Section 11.2

Confidence and Prediction Intervals for Responses

Case (i): Consider unbiased point estimator $\widehat{Y}_{x^*} = x^*\hat{\theta}$

From properties

$$\mathrm{var}\left(\sum_{i=1}^{n} x_i\theta_i\right) = \sum_{i=1}^{n} x_i^2\mathrm{var}(\theta_i) + 2\sum_{i<j} x_i x_j \mathrm{cov}(\theta_i, \theta_j)$$

$$\mathrm{var}(\hat{\theta}) = \sigma_0^2(X^T X)^{-1}$$

it follows that

$$\mathrm{var}(\widehat{Y}_{x^*}) = \sigma_0^2\left[x^*(X^T X)^{-1}x^{*T}\right]$$

Estimator:

$$\mathrm{var}(\widehat{Y}_{x^*}) = \widehat{\sigma}^2\left[x^*(X^T X)^{-1}x^{*T}\right]$$

Note: $\varepsilon_i \sim \mathcal{N}(0, \sigma_0^2)$ implies that

$$\frac{\widehat{Y}_{x^*} - \mu_{x^*}}{\sigma_0\sqrt{x^*(X^T X)^{-1}x^{*T}}} \sim \mathcal{N}(0, 1)$$

Confidence and Prediction Intervals for Responses

Note: It then follows that

$$T = \frac{\widehat{Y}_{x^*} - \mu_{x^*}}{\hat{\sigma}\sqrt{x^*(X^T X)^{-1} {x^*}^T}} \sim t(n-p)$$

Confidence Interval:

$$\left[\widehat{Y}_{x^*} \pm t_{n-p,1-\alpha/2} \cdot \hat{\sigma}\sqrt{x^*(X^T X)^{-1} {x^*}^T}\right]$$

Prediction Interval: Assume that estimators $\hat{\theta}$ and $\hat{\sigma}$ have been computed using previous data so independent from Y_{x^*}. Then $Y_{x^*} - \widehat{Y}_{x^*}$ is normally distributed with

$$\mathbb{E}(Y_{x^*} - \widehat{Y}_{x^*}) = x^* \theta^0 - x^* \mathbb{E}(\hat{\theta}) = 0$$

$$\begin{aligned} \mathrm{var}(Y_{x^*} - \widehat{Y}_{x^*}) &= \mathrm{var}(Y_{x^*}) + \mathrm{var}(\widehat{Y}_{x^*}) \\ &= \sigma_0^2 \left[1 + x^*(X^T X)^{-1} {x^*}^T\right] \end{aligned}$$

Frequentist Prediction Interval for Responses

Result:

$$\frac{\widehat{Y}_{x^*} - Y_{x^*}}{\sigma_0\sqrt{1 + x^*(X^TX)^{-1}x^{*T}}} \sim \mathcal{N}(0,1)$$

$$T = \frac{\widehat{Y}_{x^*} - Y_{x^*}}{\hat{\sigma}\sqrt{1 + x^*(X^TX)^{-1}x^{*T}}} \sim t(n-p)$$

Interval Estimator:

$$\left[\widehat{Y}_{x^*} \pm t_{n-p,1-\alpha/2} \cdot \hat{\sigma}\sqrt{1 + x^*(X^TX)^{-1}x^{*T}}\right]$$

Prediction Interval: A $100 \times (1-\alpha)\%$ prediction interval for a future observable Y_{x^*} is a random interval $[L(Y), U(Y)]$ having probability at least $1-\alpha$ of of containing Y_{x^*} under joint distribution of (Y, Y_{x^*}).

Bayesian Prediction Interval for Responses

Observation Model:

$$Y_i = f(x_i, \theta) + \varepsilon_i \ , \ \ \varepsilon_i \sim \mathcal{N}(0, \sigma^2) \ , \ \ i = 1, \ldots, n$$

Likelihood:

$$f(y|\theta, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\sum_{i=1}^{n} [y_i - f(x_i, \theta)]^2 / 2\sigma^2\right)$$

Posterior Predictive Distribution:

$$\pi^*(Y_{x^*}|y) = \int_{\mathbb{R}} \int_{\mathbb{R}^p} f(Y_{x^*}|\theta, \sigma^2)\pi(\theta, \sigma^2|y)\, d\theta\, d\sigma^2$$

where $\pi(\theta, \sigma^2|y)$ is the parameter posterior density.

