TESTING MODELS

- The **Assumptions** are valid
- Model **Structure** is sound
- Model **Parameters** are believable
- Model **Predictions** match observation

Important to remember that the model is an approximation - ignore less important features, "random errors" are expected.

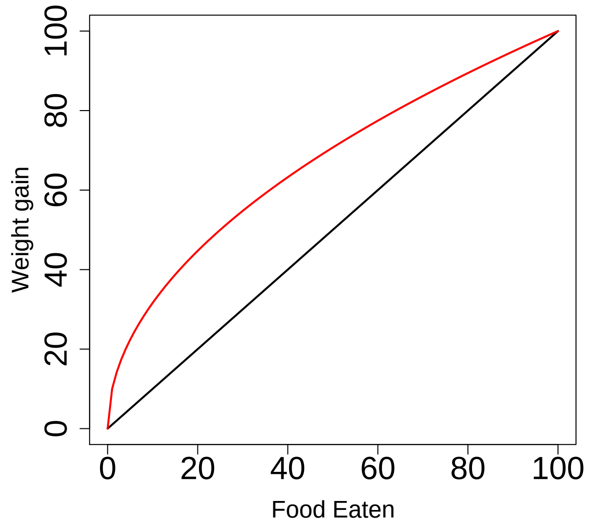
Assumptions

- Consider what the assumptions really **are**

- What **should** they be?

  e.g. Do we want a linear relationship or any increasing relationship?

Assumptions are often models and may need testing as such!



Model Structure

How sensitive is the model to changes in model structure?

Quantitive changes (predicted value changes)

Qualitative changes (nature of prediction changes)

Some structures that can change outputs:

Stochasticity

Non-linearity

Modelling physical space explicitly

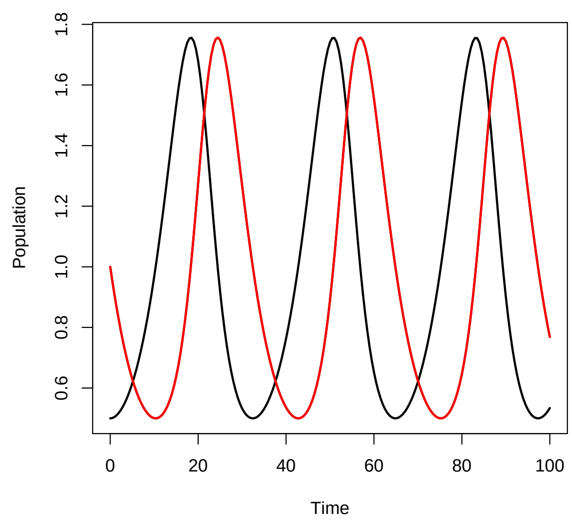
...!

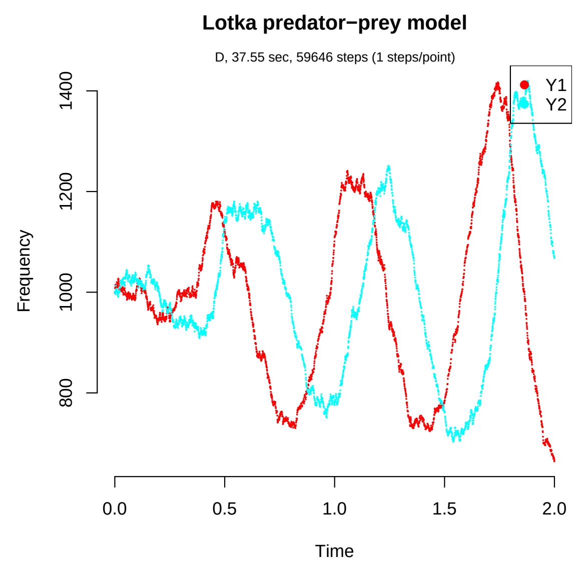
Only qualitative (different behaviour) changes matter at this stage!

