## Simulation: Solving Dynamic Models
### ABE 5646
#### Week 11 Chapter 2, Spring 2010

<table>
<thead>
<tr>
<th>Week</th>
<th>Description</th>
<th>Reading Material</th>
</tr>
</thead>
</table>
  • Comparing a model with data  
    - Graphical, errors  
    - Measures of agreement (bias in mean, variance)  
    - Evaluation of predictive quality  
  • Cross validation  
  • Bootstrap estimation  
  • Effects of errors in observations on MSEP | Wallach et al. (2006)  
Chapter 1, 2  
Also see chapters 12-13 for examples |
Weeks 11-Ch 2 Objectives

1) Learn basic methods for evaluating dynamic biophysical models
Shifting Focus

• From:
  – System modeling and simulation methods

• To:
  – Working with dynamic models (with crop examples, but methods apply to other biological and agricultural system models)

This shift means that we take a model that has already been developed and work with it

In our class, we will focus on model 1) evaluation (comparison with data from the real system), 2) analyzing uncertainty and sensitivity of a model to various factors, and 3) estimating parameters and their uncertainty.

We will not attempt to cover all methods in the book in the remaining time, but instead cover 2 or 3 of the most important methods and concepts for each topic that we do cover.
Evaluating Dynamic Models

• Evaluation vs. validation
• Importance
• Some steps
  – Objective of model, criteria for evaluation
  – Model equations, implementation accuracy
  – Sensitivity analysis, parameter estimation
  – Comparison with data
• Methods
Evaluating Model Agreement with Measurements

• Test whether model is a good representation of the real system
• In many cases, measurements are used to estimate parameters, similar to fitting a model to particular experimental results
• The main issue here is whether measurements used in an evaluation are independent
• When the predictive quality of a model is to be evaluated, comparisons must be with data not used to develop or parameterize the model. This will be considered after examining measures for comparisons
Model Error Analysis

Simple Example

Regression Model

How does this relate to dynamic models?

Example 1

Suppose that our model for predicting response \( Y \) for individual \( i \) is

\[
\hat{Y}_i = f(X_i; \theta) = \hat{\theta}^{(0)} + \hat{\theta}^{(1)} x_i^{(1)} + \hat{\theta}^{(2)} x_i^{(2)} + \hat{\theta}^{(3)} x_i^{(3)} + \hat{\theta}^{(4)} x_i^{(4)} + \hat{\theta}^{(5)} x_i^{(5)}
\]

(1)

The explanatory variables are \( X = (x^{(1)}, x^{(2)}, x^{(3)}, x^{(4)}, x^{(5)})^T \). The parameter vector is

\[
\hat{\theta} = (\hat{\theta}^{(0)}, \hat{\theta}^{(1)}, \hat{\theta}^{(2)}, \hat{\theta}^{(3)}, \hat{\theta}^{(4)}, \hat{\theta}^{(5)})^T = (1.9, 7.8, 2.5, -0.2, 0.1, 0.7)^T
\]

(2)

The hat notation is used to indicate that the parameter values are estimates. We do not need to bother here with the origin of these estimates. The data set for evaluating the model is given in Table 1.

Table 1. Measured values \( (Y_i) \), 5 explanatory variables, calculated values \( (\hat{Y}_i) \) and model errors \( (D_i) \) for 8 situations.

