Evaluation of Standards data collected from probabilistic sampling programs

Eric P. Smith
Y. Duan, Z. Li, K. Ye
Statistics Dept., Virginia Tech
This talk was not subjected to USEPA review. The conclusion and opinions are solely those of the authors and not the views of the Agency.
Outline

- Background
  - Standards assessments
- Single site analysis
- Regional analysis
  - Mixed model approach
  - Bayesian approach
- Upshot: need models that allow for additional information to be used in assessments
Standards assessment – 303d

- Clean Water Act section 303d mandates states in US to monitor and assess condition of streams
- Site impaired – list site, start TMDL process (Total Max Daily Loading)
- Impaired means site does not meet usability criteria
Linkages in 303(d)

- Set goals and WQS standards
- Conduct monitoring
- Sampling plan
- Meeting WQS?
  - Yes
  - No
    - Tests
    - Apply Antidegradation
      - Local to regional
      - List
      - TMDLs
      - NPDES, 319, SRF, etc.
Impaired sites

- Site impaired if standards not met
- Standards – defined through numerical criteria
  - Involve frequency, duration, magnitude
- Old method
  - Site impaired if >10% of samples exceed criteria
  - Implicit statistical decision process- error rates
Test of impairment

![Graph showing density against concentration with labeled categories: don't list and list, with concentration levels 0.05, 0.1, 0.15, 0.2, 0.25, 5, 10, and 15, and density levels 0, 0.05, 0.1, 0.15, 0.2, 0.25, indicating standard and cutoff levels of >=0.1 and <0.1.]
Newer approach to evaluation

- **Frequency:**
  - Binomial method
  - Test p<0.1

- **Magnitude**
  - Acceptance sampling by variables
  - Tolerance interval on percentile
  - Test criteria by computing mean for the distribution of measurements and comparing with what is expected given the percentile criteria
Problems

- Approach is local
  - Limited sampling budget; many stations means small sample sizes per station
  - Impairment may occur over a region
  - Modeling must be relatively simple (hard to account for seasonality, temporal effects)
  - Does not complement current approaches to sampling
  - Site history is ignored
  - Not linked to TMDL analysis (regional) and 305 reporting
Probabilistic sampling schemes

Rotating panel surveys

- Some sites sampled at all possible times
- Other sites sampled on rotational basis
- Sites in second group may be randomly selected
Making the assessment regional

\[ Y = \text{mean} + \text{site} \]
\[ Y = \text{mean} + \text{time} + \text{site} \]

\[ y = X\beta + Zu + e \]

- X defines fixed effects (time), Z defines random ones (site, location), \( \beta, u \) are parameters

- Covariances

\[ e \sim \text{MVN}(0, \Gamma) \]
\[ u \sim \text{MVN}(0, G) \]

\[ V(y) = ZGZ' + \Gamma \]
Regional Mixed Model

- Allows for covariates
- Allows for a variety of error structures
  - Temporal, spatial, both
- Does not require equal sample sizes etc
- Allows estimation of means for sites with small sample sizes
  - Improves estimation by borrowing information from other sites
Simple model

\[ y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \]

- Random site effect
- Error term allows for modeling of temporal or spatial correlation

- Testing is based on estimate and variance of mean for site i \((\mu_i)\)
- Can also test for regional impairment using distribution of grand mean
Error and stochastic components

\[ y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \]

Random site effect

\[ \varepsilon_{ij} \sim N(0, \sigma^2) \]

Covariance Structure without correlation (one random effect model)

\[ \text{Var}(\varepsilon) = \sigma^2 \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{21} & 1 & \rho_{23} & \rho_{24} \\ \rho_{31} & \rho_{32} & 1 & \rho_{34} \\ \rho_{41} & \rho_{42} & \rho_{43} & 1 \end{bmatrix} \]

Error term allows for modeling of temporal or spatial correlation
Test based on OLS estimations for each site $i$

\[ \frac{\bar{y}_i - \text{baseline}}{\hat{\sigma} / \sqrt{n_i}} \sim t_{df, \delta} \]

where $\bar{y}_i$ and $\hat{\sigma}$ are OLS estimates of $\mu$ and $\sigma$;

$df = n_i - 1$, $\delta = \text{noncentrality}$

Baseline is the standard. For DO, we use 5, and for PH 6.