Example: Changing Model Structure

Lotka-Volterra Predator Prey model:

Deterministic model gives stable cycles... Adding ANY noise produces unstable cycles!





Parameters

Question: How sensitive is the model to the parameters?

Another: how do you find the "best" parameters?

Need Goodness of fit measure.

Most commonly used are:

- Sum-squared error
- **Likelihood**

Estimating model parameters

Deterministic models

- Minimise mean square error
- Assume normal and uncorrelated errors $\rightarrow$ std. errors
- Pretend model is true, discrepancy is observation error
- Likelihood $L(D \mid p)$: probability that the observed errors could happen

  Likelihood $= \exp(-$ sum of [squared error/standard deviation])

Stochastic models

- Likelihood $L(D \mid p)$ is a natural concept
- Only use model assumptions $\rightarrow$ parameter distributions
- Missing observations must be averaged out

Estimating model parameters

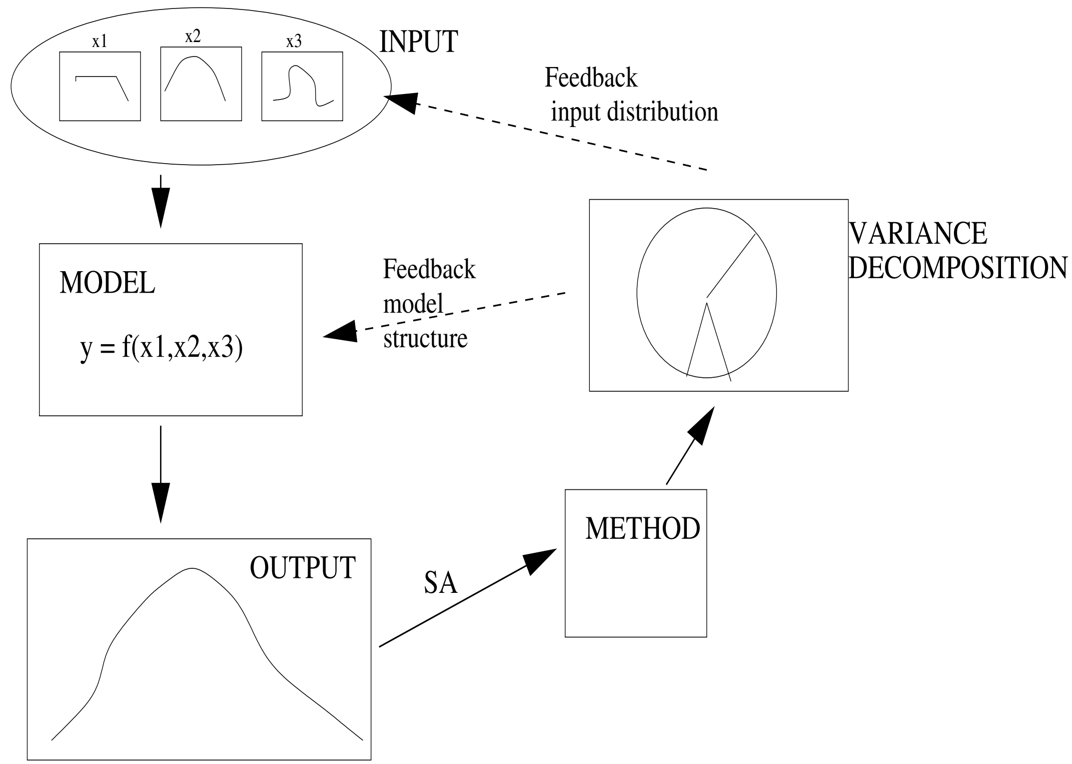
To obtain "good" model parameters:

Maximise Likelihood (or minimise errors)
Many methods to do this:

- Solve for maximum (usually impossible!)
- Numerically find "Maximum Likelihood" (easy, but only local maxima found)
- Simulated Annealing (hard, should find global maxima)
- MCMC (hard, should find global maxima AND give correct parameter distribution around it)
- ...