<table>
<thead>
<tr>
<th>( i )</th>
<th>( Y_i )</th>
<th>( x_i^{(1)} )</th>
<th>( x_i^{(2)} )</th>
<th>( x_i^{(3)} )</th>
<th>( x_i^{(4)} )</th>
<th>( x_i^{(5)} )</th>
<th>( \hat{Y}_i )</th>
<th>( D_i ) = ( Y_i - \hat{Y}_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-9.39</td>
<td>-1.63</td>
<td>0.80</td>
<td>0.44</td>
<td>-0.45</td>
<td>-0.47</td>
<td>-9.31</td>
<td>-0.08</td>
</tr>
<tr>
<td>2</td>
<td>-3.23</td>
<td>-0.95</td>
<td>1.07</td>
<td>0.50</td>
<td>0.53</td>
<td>-0.33</td>
<td>-3.10</td>
<td>-0.13</td>
</tr>
<tr>
<td>3</td>
<td>0.37</td>
<td>-0.25</td>
<td>0.20</td>
<td>0.51</td>
<td>0.82</td>
<td>0.45</td>
<td>0.72</td>
<td>-0.05</td>
</tr>
<tr>
<td>4</td>
<td>7.10</td>
<td>0.38</td>
<td>1.02</td>
<td>-0.22</td>
<td>0.82</td>
<td>-1.93</td>
<td>6.18</td>
<td>0.92</td>
</tr>
<tr>
<td>5</td>
<td>5.82</td>
<td>0.15</td>
<td>1.14</td>
<td>1.07</td>
<td>1.63</td>
<td>-0.55</td>
<td>5.45</td>
<td>0.38</td>
</tr>
<tr>
<td>6</td>
<td>-11.21</td>
<td>-1.20</td>
<td>-1.79</td>
<td>0.35</td>
<td>-0.26</td>
<td>0.26</td>
<td>-11.84</td>
<td>0.63</td>
</tr>
<tr>
<td>7</td>
<td>-5.81</td>
<td>-0.97</td>
<td>-0.15</td>
<td>0.11</td>
<td>1.13</td>
<td>0.10</td>
<td>-5.90</td>
<td>0.09</td>
</tr>
<tr>
<td>8</td>
<td>2.82</td>
<td>0.39</td>
<td>-1.20</td>
<td>2.01</td>
<td>-0.79</td>
<td>-1.44</td>
<td>0.47</td>
<td>2.35</td>
</tr>
</tbody>
</table>
Comparisons using Graphics

Predicted vs. Observed

Figure 1. Calculated versus measured values, using the model and data of Table 1.

Error (residue) = Predicted – Observed

Figure 2. Residues for model and data of Table 1.
Comparisons using Graphs vs. Time

Simulated and Measured Soybean

Gainesville, FL
1978

Yields

Dry Weight (kg/ha)

Day of Year

Total Crop - IRRIGATED

Grain - IRRIGATED

Total Crop - NOT IRRIGATED

Grain - NOT IRRIGATED
Graphing Final Yield
Predicted vs. Observed

On farm crop model tests in Pampas Region, Argentina. Magrin et al.
Simple Measures of Agreement

*Difference between observation and prediction*

\[ D_i = Y_i - \hat{Y}_i \]

and

\[ Bias = \frac{1}{N} \sum_{i=1}^{N} D_i \]

\[ \text{Mean Square Error} = \text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (D_i)^2 \]

\[ \text{Root Mean Square Error} = \text{RMSE} = \sqrt{\text{MSE}} \]
Simple Measures of Agreement

Mean Absolute Error = \( \frac{1}{N} \sum_{i=1}^{N} |D_i| \)

Relative Root Mean Square Error = \( \frac{RMSE}{\bar{Y}} \)

with \( \bar{Y} = \frac{1}{N} \sum_{i=1}^{N} Y_i \)
Modeling Efficiency

\[ EF = 1 - \frac{\sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{N} (Y_i - \bar{Y}_i)^2} \]
Correlation Coefficient

\[ r = \frac{\hat{\sigma}_{Y\hat{Y}}}{\hat{\sigma}_Y \hat{\sigma}_{\hat{Y}}} \]

[there was a mistake in book] \hspace{1cm} (10)

where \( \hat{\sigma}_Y^2 \), \( \hat{\sigma}_{\hat{Y}}^2 \) and \( \hat{\sigma}_{Y\hat{Y}} \) are sample estimates of the variance of \( Y \), the variance of \( \hat{Y} \) and the covariance of \( Y \) and \( \hat{Y} \) respectively.

\[
\hat{\sigma}_Y^2 = \frac{1}{N} \sum_{i=1}^{N} [(Y_i - \bar{Y})^2]
\]

\[
\hat{\sigma}_{\hat{Y}}^2 = \frac{1}{N} \sum_{i=1}^{N} [(\hat{Y}_i - \bar{Y})^2]
\]

\[
\hat{\sigma}_{Y\hat{Y}} = \frac{1}{N} \sum_{i=1}^{N} [(Y_i - \bar{Y})(\hat{Y}_i - \bar{Y})]
\]

\[
\bar{Y} = \frac{1}{N} \sum_{i=1}^{N} \hat{Y}_i
\]
Willmott D Index

\[
index = 1 - \frac{\sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{N} (|\hat{Y}_i - \bar{Y}| + |Y_i - \bar{Y}|)^2}
\]
Decomposing Error due to Different Sources

Decomposing MSE (Kobayashi and Salam (2000):

\[
MSE = (Bias)^2 + SDSD + LCS
\]

with

\[
SDSD = (\sigma_Y - \sigma_{\hat{Y}})^2
\]

\[
LCS = 2\sigma_Y \sigma_{\hat{Y}} (1 - r) \quad \text{or} \quad MSE - (Bias)^2 - SDSD
\]

and

SDSD is due to differences between standard deviation of measurements vs. modeled