Model based: same idea but mean and variance are estimated from model
Located in SW Virginia
Good bass fishing
DO data collected at four stations of PHILPOTT RESERVOIR (years 2000, 2001 & 2002)
## Evaluation based on Do data of PHILPOTT RESERVIOR (2000-2002)

<table>
<thead>
<tr>
<th></th>
<th>4ASRE046.90</th>
<th>Model based</th>
<th>4ASRE052.31</th>
<th>4ASRE056.06</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>n</strong></td>
<td>28</td>
<td>31</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td><strong>Sample mean</strong></td>
<td>7.55</td>
<td>6.66</td>
<td>6.67</td>
<td></td>
</tr>
<tr>
<td><strong>Sample variance</strong></td>
<td>5.81</td>
<td>9.56</td>
<td>16.15</td>
<td></td>
</tr>
<tr>
<td><strong>% exceeding</strong></td>
<td>11</td>
<td>26</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td><strong>Binomial p-value</strong></td>
<td>.5406</td>
<td>.0096</td>
<td>.0033</td>
<td></td>
</tr>
<tr>
<td><strong>Test statistic</strong></td>
<td>5.6</td>
<td>4.27</td>
<td>2.99</td>
<td>2.35</td>
</tr>
<tr>
<td><strong>critical value</strong></td>
<td>4.75</td>
<td>5.05</td>
<td>5.19</td>
<td>5.2</td>
</tr>
<tr>
<td><strong>conclusion</strong></td>
<td>Fail to reject</td>
<td>reject</td>
<td>reject</td>
<td>reject</td>
</tr>
</tbody>
</table>

Single site analysis
Bayesian approach

\[ y_{ij} = \mu + a_i + \varepsilon_{ij} \]

- \( a \) is a random site effect
- Error term may include temporal correlation or spatial
- Priors on parameters
  - Mean - uniform
  - \( a \) is normal (random effect) variance has prior

\[
\pi(\sigma^2, \sigma_a^2) \propto \frac{1}{\sigma^2} \frac{1}{\sigma^2 + \sigma_a^2}
\]
Alternative: Using historical data

- Power prior – Chen, Ibrahim, Shao 2000
- Use likelihood from the previous assessment ($D_0$). Basic idea: weight new data by prior data
- Power term, $\delta$, determines influence of historical data.
- Modification to work with Winbugs
Incorporate Historical Data using Power Priors

Make $\delta$ random, and assign a prior $\pi(\delta) = \text{Beta}(\alpha, \beta)$ on it. The joint posterior of $(\theta, \delta)$ becomes

$$
\pi(\theta, \delta | D_0, D) \propto \frac{L(\theta | D)(L(\theta | D_0))^\delta \pi(\theta)\pi(\delta)}{\int (L(\theta | D_0))^\delta \pi(\theta)d\theta} I_A(\delta)
$$

where $D$ is current data and $D_0$ is past data

$$
A = \{\delta : 0 < \int \pi(\theta)(L(\theta | D_0))^\delta d\theta < \infty\}
$$

Advantage: Improve the precision of estimates.
PH data collected at four stations: use past information to build prior

![Graph showing PH data collected at four stations.](image)

- 1AUMC004.43: n = 16, n₀ = 62
- 1AUMC009.61: n = 12, n₀ = 31
- 1ACHO003.65: n = 24, n₀ = 84
- 1APOM002.41: n = 21, n₀ = 75
Evaluate site impairment based on PH data with power priors