Sensitivity Analysis

(Sensitivity Analysis eds. Saltelli, Chan, Scott.)



Sensitivity Analysis Methods

LOCAL

(directional) derivative:

$S_i = \frac{x_j}{y_i}\frac{\partial y_i}{\partial x_j}$

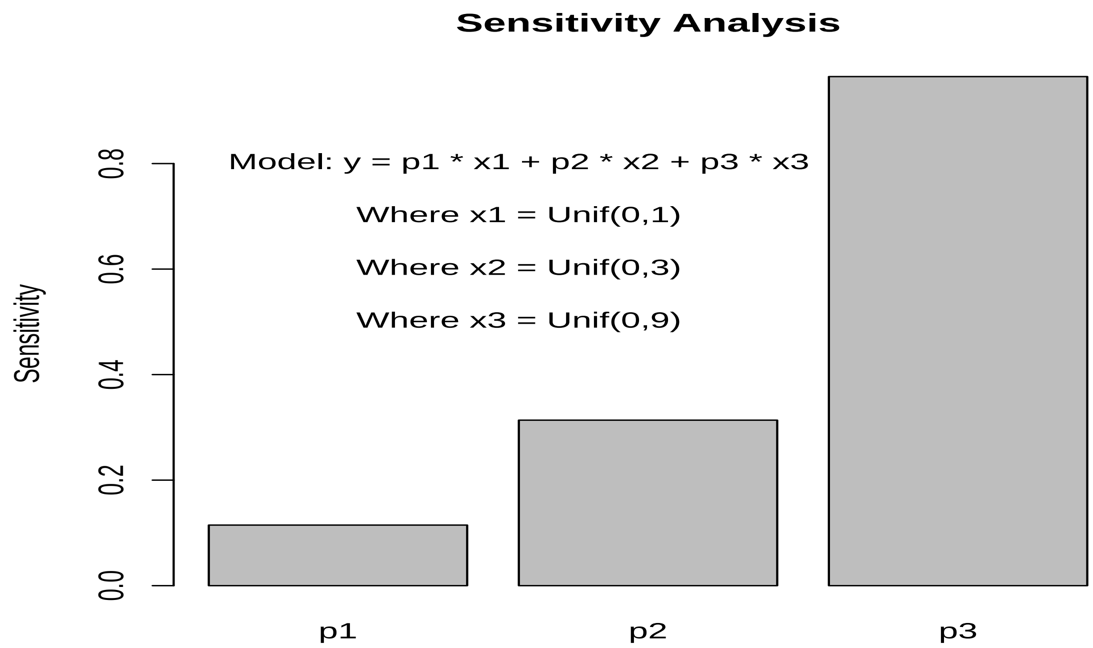
i.e. the relative change from making a small change to each parameter

GLOBAL

- Sampling based methods (Monte Carlo, Latin Hypercube Sampling)
  - Sensitivity Indices (Importance Measures, SOBOL)
  - FAST sampling (Fourier methods)

These are all implemented in the R package "sensitivity".