LCS is related to how well the model mimics observed variation across situations

There are other ways to decompose the error (i.e., Gauch et al. (2003), Willmott (1981)
Summary of Measures

<table>
<thead>
<tr>
<th>Name</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias</td>
<td>$\text{Bias} = \frac{1}{N} \sum_{i=1}^{N} D_i$</td>
</tr>
<tr>
<td>Mean squared error</td>
<td>$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (D_i)^2$</td>
</tr>
<tr>
<td>Root mean squared error</td>
<td>$\text{RMSE} = \sqrt{\text{MSE}}$</td>
</tr>
<tr>
<td>Mean absolute error</td>
<td>$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N}</td>
</tr>
<tr>
<td>Relative root mean squared error</td>
<td>$\text{RRMSE} = \frac{\text{RMSE}}{\bar{Y}}$</td>
</tr>
<tr>
<td>Relative mean absolute error</td>
<td>$\text{RMAE} = \frac{1}{N} \sum_{i=1}^{N} \frac{</td>
</tr>
<tr>
<td>Modeling efficiency</td>
<td>$\text{EF} = 1 - \frac{\sum_{i=1}^{N} (Y_i - \hat{Y}<em>i)^2}{\sum</em>{i=1}^{N} (Y_i - \bar{Y})^2}$</td>
</tr>
<tr>
<td>Correlation coefficient</td>
<td>$r = \frac{\sum_{i=1}^{N} (Y_i - \bar{Y})(\hat{Y}<em>i - \bar{Y})}{\sqrt{\sum</em>{i=1}^{N} (Y_i - \bar{Y})^2} \sum_{i=1}^{N} (\hat{Y}_i - \bar{Y})^2}}$</td>
</tr>
<tr>
<td>Agreement index</td>
<td>index $= 1 - \frac{\sum_{i=1}^{N} (Y_i - \hat{Y}<em>i)^2}{\sum</em>{i=1}^{N} (</td>
</tr>
<tr>
<td>Concordance correlation coefficient</td>
<td>$\rho_c = \frac{2\sigma_{Y\hat{Y}}}{\sigma_Y^2 + \sigma_{\hat{Y}}^2 + (\mu_{\hat{Y}} - \mu_Y)^2}$</td>
</tr>
<tr>
<td>Total deviation index</td>
<td>$\text{TDI}(p) = \text{minimal value of } d \text{ such that }</td>
</tr>
<tr>
<td>Coverage probability</td>
<td>$\text{CP}(d) = \text{the smallest value of } p \text{ such that, for a percentage } p \text{ of the observed situations, }</td>
</tr>
</tbody>
</table>
Evaluating the Predictive Quality of a Model

- When parameters are fixed, a model can be used to predict independent observations
- MSEP is the Mean Square Error of Prediction
- MSEP is not equal to MSE
  - MSE may include data that were used to develop a model or to estimate its parameters
  - Data used to compute MSE may not represent the full range of interest for the model purposes
  - MSE may be a poor estimator of MSEP
MSEP

\[ MSEP(\hat{\theta}) = E\{[Y - f(X; \hat{\theta})|\hat{\theta}]^2 \} \]

\[ RMSEP(\hat{\theta}) = \sqrt{MSEP(\hat{\theta})} \]
Figure 3. Response $Y$ as a function of a single explanatory variable $x$. For 3 specific $x$ values, the variability of $Y$ is shown. The solid line is $E(Y|x)$ and the dashed line a hypothetical model.

We can develop the mean squared error of prediction as

$$MSEP(\hat{\theta}) = E\left\{ [Y - E(Y|X) + E(Y|X) - f(X; \theta)]^2 \right\}$$

$$= E_X\left\{ E_Y\{ [Y - E_Y(Y|X) + E_Y(Y|X) - f(X; \hat{\theta})]|X\}^2 \right\}$$

$$= E_X\left\{ E_Y\{ [Y - E_Y(Y|X)]^2 \} + E_X\{ [E_Y(Y|X) - f(X; \hat{\theta})]|X\}^2 \right\}$$

$$= \Lambda + \Delta \quad (19)$$

where

$$\Lambda = E_X\{E_Y\{[Y - E_Y(Y|X)]^2]\} = E_X [\text{var}(Y|X)] = \text{population variance} \quad (20)$$

$$\Delta = E_X\{[E_Y(Y|X) - f(X; \hat{\theta})]^2\} = \text{squared bias} \quad (21)$$
Estimating MSEP(\( \hat{\theta} \))

