Prediction Limits vs. Control Charts

Kirk Cameron, Ph.D.
kcmacstat@gmail.com
719-532-0453

Just For Fun

“Forecasting is like trying to drive a car blindfolded while following directions given by someone looking out of the rear window!”
— Unknown

Prediction Limit Overview

- Core method for detection monitoring
- Prediction limits estimate likely range of compliance concentrations based on observed BG
- Compliance values outside range suggest possible contamination

Flexibility

- Interwell or intrawell
- Parametric or non-parametric
- Parametric uses normalized measurements
- Non-parametric uses order statistics
Key Benefits

- Prediction limits can be designed to test individual new values from each compliance well
  - Typical use: 1 new measurement from each well
  - Fits modern sampling schedules
  - Good for sites with slow groundwater flow, autocorrelation

Benefits (cont.)

- Well-studied:
  - Can be applied — in conjunction with retesting — to either small or large monitoring networks
  - Known site-wide false positive rates (SWFPR) and statistical power

Prediction Limit Cons

- Not always as powerful as other tests (ANOVA, control charts)
  - Tradeoff between power, sampling requirements, & ease of use/interpretation
  - Retesting helps a good bit

Parametric Example

- Prediction limit for $m$ individual observations

- Suppose one has $n = 10$ normal BG measurements
  - Want to predict $m = 4$ compliance values with 95% confidence

$$PL = \bar{x} + \kappa s$$
**Arsenic Example**

<table>
<thead>
<tr>
<th>Event</th>
<th>Conc (ppb)</th>
<th>Event</th>
<th>Conc (ppb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.7</td>
<td>11</td>
<td>9.3</td>
</tr>
<tr>
<td>2</td>
<td>14.1</td>
<td>12</td>
<td>10.8</td>
</tr>
<tr>
<td>3</td>
<td>11.4</td>
<td>13</td>
<td>9.6</td>
</tr>
<tr>
<td>4</td>
<td>3.7</td>
<td>14</td>
<td>19.5</td>
</tr>
<tr>
<td>5</td>
<td>11.1</td>
<td>6</td>
<td>12.2</td>
</tr>
<tr>
<td>7</td>
<td>8.3</td>
<td>8</td>
<td>11.9</td>
</tr>
<tr>
<td>9</td>
<td>6.6</td>
<td>10</td>
<td>9.5</td>
</tr>
</tbody>
</table>

*95% Prediction Limit = 18.4 ppb*

*1 of next 4 observations exceeds PL*

*Statistically significant exceedance on last event*

**Arsenic Exceedance**

**Practical Notes**

*Formula given for κ applies to single constituent at a single well, without retesting*

*Use κ tables in Unified Guidance to design tests for well networks using a retesting scheme*

*Guidance tables control for SWFPR and desired statistical power*

**NonParametric PLs**

*Upsides*

*Normalized data not required*

*Easy to compute, wide application*

*Fallback option when parametric control chart is not applicable*

*Drawback*

*Much larger BG sample size needed to achieve same accuracy and power (particularly without retesting)*
**Computation**

- Easy set-up: prediction limit = large BG value (e.g., max)
- Next m compliance observations compared to PL
- Confidence level not set by user; depends on BG size (n), m, and which BG value is used as PL
- Computed after the fact'

**PCE Example**

<table>
<thead>
<tr>
<th>Qtr</th>
<th>Well 1</th>
<th>Well 2</th>
<th>Well 3</th>
<th>Well 4</th>
<th>Well 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>&lt;5</td>
<td>9.2</td>
<td>&lt;5</td>
<td>&lt;5</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>&lt;5</td>
<td>&lt;5</td>
<td>5.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>7.5</td>
<td>&lt;5</td>
<td>6.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>&lt;5</td>
<td>6.1</td>
<td>&lt;5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>&lt;5</td>
<td>8</td>
<td>&lt;5</td>
<td>6.2</td>
<td>&lt;5</td>
</tr>
<tr>
<td>6</td>
<td>&lt;5</td>
<td>5.9</td>
<td>&lt;5</td>
<td>&lt;5</td>
<td>&lt;5</td>
</tr>
<tr>
<td>7</td>
<td>6.4</td>
<td>&lt;5</td>
<td>&lt;5</td>
<td>7.8</td>
<td>5.6</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>&lt;5</td>
<td>&lt;5</td>
<td>10.4</td>
<td>&lt;5</td>
</tr>
</tbody>
</table>