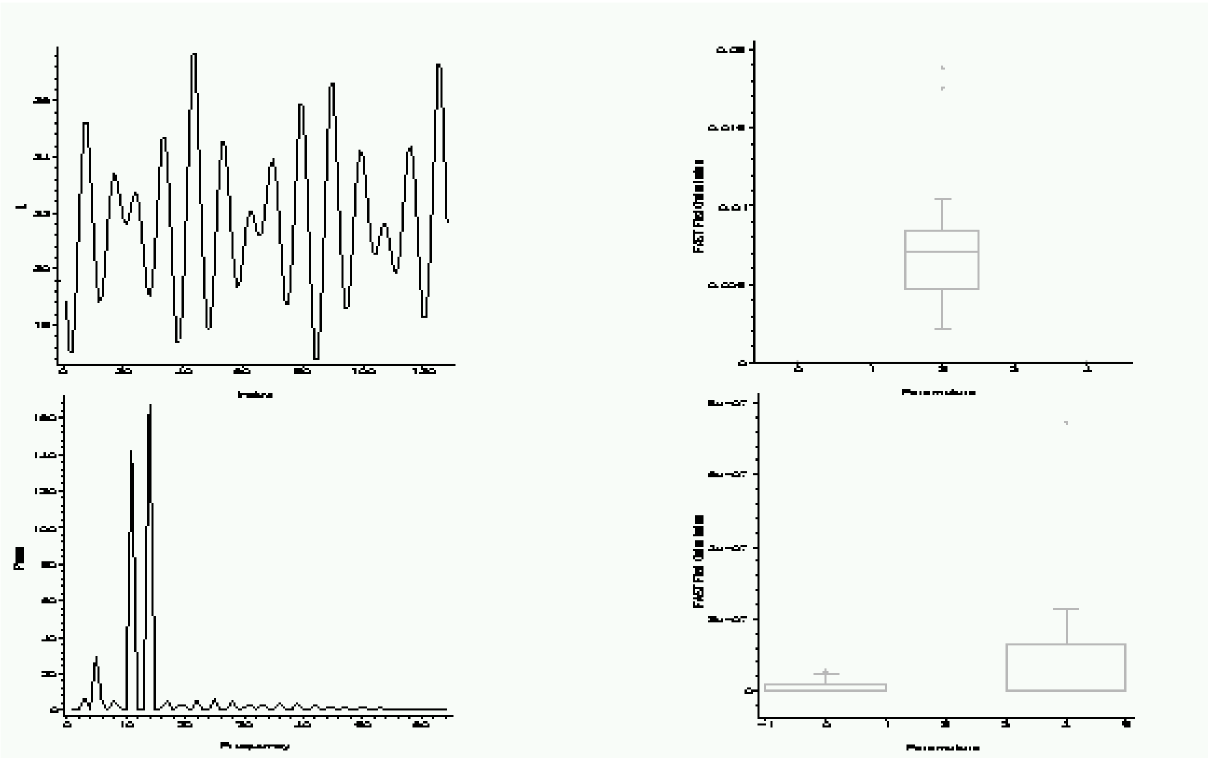
Sensitivity Analysis Example



Relative change in output from each input

Sensitivity Analysis Example 2

HeathMod Grazing System (MLURI)



Prediction

Test model prediction against NEW and INDEPENDENT data!
Why:

- We used the old data to *fit* the model - avoid overfitting
- Want *general* prediction - not for specific case

  e.g. Fit model for a farm. Better to test on *different* farm, rather than new data for old farm. Otherwise only get good model for one specific farm!

Sometimes hard to get new data; instead, resample current data.

- Split data into chunks (e.g. leave one out)
- Randomize over which data used to fit model, which to predict *from* model

Reasons for Prediction Errors

A) "Random" variation

B) Excluded effects, i.e. incomplete model

C) Wrong model:

- poor model structure
- poor parameter estimates

Summary Statistics

Bias

$$B = \sum_i (O_i - P_i)/n$$

Standard deviation

$$SD = \{\sum_i [(O_i - P_i) - B]^2/n\}^{\frac{1}{2}}$$

Prediction mean square error

$$PMSE = \sum_i (O_i - P_i)^2/n = B^2 + SD^2$$

Comparison of Models

Issues to consider include:

- generality
- sensitivity
- predictive ability

NB equations with different functional forms can give similar predictions

Subjective element - what is it you *want* from the model?