Data are independent from model & parameters

\[
\hat{MSEP}(\hat{\theta}) = MSE = \frac{1}{N} \sum_{i=1}^{N} [Y_i - f(X_i; \hat{\theta})]^2
\]

If different data are used to estimate MSEP, one can also compute an estimate of the variance of MSEP (eq. 23, p.34)
Cross Validation

- Method for estimating $\hat{\theta}$, uncertainty in prediction
- Useful when number of observations is small
- Assume data are random sample from target population
- Could split data into two parts, one for parameter estimation and one for testing (independent)
- But, lose important information for estimating parameter
- Cross Validation also uses data splitting
  - First, using all but one data point to estimate parameters
  - Then, calculate error for the one data point that was left out
  - Repeat this N times (# data points); end up with N errors
  - This approach uses all data to estimate parameters & evaluate predictions
MSEP computed by Cross Validation is:

\[ \hat{MSEP}_{CV}(\hat{\theta}) = \frac{1}{N} \sum_{i=1}^{N} [Y_i - f(X_i; \hat{\theta}_{-i})]^2 \]

Where the \( \hat{\theta}_i \) is the set of parameters leaving out data point \( i \)
Cross Validation Example

• Using the simple model (5 inputs, 6 parameters):

\[ f_5(X; \hat{\theta}) = \hat{\theta}^{(0)} + \hat{\theta}^{(1)} x^{(1)} + \hat{\theta}^{(2)} x^{(2)} + \hat{\theta}^{(3)} x^{(3)} + \hat{\theta}^{(4)} x^{(4)} + \hat{\theta}^{(5)} x^{(5)} \]

• And the data from Table 1 (N=8 data points)

<table>
<thead>
<tr>
<th>i</th>
<th>( Y_i )</th>
<th>( x_i^{(1)} )</th>
<th>( x_i^{(2)} )</th>
<th>( x_i^{(3)} )</th>
<th>( x_i^{(4)} )</th>
<th>( x_i^{(5)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-9.39</td>
<td>-1.63</td>
<td>0.80</td>
<td>0.44</td>
<td>-0.45</td>
<td>-0.47</td>
</tr>
<tr>
<td>2</td>
<td>-3.23</td>
<td>-0.95</td>
<td>1.07</td>
<td>0.50</td>
<td>0.53</td>
<td>-0.33</td>
</tr>
<tr>
<td>3</td>
<td>0.37</td>
<td>-0.25</td>
<td>0.20</td>
<td>0.51</td>
<td>0.82</td>
<td>0.45</td>
</tr>
<tr>
<td>4</td>
<td>7.10</td>
<td>0.38</td>
<td>1.02</td>
<td>-0.22</td>
<td>0.82</td>
<td>-1.93</td>
</tr>
<tr>
<td>5</td>
<td>5.82</td>
<td>0.15</td>
<td>1.14</td>
<td>1.07</td>
<td>1.63</td>
<td>-0.55</td>
</tr>
<tr>
<td>6</td>
<td>-11.21</td>
<td>-1.20</td>
<td>-1.79</td>
<td>0.35</td>
<td>-0.26</td>
<td>0.26</td>
</tr>
<tr>
<td>7</td>
<td>-5.81</td>
<td>-0.97</td>
<td>-0.15</td>
<td>0.11</td>
<td>1.13</td>
<td>0.10</td>
</tr>
<tr>
<td>8</td>
<td>2.82</td>
<td>0.39</td>
<td>-1.20</td>
<td>2.01</td>
<td>-0.79</td>
<td>-1.44</td>
</tr>
</tbody>
</table>
Cross Validation Example

Table 4. Calculation of cross-validation estimate of $MSEP(\hat{\theta})$ for model $f_5(X; \hat{\theta})$.