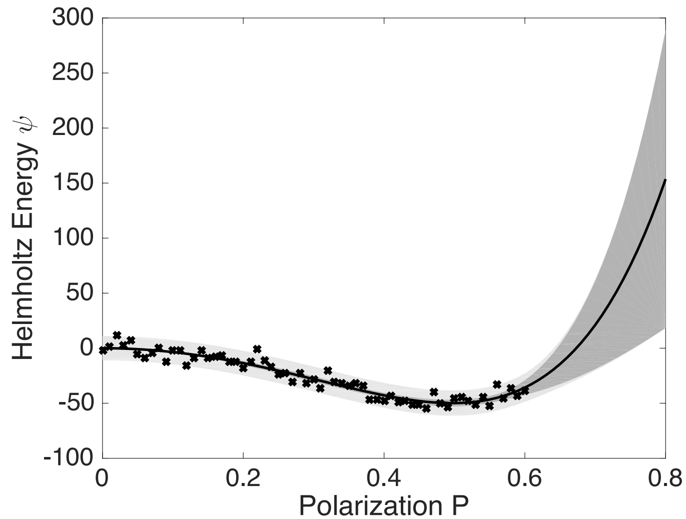
Reference: Section 13.3

Prediction Intervals: Linear Models

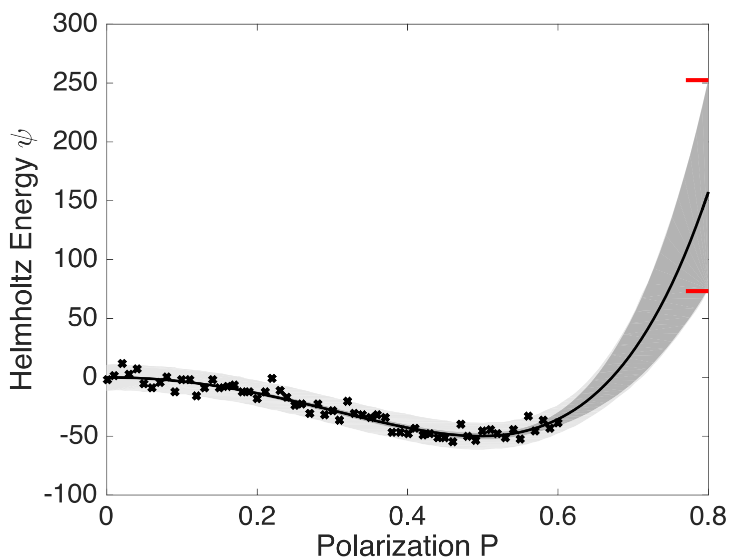
Example: Helmholtz energy with $q = \theta = [\alpha_1, \alpha_{11}, \alpha_{111}]$ and $P_i \in [0, 6]$

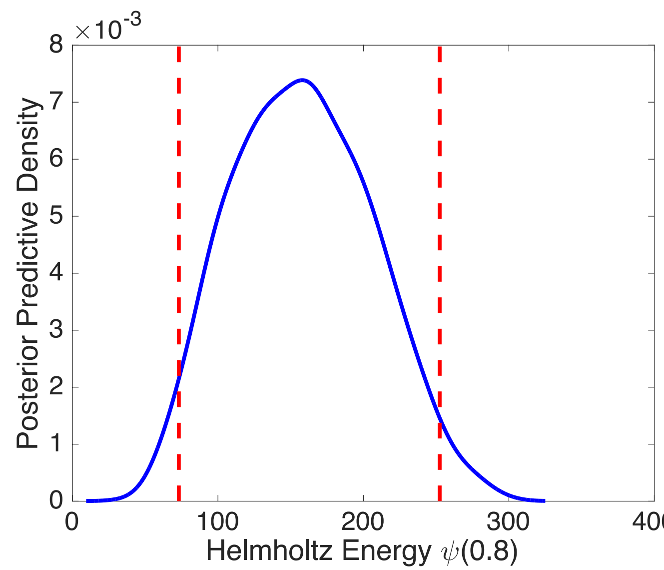
$$Y_i = \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \alpha_{111} P_i^6 + \varepsilon_i \ , \ i = 1, \ldots, n$$