Information criterion

With Likelihood L, and using k parameters for the model, consider:

- AIC (Akaike Information Criterion)
- BIC (Bayesian Information Criterion)
- DIC (Deviance Information Criterion)

Each has a slightly different form and is built on different assumptions - but usually agree

Many (stats) programs provide them routinely

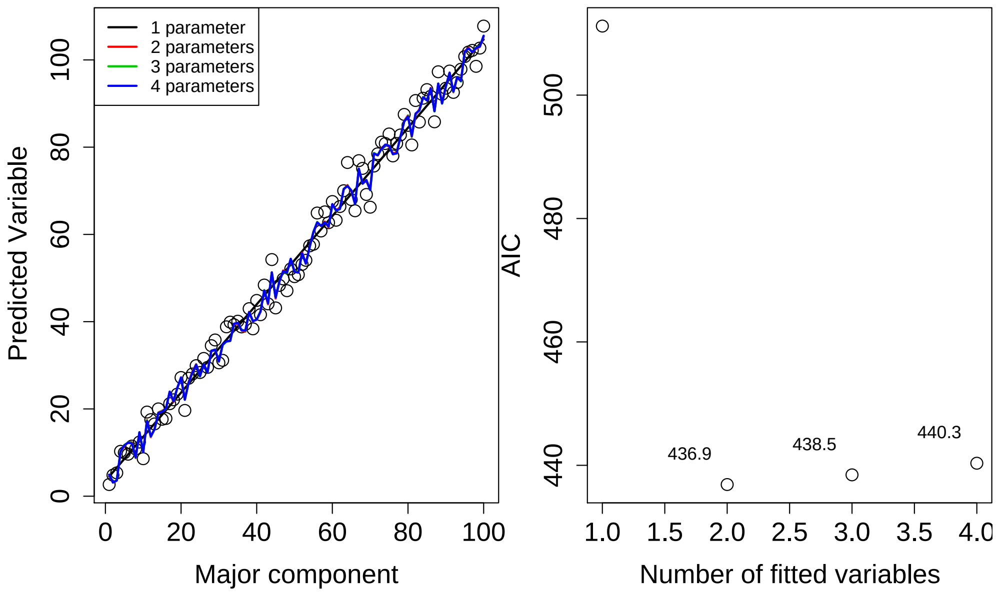
Take the form $-2\log(L) + f(n)$
($f = 2k$ for AIC, $klog(n)$ for BIC, where n is the number of observations)

Choose minimum IC value; difference of 10 is very significant, $5 - 10$ is strong, $0 - 5$ means both models could be right

Example: AIC

Model: linear model with 4 components: Each a factor a less important than the last, with noise.

Linear model: $y_i = \sum_{j=1:4}(a^j x_{ij} + \sigma_{ij})$



Use lowest AIC for predictions: Only use 2 variables. WHY? The noise swamps out the other 2 and so it isnt worth the extra complexity.

Comparing models for cattle growth

Two models for prediction of liveweight gain in growing cattle

Notation:

$$
\begin{array}{lcl}
MEI & = & \text{metabolisable energy of daily ration (MJ/d)} \\
q & = & \text{ration of metabolisable to gross energy in the diet} \\
E_m & = & \text{energy of maintenance (MJ/d)} \\
k_m & = & \text{efficiency of utilisation of dietary ME for maintenance} \\
L & = & MEI * k_m/E_m \text{(level of feeding)} \\
E_g & = & \text{energy retained in daily weight change (MJ/d)} \\
k_g & = & \text{efficiency of utilisation of dietary ME for weight change} \\
EV_g & = & \text{the energy value of tissue lost or gained (MK/kg)} \\
W & = & \text{liveweight (kg)} \\
\Delta W & = & \text{liveweight change (kg/d)}
\end{array}
$$