<table>
<thead>
<tr>
<th>Station of interest</th>
<th>1AUMC004.43</th>
<th>1AUMC009.61</th>
<th>1Acho003.65</th>
<th>1APOM002.41</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>16 (yr.99-02)</td>
<td>12 (yr.99-01)</td>
<td>24 (yr.99-01)</td>
<td>21 (yr.99-00)</td>
</tr>
<tr>
<td>No. obs &lt;6</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>sample mean</td>
<td>6.91</td>
<td>6.78</td>
<td>6.43</td>
<td>7.87</td>
</tr>
<tr>
<td>sample variance</td>
<td>0.82</td>
<td>1.06</td>
<td>0.78</td>
<td>1.23</td>
</tr>
<tr>
<td>n₀</td>
<td>62 (yr.90-98)</td>
<td>31 (yr.90-98)</td>
<td>84 (yr.90-98)</td>
<td>75 (yr.90-98)</td>
</tr>
<tr>
<td>sample mean of D₀</td>
<td>7.05</td>
<td>6.73</td>
<td>6.95</td>
<td>7.88</td>
</tr>
<tr>
<td>Percent exceed the EPA standard</td>
<td>0.13</td>
<td>0.17</td>
<td>0.25</td>
<td>0.10</td>
</tr>
<tr>
<td>P-value of Binomial test (H₀: p=0.1, Hₐ: p&gt;0.1)</td>
<td>0.4853</td>
<td>0.3410</td>
<td>0.0277</td>
<td>0.6353</td>
</tr>
</tbody>
</table>

Bayesian test. (H₀: L=6, Hₐ: L<6), L is the lower 10th percentile of PH

With Reference Prior:

<table>
<thead>
<tr>
<th></th>
<th>P(H₀)</th>
<th>posterior s.d. of ?</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(H₀)</td>
<td>0.1663</td>
<td>0.0502</td>
</tr>
<tr>
<td>posterior s.d. of ?</td>
<td>0.3399</td>
<td>0.4708</td>
</tr>
</tbody>
</table>

With Power Prior:

<table>
<thead>
<tr>
<th></th>
<th>P(H₀)</th>
<th>posterior s.d. of ?</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(H₀)</td>
<td>0.4868</td>
<td>0.03525</td>
</tr>
<tr>
<td>posterior s.d. of ?</td>
<td>0.2566</td>
<td>0.2562</td>
</tr>
<tr>
<td></td>
<td>0.2381</td>
<td>0.2477</td>
</tr>
</tbody>
</table>
If multiple historical data sets are available, assign a different $\delta_j$ for each historical data set.

$$L(\theta | D) \left( \prod_{j=1}^{m} (L(\theta | D_{0j}))^{\delta_j} \pi(\delta_j) \right) \pi(\theta)$$

$$\pi(\theta, \delta | D_0, D) \propto \frac{\int \left( \prod_{j=1}^{m} (L(\theta | D_{0j}))^{\delta_j} \pi(\theta) \right) d\theta}{I_B(\delta)}$$

where

$$B = \left\{ (\delta_1, \ldots, \delta_m) : 0 < \int \left( \prod_{j=1}^{m} (L(\theta | D_{0j}))^{\delta_j} \pi(\theta) \right) d\theta < \infty \right\}$$

Data collected at adjacent stations could be used as “historical” data.
DO data collected at four stations of PHILPOTT RESERVOIR (years 2000, 2001 & 2002)
Evaluate site impairment based on DO data collected at four stations of PHILPOTT RESERVOIR (years 2000, 2001 & 2002)