Frequentist confidence and prediction



Bayesian credible and prediction





Perturbation Methods for Nonlinear Models

Model:

$$Y = f(q) + \varepsilon$$

Assumption:

$$\int_{\mathbb{R}^p} (q_i - \bar{q}_i)\rho(q)dq = 0 \ , \ \ i = 1, \ldots, p$$

Note: For $q = \bar{q} + \delta q$,

$$\begin{aligned} f(q) &= f(\bar{q}_1 + \delta q_1, \ldots, \bar{q}_p + \delta q_p) \\ &\approx f(\bar{q}) + \sum_{i=1}^{p} \left.\frac{\partial f}{\partial q_i}\right|_{\bar{q}} \delta q_i \\ &\approx \bar{y} + \sum_{i=1}^{p} s_i \delta q_i \end{aligned}$$

and

$$\mathbb{E}(q_i) = \int_{\mathbb{R}^p} q_i \rho(q)dq = \bar{q}_i,$$

$$\mathrm{var}(q_i) = \int_{\mathbb{R}^p} (q_i - \bar{q}_i)^2 \rho(q)dq = \int_{\mathbb{R}^p} (\delta q_i)^2 \rho(q)dq,$$

$$\mathrm{cov}(q_i, q_j) = \int_{\mathbb{R}^p} (q_i - \bar{q}_i)(q_j - \bar{q}_j)\rho(q)dq = \int_{\mathbb{R}^p} \delta q_i \delta q_j \rho(q)dq,$$

Perturbation Methods for Nonlinear Models

It follows that

$$\mathbb{E}[f(q)] = \bar{y} \int_{\mathbb{R}^p} \rho(q)dq + \sum_{i=1}^{p} s_i \int_{\mathbb{R}^p} (q_i - \bar{q}_i)\rho(q)dq = \bar{y}$$

and

$$\begin{aligned}
\mathrm{var}[f(q)] &= \mathbb{E}[(f(q) - \bar{y})^2] \\
&= \int_{\mathbb{R}^p} \left(\sum_{i=1}^{p} s_i \delta q_i \right)^2 \rho(q)dq \\
&= \sum_{i=1}^{p} s_i^2 \int_{\mathbb{R}^p} (\delta q_i)^2 \rho(q)dq + \sum_{i=1}^{p} \sum_{\substack{j=1 \\ j \neq i}}^{p} s_i s_j \int_{\mathbb{R}^p} (\delta q_i)(\delta q_j)\rho(q)dq \\
&= \sum_{i=1}^{p} s_i^2 \mathrm{var}(q_i) + \sum_{i=1}^{p} \sum_{\substack{j=1 \\ j \neq i}}^{p} s_i s_j \mathrm{cov}(q_i, q_j),
\end{aligned}$$

Notes:

- S and V are the local sensitivity vector and covariance matrix.
- This is often termed the “sandwich relation”.
- Suppose $q_i \sim \mathcal{N}(0, \sigma_i^2)$ are mutually independent. Then $f(q) \sim \mathcal{N}(\bar{y}, s^T V s)$.

Perturbation and Sampling for Forward Propagation

Example: Consider

$$m\frac{d^2z}{dt^2} + c\frac{dz}{dt} + kz = f_0\cos(\omega_F t),$$

$$z(0) = z_0\ ,\ \frac{dz}{dt}(0) = z_1$$

with q = [m,c,k]. This has the amplitude

$$Z_0(q) = \frac{f_0}{\sqrt{m^2(\omega_0^2 - \omega_F^2)^2 + c^2\omega_F^2}}\ ,\ \omega_0 = \sqrt{k/m}$$

Consider the response

$$y = f(\omega_F, q) = \frac{Z_0(q)}{f_0} = \frac{1}{\sqrt{(k - m\omega_F^2)^2 + (c\omega_F)^2}}$$

Take $q = [m, c, k] \sim N(\bar{q}, V)$ where

$$\bar{q} = [2.7, 0.24, 8.5] \quad , \quad V = \begin{bmatrix} 0.002^2 & 0 & 0 \\ 0 & 0.065^2 & 0 \\ 0 & 0 & .001^2 \end{bmatrix}$$

Perturbation and Sampling for Forward Propagation

Notes:

- Natural frequency $\bar{\omega}_0 = 1.7743 \text{ Hz}$
- Analytic sensitivity relations

$$s^T = \left[\frac{\partial f}{\partial m}, \frac{\partial f}{\partial c}, \frac{\partial f}{dk} \right]$$

Compare:

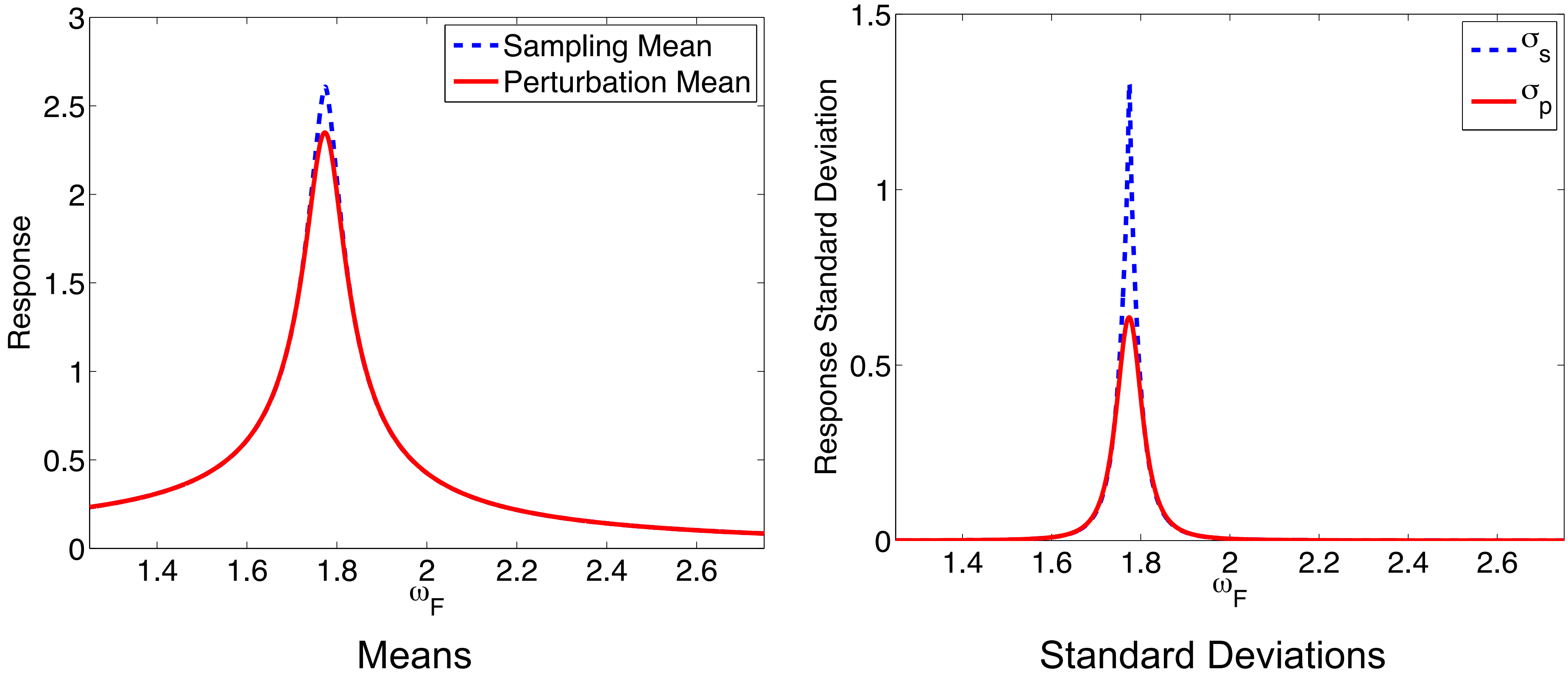
- Perturbation result
- Sample M = 10,000 and construct sample mean and variance

$$\bar{y}_s(\omega_F) = \frac{1}{M} \sum_{m=1}^{M} f(\omega_F, q^m)$$

$$\omega_s^2(\omega_F) = \frac{1}{M-1} \sum_{m=1}^{M} \left[f(\omega_F, q^m) - \bar{y}_s(\omega_F) \right]^2$$