Comparing models for cattle growth

General

The daily energy balance in growing cattle may be represented as follows:

$$MEI = \frac{E_m}{k_m} + \frac{E_g}{k_g} \tag{1}$$

writing $E_g = EV_g \times \Delta W$ we obtain

$$MEI = \frac{E_m}{k_m} + \frac{EV_g \times \Delta W}{k_g} \tag{2}$$

Comparing models for cattle growth

MODEL 1: (*AFRC 1980*)

$$
\begin{aligned}
E_m &= 0.53(W/1.08)^{0.67} + 0.0043W \\
k_m &= 0.35q + 0.503 \\
k_f &= 0.78q + 0.006 \\
k_g &= \frac{k_f}{L-1} \\
EV_g &= \frac{(4.1 + 0.0332W - 0.000009W^2)}{1 - 0.1475\Delta W}
\end{aligned}
$$

$$
\Delta W = \frac{E_g}{4.1 + 0.0332W - 0.000009W2 + 0.1475E_g}
$$

$$
E_g = k_g \times (MEI - \frac{E_m}{k_m}).
$$

Comparing models for cattle growth

MODEL 2: ($TB33$)

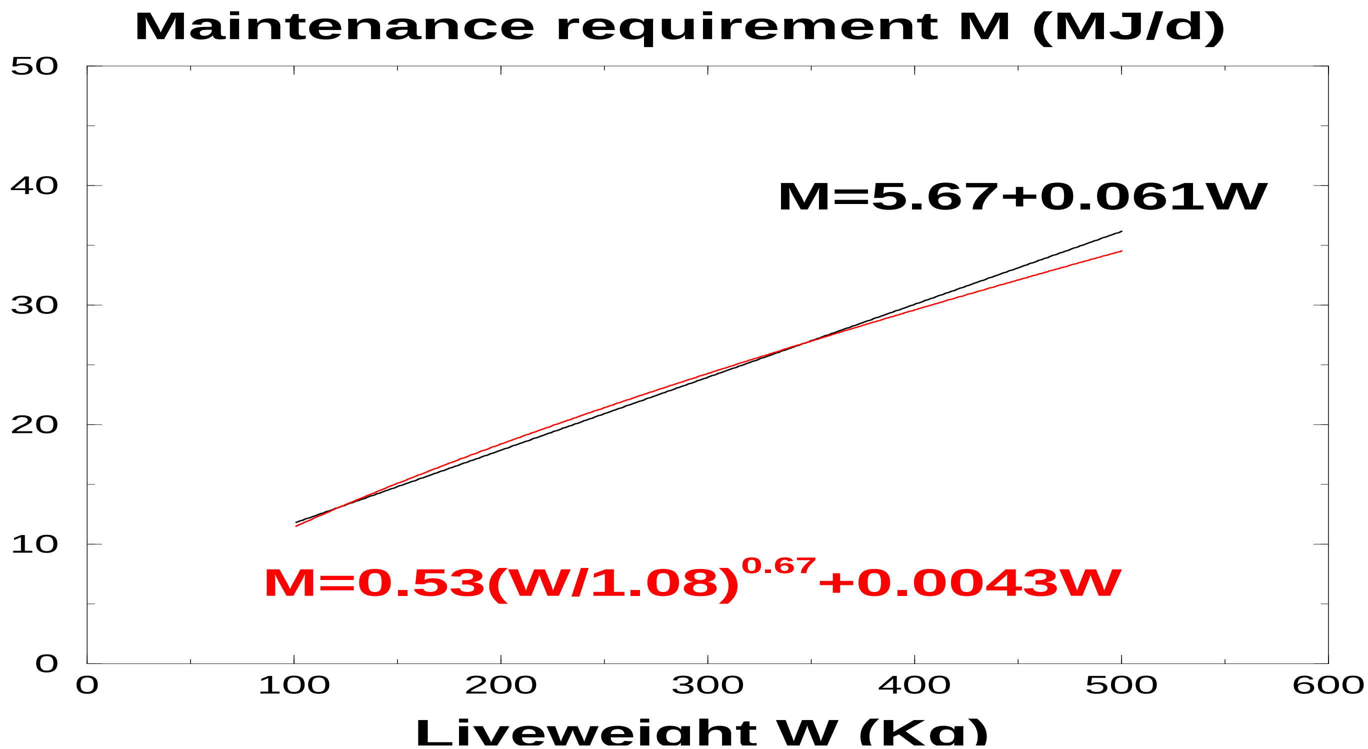
$$
\begin{aligned}
E_m &= 5.67 + 0.061W \\
k_m &= 0.72 \\
k_g &= 0.9q \\
EV_g &= 6.28 + 0.3E_g + 0.0188W
\end{aligned}
$$