<table>
<thead>
<tr>
<th>$i$</th>
<th>Data used for parameter adjustment</th>
<th>$Y_i$</th>
<th>$f_5(X_i; \hat{\theta}_{-i})$</th>
<th>$[Y_i - f_5(X_i; \hat{\theta}_{-i})]^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$Y_2, Y_3, Y_4, Y_5, Y_6, Y_7, Y_8$</td>
<td>-9.39</td>
<td>9.44</td>
<td>1.06</td>
</tr>
<tr>
<td>2</td>
<td>$Y_1, Y_3, Y_4, Y_5, Y_6, Y_7, Y_8$</td>
<td>-3.23</td>
<td>11.96</td>
<td>0.97</td>
</tr>
<tr>
<td>3</td>
<td>$Y_1, Y_2, Y_4, Y_5, Y_6, Y_7, Y_8$</td>
<td>0.37</td>
<td>-2.45</td>
<td>0.12</td>
</tr>
<tr>
<td>4</td>
<td>$Y_1, Y_2, Y_3, Y_5, Y_6, Y_7, Y_8$</td>
<td>7.10</td>
<td>1.01</td>
<td>0.00</td>
</tr>
<tr>
<td>5</td>
<td>$Y_1, Y_2, Y_3, Y_4, Y_6, Y_7, Y_8$</td>
<td>5.82</td>
<td>-8.33</td>
<td>0.03</td>
</tr>
<tr>
<td>6</td>
<td>$Y_1, Y_2, Y_3, Y_4, Y_5, Y_7, Y_8$</td>
<td>-11.21</td>
<td>9.14</td>
<td>0.15</td>
</tr>
<tr>
<td>7</td>
<td>$Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_8$</td>
<td>-5.81</td>
<td>-8.48</td>
<td>0.06</td>
</tr>
<tr>
<td>8</td>
<td>$Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_7$</td>
<td>2.82</td>
<td>-4.04</td>
<td>0.10</td>
</tr>
</tbody>
</table>

$\hat{MSEP}_{CV}(\hat{\theta})$ 0.31

What about values of the parameters?
Bootstrap Estimation

- Data re-sampling for estimating $\text{MSEP}(\hat{\theta})$
- Useful when number of observations is small
- Assume data are the full target population
- Could split data into two parts, one for parameter estimation and one for testing (independent)
- But, lose important information for estimating parameter
- Bootstrap also uses data splitting
  - Sampling with replacement
  - $N=$ number of data points, pick a data point, replace it, then pick a second data point, then a $3^{rd}$, etc. to get $N$ data points
  - Repeat this $b$ times
Bootstrap Estimation Example

Let $N = 8$, using data from CV example
Let $B = 3$ bootstrap samples
Then we have 3 samples of 8 observations each
Estimate parameters for first sample ($b=1$)
Then use those parameters to compute $\text{MSEP}_b$ for the original 8 samples AND to estimate MSE using the 8 data points in sample $b$
Compute a correction term for sample $b$
Repeat this for $b=2$, then $b=3$
Average the correction term and add it to MSE computed from the original dataset
Bootstrap Estimation of MSEP

Correction term:

\[ op_b = MSEP_b(\hat{\theta}_b) - MSE \]

by first computing \( MSEP_b \) and \( MSE_b \)

\[ MSEP_b(\hat{\theta}_b) = \frac{1}{N} \sum_{i=1}^{N} [Y_i - f(X_i;\hat{\theta}_b)]^2 \]

\[ MSE_b = \frac{1}{N} \sum_{i=1}^{N} [Y_{bi} - f(X_{bi};\hat{\theta}_b)]^2 \]
Bootstrap Estimation of MSEP

\[ \hat{MSEP}_{bootstrap}(\hat{\theta}) = MSE + \hat{op} \]

where

\[ \hat{op} = \frac{1}{B} \sum_{b=1}^{B} op_b \]
Measurement Errors in Y and MSEP

- If Y has significant measurement error, then there are really two MSEP values
  - one for the difference between predicted and measured values
  - the other for the difference between predicted and the TRUE value

\[ Y_{\text{obs}} = Y + \eta \]

\(\eta\) is measurement error, mean 0
MSEP with Errors in Measurements

\[ \text{MSEP}^\text{obs}(\hat{\theta}) = \sigma^2_\eta + \text{MSEP}(\hat{\theta}) \]

and

\[ \text{MSEP}(\hat{\theta}) = \text{MSEP}^\text{obs}(\hat{\theta}) - \hat{\sigma}^2_\eta \]

if \( \eta \) is independent of \( Y \) and \( X \); \( \eta \sim (0, \sigma^2_\eta) \)
Discussion

• Random sampling for stochastic models
• Comparing a model with data
  - Graphical, errors
  - Measures of agreement (bias in mean, variance)
  - Evaluation of predictive quality
• Cross validation
• Bootstrap
• Errors in measurements
Homework Chapter 2

• Chapter 2 in Wallach et al.
  – Problem 1
  – Problem 4

• Due on March 25