**PCE Example (cont.)**

- Interwell prediction limit with n=24 pooled BG values
- PL = BG maximum = 9.2 ppb
- One violation in Well 4
- Calculate achieved confidence level at each well
- \((1 - \alpha) = 86\%\)

**PCE Follow-Up**

- Fairly high chance of false positives
- \(\alpha = 14\%\) per well
- Would need \(n = 76\) to achieve 95% confidence
- Upshot: need to include retesting plan!
- Allows for smaller BG sample size
- But still allows target false positive rate to be achieved
Prediction Limit Summary

- Prediction limits a preferred method of choice in Unified Guidance
- Variety of PLs provided for flexibility
- Choose specific PL according to:
  - Whether assumptions are met
  - Site configuration; need for retesting plan

Using Control Charts

- Important detection monitoring technique
  - Typically used for intrawell testing, though can be structured for interwell
  - Not as flexible as prediction limits
  - No non-parametric version
  - Less published about statistical power, false positive rates

More Just For Fun

- “Statistics play an important role in genetics. For instance, statistics prove that numbers of offspring is an inherited trait. If your parents didn’t have any kids, chances are you won’t either.” — Hugh Graham

Indiana Jones discovers Han Solo

Control Chart Advantages

- Graphical tool to track & identify:
  - Rapid changes in concentration
  - Long-term trends
  - Data plotted cumulatively
  - New data added as it arrives
**Use & Interpretation**

- Control chart limit similar to prediction limit
  \[ CCL = \bar{x}_{BG} + h_s_{BG} \]
- Process "out-of-control" when either \( x_i \) OR CUSUM (\( S_i \)) exceeds threshold (i.e., control limit)
- Signals that compliance data have changed compared to baseline conditions

**Barium Prediction Limit**

**Barium Control Chart**

**Nickel Example**
Nickel Control Chart

New Guidance

- Common thresholds for control charts ($h=4-5$) don’t account for retesting or multiple comparisons
- No ‘easy’ way to adjust $h$ to get pre-specified power and false positive rate
- Unified Guidance recommends site-specific simulation before implementing control charts

Control Chart Prospects Improving

- Control chart simulations scare many away
- Many use older limits or don’t use at all
- New study shows benefit of control charts in side-by-side comparison to prediction limits (Fernald)

Fernald Comparison

- Control charts tested vs. prediction limits both for:
  - Theoretical performance (false positive rate, statistical power)
  - Real-life performance (400+ sets of monitoring data)
**Theoretical Findings**

- After fixing same site-wide false positive rate (SWFPR) for both methods:
  - Control limits (h) were always higher than prediction limits (κ), BUT
  - Control charts were almost always more powerful
- Why? CUSUM!

**Real-Life Performance**

- Using historical, quarterized data, control charts captured more SSIs than comparable prediction limits
- Control charts also had shorter ARLs to the first SSI than prediction limits
- Retesting was critical for screening out false positives

**Example: Boron (1-of-2)**

- Using historical, quarterized data, control charts captured more SSIs than comparable prediction limits
- Control charts also had shorter ARLs to the first SSI than prediction limits
- Retesting was critical for screening out false positives
Example: Magnesium (1-of-3)

Control Chart Summary

- Control charts a viable, powerful intrawell tool
  - Less well-studied than prediction limits, BUT
  - New study shows control charts can perform better than prediction limits

- Control limits must account for larger sites, multiple comparisons
  - Simulations needed to adjust control h-limit to account for size of network, adequate power, target site-wide false positive rate, retesting strategy, BG updating, etc.