Rearranging (2) and substituting for EV_g gives

$$\Delta W = E_g/(6.28 + 0.0188W + 0.3E_g)$$

where $E_g = k_g \times (MEI - E_m/k_m)$.

Comparing models for cattle growth



Comparing models for cattle growth

Table A
Predictions of liveweight gains (g/d) according to the competing models

W (kg)	100				500			
MEI(MJ/d)	20		40		60		100	
q	.46	.68	.57	.68	.46	.68	.57	.68
Model 1	170	351	1112	1322	197	408	1044	1239
Model 2	157	226	947	1070	215	307	1010	1137

Comparing models for cattle growth

Table B
Bias in predicting liveweight gain (g/d) using independent data (standard deviations in parentheses)

	Data Set	Mean liveweight gain (g/d)	Model 1	Model 2
1.	(Food, Reading)	1080	210(80)	130(80)
2*a*.	(Hinks, Edinburgh)	660	180(100)	150(100)
2*b*.		890	130(180)	100(170)
2*c*.		820	70(120)	50(120)
2*d*.		970	160(130)	220(130)
3.	(Drayton EHF)	730	−10(120)	0(120)
4.	(MLC, Nottingham)	910	170(220)	140(210)

Working Party Report

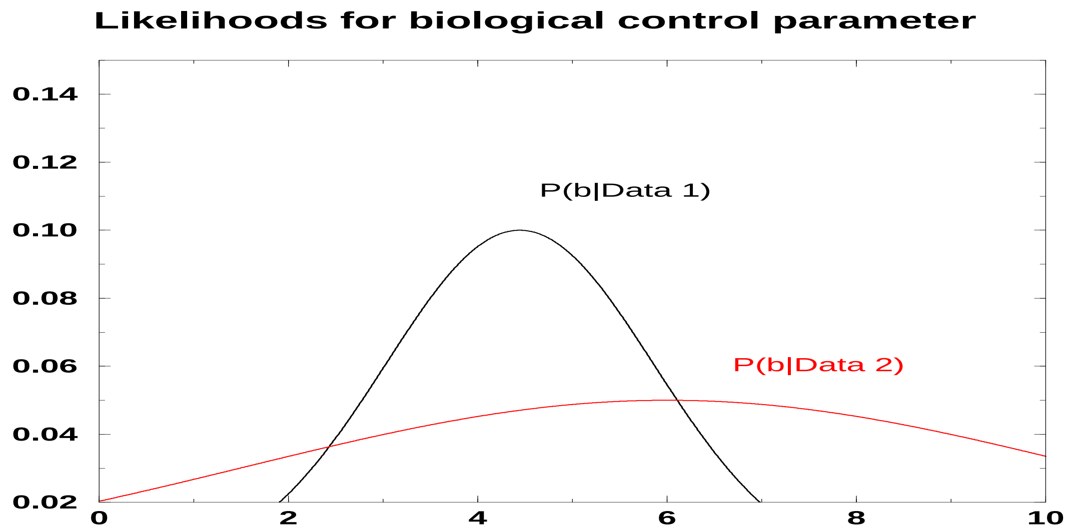
Concludes there is little difference in predictive ability
Recommends Model 1 because:

a) no need for linearizing approximations of Model 2

b) includes terms for known effects absent in Model 1

c) better platform for future development

Comparison of two models via precision of parameter estimates



Treatment 1 supports model including the biocontrol parameter b
Treatment 2 supports simpler model without b