Phase I Monitoring of Nonlinear Profiles

James D. Williams

William H. Woodall

Jeffrey B. Birch

May 22, 2003
Profile Monitoring

Scenario

- Monitor a process or product whose quality cannot be assessed by a single quality characteristic
- Measure across some continuum (a sequence of time, space, etc.) producing a “profile”
- Various profile shapes:
  - Linear Profiles: (Kang and Albin (2000), Kim, Mahmoud, and Woodall (2003), Mahmoud and Woodall (2003))
  - Nonlinear Profiles: (Brill (2001))
- Very little work has been done to address monitoring nonlinear profiles (Woodall, et. al. (2003))
Profile Monitoring

Path forward

- Brill’s (2001) method
- Suggest two more methods
- Illustrate methods with nonlinear profile data
- Recommendations
Example 1: Vertical Density Profile (VDP)

Board A1 from Walker and Wright (2002, *JQT*)
Example 2: Dose-Response Profile of a Drug
Phase I Analysis: Historical Data

Response

Nonlinear Profile

<table>
<thead>
<tr>
<th>Sample</th>
<th>1</th>
<th>2</th>
<th>...</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$y_{1,1}$</td>
<td>$y_{1,2}$</td>
<td>...</td>
<td>$y_{1,n}$</td>
</tr>
<tr>
<td>2</td>
<td>$y_{2,1}$</td>
<td>$y_{2,2}$</td>
<td>...</td>
<td>$y_{2,n}$</td>
</tr>
<tr>
<td>:</td>
<td>:</td>
<td>:</td>
<td>:</td>
<td>:</td>
</tr>
<tr>
<td>m</td>
<td>$y_{m,1}$</td>
<td>$y_{m,2}$</td>
<td>...</td>
<td>$y_{m,n}$</td>
</tr>
</tbody>
</table>
Brief Intro to Nonlinear Regression Models

Simple Case: One Y and one X

\[ y_i = f(x_i, \beta) + \varepsilon_i \quad i = 1, \ldots, n \]

where

- \( y_i \) is the \( i^{th} \) response
- \( f(x_i, \beta) \) is an appropriate nonlinear function
- \( x_i \) is the \( i^{th} \) regressor variable value
- \( \beta \) is the \( p \times 1 \) vector of parameters to estimate
- \( \varepsilon_i \) is the \( i^{th} \) residual error
Brief Intro to Nonlinear Regression Models

\[ \hat{\beta}_i \quad \text{obtained iteratively for each sample} \]

\[ Var(\hat{\beta}_i) = \hat{\sigma}^2 \left( \hat{D}_i' \hat{D}_i \right)^{-1} = C_i \]

where \( \hat{D}_i \) is the estimated derivative matrix used in the estimation of the nonlinear regression parameters
## Parameter Estimates from Historical Data

<table>
<thead>
<tr>
<th>Sample</th>
<th>Parameter</th>
<th>( \hat{\beta}_{1,1} )</th>
<th>( \hat{\beta}_{1,2} )</th>
<th>( \cdots )</th>
<th>( \hat{\beta}_{1,p} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \hat{\beta}_{2,1} )</td>
<td>( \hat{\beta}_{2,2} )</td>
<td>( \cdots )</td>
<td>( \hat{\beta}_{2,p} )</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td></td>
</tr>
<tr>
<td>( m )</td>
<td>( \hat{\beta}_{m,1} )</td>
<td>( \hat{\beta}_{m,2} )</td>
<td>( \cdots )</td>
<td>( \hat{\beta}_{m,p} )</td>
<td></td>
</tr>
</tbody>
</table>
How to Monitor Nonlinear Profiles

- Ideally, monitor each parameter independently
- **Problem**: parameter estimates are correlated in nonlinear regression
- Cannot monitor each parameter separately, so use a multivariate $T^2$ control chart to monitor the parameters simultaneously
Multivariate $T^2$ Control Chart Statistic

General form of the $T^2$ statistic:

$$T_i^2 = \left( \hat{\beta}_i - \bar{\beta} \right) S^{-1} \left( \hat{\beta}_i - \bar{\beta} \right)$$

$i = 1, \ldots, m$

$S$ is the covariance matrix of parameter estimates

$$\hat{\beta}_i = \begin{pmatrix} \hat{\beta}_{i,1} \\ \hat{\beta}_{i,2} \\ \vdots \\ \hat{\beta}_{i,p} \end{pmatrix}$$

$$\bar{\beta} = \begin{pmatrix} \frac{1}{m} \sum_{i=1}^{m} \hat{\beta}_{i,1} \\ \frac{1}{m} \sum_{i=1}^{m} \hat{\beta}_{i,2} \\ \vdots \\ \frac{1}{m} \sum_{i=1}^{m} \hat{\beta}_{i,p} \end{pmatrix}$$
Three Choices for $S$

Method 1: Sample Covariance Matrix (Brill, 2001)

$$S_1 = \frac{1}{m-1} \sum_{i=1}^{m} \left( \hat{\beta}_i - \bar{\beta} \right) \times \left( \hat{\beta}_i - \bar{\beta} \right)'$$

Pros:
- Easy to calculate
- Widely used and easily understood

Cons:
- Greatly affected by shifts in mean vector
- Results in low power for the $T^2$ control chart
Three Choices for S

Method 2: Successive Differences (Holmes and Mergen, 1993)

Let \( \mathbf{v}_i = \hat{\beta}_{i+1} - \hat{\beta}_i \quad i = 1, \ldots, m - 1 \)

\[
\mathbf{v} = \begin{bmatrix}
\mathbf{v}_1' \\
\mathbf{v}_2' \\
\vdots \\
\mathbf{v}_{m-1}'
\end{bmatrix}
\]