<table>
<thead>
<tr>
<th>Station of interest</th>
<th>4ASRE046.90</th>
<th>4ASRE048.98</th>
<th>4ASRE052.31</th>
<th>4ASRE056.06</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>28</td>
<td>31</td>
<td>31</td>
<td>32</td>
</tr>
<tr>
<td>No. obs &lt;5</td>
<td>3</td>
<td>5</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>sample mean</td>
<td>7.55</td>
<td>7.10</td>
<td>6.66</td>
<td>6.67</td>
</tr>
<tr>
<td>sample variance</td>
<td>5.81</td>
<td>8.28</td>
<td>9.56</td>
<td>16.15</td>
</tr>
<tr>
<td>Percent exceed the EPA standard</td>
<td>0.11</td>
<td>0.16</td>
<td>0.26</td>
<td>0.28</td>
</tr>
<tr>
<td>P-value of Binomial test (H₀: p=0.1  Hₐ: p&gt;0.1)</td>
<td>0.5406</td>
<td>0.1932</td>
<td>0.0096</td>
<td>0.0033</td>
</tr>
<tr>
<td>Bayesian test. (H₀: L=5  Hₐ: L&lt;5), L is the lower 10th percentile of DO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With Reference Prior:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P(H₀)</td>
<td>0.1640</td>
<td>0.0038</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>posterior s.d. of ?</td>
<td>0.6514</td>
<td>0.7325</td>
<td>0.7875</td>
<td>1.008</td>
</tr>
<tr>
<td>With Power Prior:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P(H₀)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>posterior s.d. of ?</td>
<td>0.5485</td>
<td>0.5371</td>
<td>0.5439</td>
<td>0.6162</td>
</tr>
</tbody>
</table>
DO data collected at four stations of MOOMAW RESERVOIR (years 2000 & 2001)

n=21, n=20, n=16, n=18

DO PROBE

STATIONS

2-JKS044.60  2-JKS046.40  2-JKS048.90  2-JKS053.48
Evaluate site impairment based on DO data collected at four stations of MOOMAW RESERVOIR (years 2000 & 2001)

<table>
<thead>
<tr>
<th>Station of interest</th>
<th>2-JKS044.60</th>
<th>2-JKS046.40</th>
<th>2-JKS048.90</th>
<th>2-JKS053.48</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>21</td>
<td>20</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>No. obs &lt;5</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>sample mean</td>
<td>8.16</td>
<td>8.06</td>
<td>8.19</td>
<td>9.81</td>
</tr>
<tr>
<td>sample variance</td>
<td>9.73</td>
<td>10.07</td>
<td>12.14</td>
<td>1.32</td>
</tr>
<tr>
<td>Percent exceed the EPA standard</td>
<td>0.14</td>
<td>0.15</td>
<td>0.19</td>
<td>0.00</td>
</tr>
<tr>
<td>P-value of Binomial test (H₀: p=0.1 Hₐ: p&gt;0.1)</td>
<td>0.3516</td>
<td>0.3231</td>
<td>0.2108</td>
<td>1.0000</td>
</tr>
<tr>
<td>Bayesian test. (H₀: L=5 Hₐ: L&lt;5), L is the lower 10th percentile of DO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With Reference Prior:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P(H₀)</td>
<td>0.1497</td>
<td>0.1149</td>
<td>0.1022</td>
<td>0.9968</td>
</tr>
<tr>
<td>posterior s.d. of ?</td>
<td>1.0030</td>
<td>1.0500</td>
<td>1.3110</td>
<td>0.7219</td>
</tr>
<tr>
<td>With Power Prior:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P(H₀)</td>
<td>0.1338</td>
<td>0.1206</td>
<td>0.1163</td>
<td>0.3301</td>
</tr>
<tr>
<td>posterior s.d. of ?</td>
<td>0.6698</td>
<td>0.6832</td>
<td>0.7132</td>
<td>0.7469</td>
</tr>
</tbody>
</table>
Comments

⚠️ Advantages
- Greater flexibility in modeling
- Allows for site history to be included
- Can include spatial and temporal components
- Can better connect to TMDL analysis and probabilistic sampling

⚠️ Disadvantage
- Requires more commitment to the modeling process
- Greater emphasis on the distributional assumptions