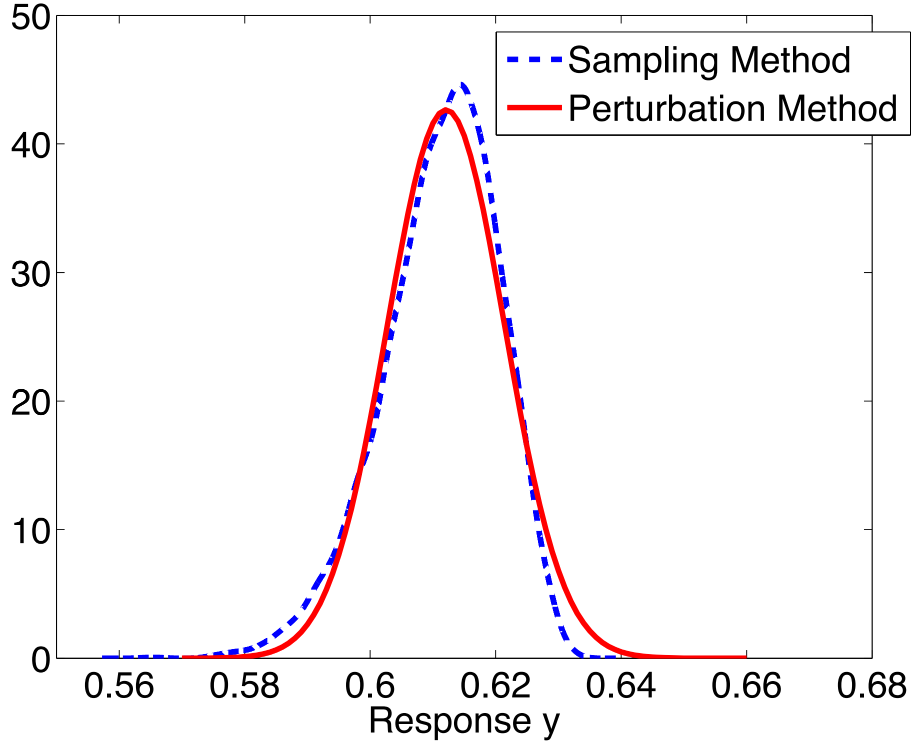
Perturbation and Sampling for Forward Propagation

Results:

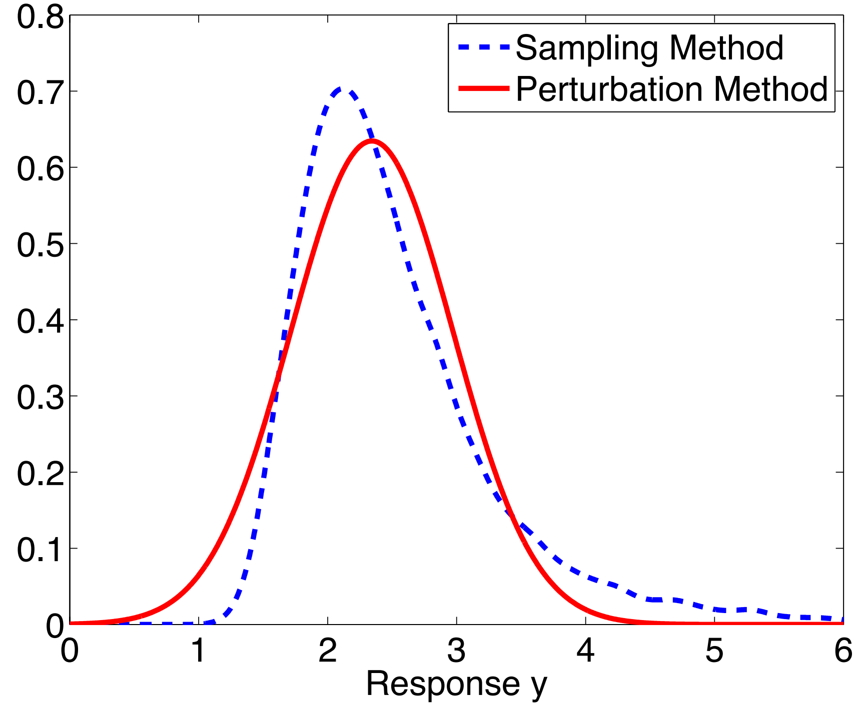


Perturbation and Sampling for Forward Propagation

Results: Recall the natural frequency $\bar{\omega}_0 = 1.7743 \text{ Hz}$



1.60 Hz



1.77 Hz

Uncertainty Propagation in Models

Example: HIV model

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

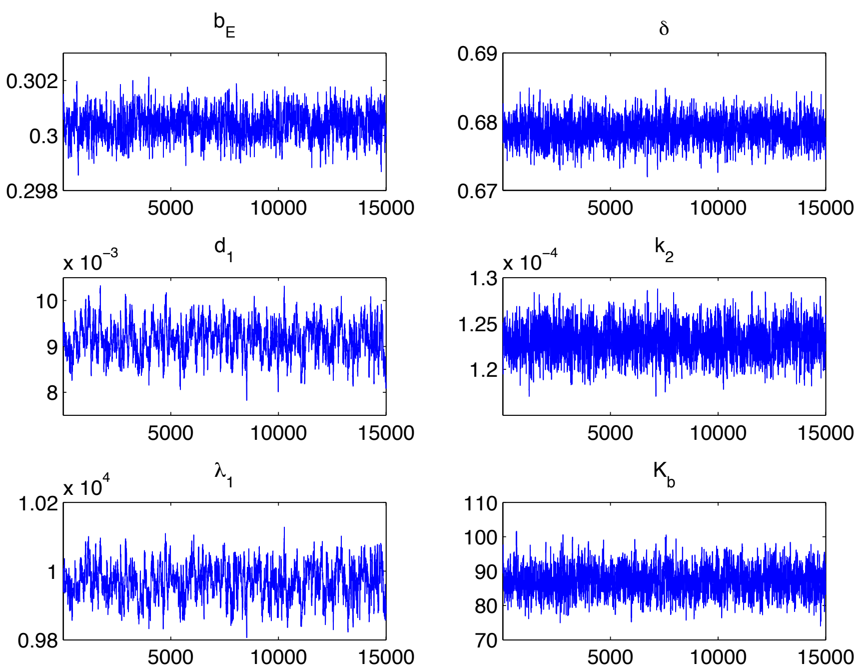
$$\dot{T}_2 = \lambda_2 - d_2 T_2 - (1-f\varepsilon) k_2 V T_2$$

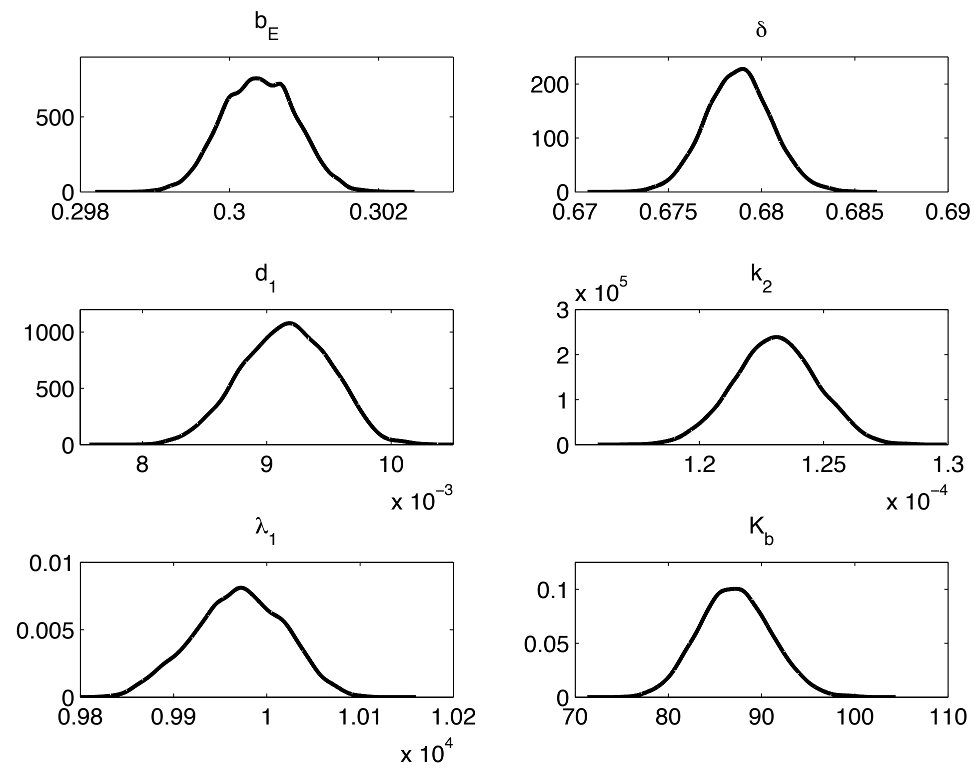
$$\dot{T}_1^* = (1-\varepsilon) k_1 V T_1 - \delta T_1^* - m_1 E T_1^*$$

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

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

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





Uncertainty Propagation in Models

Example: HIV model