Then \( S_2 = \frac{\mathbf{V}'\mathbf{V}}{2(m-1)} \)

Pros:
- Like moving range with individual observations
- Not effected by shifts in the mean vector
- High power

Cons:
- Less statistical theory developed to date
Three Choices for S

Method 3: Intra-Profile Pooling

For each of the $m$ samples: \[
\text{Var}(\hat{\beta}_i) = \hat{\sigma}^2 \left( \hat{D}_i' \hat{D}_i \right)^{-1} = C_i
\]

Then \[
S_3 = \frac{1}{m} \sum_{i=1}^{m} C_i
\]

Pros: • Uses information from nonlinear regression estimation

Cons: • Does not account for profile-to-profile common cause variability
Three Choices for $T_i^2$

Three formulations of the $T^2$ statistic:

Method 1: Sample Covariance Matrix

$$T_{1,i}^2 = \left( \hat{\beta}_i - \bar{\beta} \right)' S_1^{-1} \left( \hat{\beta}_i - \bar{\beta} \right)$$

Method 2: Successive Differences

$$T_{2,i}^2 = \left( \hat{\beta}_i - \bar{\beta} \right)' S_2^{-1} \left( \hat{\beta}_i - \bar{\beta} \right)$$

Method 3: Intra-Profile Pooling

$$T_{3,i}^2 = \left( \hat{\beta}_i - \bar{\beta} \right)' S_3^{-1} \left( \hat{\beta}_i - \bar{\beta} \right)$$
Upper Control Limits

Method 1: Sample Covariance Matrix

\[ T_1^2 \frac{m}{(m-1)^2} \sim \text{Beta} \left( \frac{p}{2}, \frac{m-p-1}{2} \right) \]

As discussed by Sullivan and Woodall (1996)

\[ UCL_1 = \frac{(m-1)^2}{m} B_{1-\alpha, p/2, (m-p-1)/2} \]
Upper Control Limits

Method 2: Successive Differences

Approximately

\[ T_2^2 \frac{m}{(m-1)^2} \sim \text{Beta} \left( \frac{p}{2}, \frac{f-p-1}{2} \right) \]

where

\[ f = \frac{2(m-1)^2}{3m - 4} \]

For more information, see Scholz and Tosch (1994)

\[ UCL_2 = \frac{(m-1)^2}{m} B_{1-\alpha, p/2, (f-p-1)/2} \]
Upper Control Limits

Method 3: Intra-Profile Pooling

We think that approximately

\[ T_3^2 \frac{m(m - p)}{p(m - 1)(m + 1)} \sim F(p, m - p) \]

\[ UCL_3 = \frac{m(m - p)}{p(m - 1)(m + 1)} F_{1-\alpha, p, m-p} \]

Control limits are best approximations so far
Illustration: VDP Data

VDP of 24 Particle Boards
Nonlinear Function to Model VDP Data

Use a “bathtub” function to model each board from the VDP data

\[
f(x_i, \beta) = \begin{cases} 
    a_1(x_i - d)^{b_1} + c & x_i > d \\
    a_2(-x_i + d)^{b_2} + c & x_i \leq d 
\end{cases}
\]

where \( x_i \) is the \( i \)th regressor variable value

\[
\beta = \begin{pmatrix} 
    a_1 \\
    a_2 \\
    b_1 \\
    b_2 \\
    c \\
    d 
\end{pmatrix}
\]

determine the width of the “bathtub”

determine the “flatness” of the “bathtub”

\( c \rightarrow \) is the bottom of the “bathtub”

\( d \rightarrow \) is the center of the “bathtub”
Nonlinear Function to Model VDP Data

Board #1 from Walker and Wright (2002, JQT)
Nonlinear Function to Model VDP Data

Estimated nonlinear profile of Board #1

\[
f(x_i, \hat{\beta}) = \begin{cases} 
5708(x_i - 0.313)^{5.14} + 46.0 & x_i > 0.313 \\
3921(-x_i + 0.313)^{4.87} + 46.0 & x_i \leq 0.313 
\end{cases}
\]

- Estimate profile for each board
- Calculate \( S_1 \), \( S_2 \), and \( S_3 \).
- Calculate \( T_1^2 \), \( T_2^2 \), and \( T_3^2 \).
$T_1^2$ Control Chart

$T_1^2$ UCL = 9.6
$T^2_2$ Control Chart

$T^2_2$ UCL = 14.1

Board #15

Board #18
$T^2_1$ and $T^2_2$ Control Charts

$T^2_2$ UCL = 14.1

$T^2_1$ UCL = 9.6
According to our UCL, all of the boards are out-of-control.
Conclusions

- Method 1 (sample covariance matrix) does not take into account the sequential sampling structure of the data:
  - The overall probability of detecting a shift in the mean vector will decrease (See Sullivan and Woodall, 1996)
  - Should not be used

- Method 2 (successive differences) accounts for the sequential sampling scheme, and gives a more robust estimate of the covariance matrix

- In the VDP example, both Methods 1 and 2 gave same result because
  - No apparent shift in the mean vector
  - There were only about two outliers
Conclusions

- Method 3 (intra-profile pooling) should be used when there is no profile-to-profile common cause variability.

- Comparison of the three methods:
  - Method 1 assumes all variability is due to common cause.
  - Method 3 assumes that no variability is due to common cause.
  - Method 2 is somewhere in the middle.

Issue: Monitoring parameters versus monitoring the fitted curves.
References


