Building of the Bayesian Network Relative Risk Model for the Upper San Francisco Estuary and the analysis of Risk

Wayne G. Landis, Mikayla Bowers, Ethan Brown, Steven Eikenbary, Skyler Elmstron, Eric Lawrence, April J. Markiewicz, Emma E. Sharpe, Erika Whitney.

Institute of Environmental Toxicology and Chemistry, Huxley College of the Environment, Western Washington University, Bellingham WA

Institute of Environmental Toxicology and Chemistry
Huxley College of the Environment
Western Washington University

Funding
CA Metropolitan Water District
State Water Contractors
CA Department of Pesticide Regulation
CA Delta Program

EcoRisk Projects
https://www.youtube.com/channel/UCBb99MTRwREF0y5Sp0NYiA
Outline for this morning....

A rapid review of Bayesian networks

BN-Relative Risk model for USFE

BN-RRM for pesticides and fish toxicity

Derivation of models for water quality and macroinvertebrate community structure

Ongoing work and next steps
Why Bayesian networks?....

- Adaptable—Chemicals, water quality, microplastics, rainfall, and restoration options.
- Multiple stressors are normal and can be calculated.
- Interactions of management methods can be evaluated.
- Pictures and number can aid in communication.
Bayesian Network Relative Risk Model

The methods have been published for other sites and a variety of stressors.
Bayesian Network Relative Risk Model

The methods have been published for other sites and a variety of stressors:

![Diagram of Bayesian Network]

- **Parent Nodes**
  - Stressor A
    - States: zero, low, med, high
    - Conditional Probability Table
- **Child Node**
  - Condition 1
    - States: zero, low, med, high
    - Conditional Probability Table

![Conditional Probability Table]

<table>
<thead>
<tr>
<th>Stressor A</th>
<th>Stressor B</th>
<th>zero</th>
<th>low</th>
<th>med</th>
<th>high</th>
</tr>
</thead>
<tbody>
<tr>
<td>zero</td>
<td>zero</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>low</td>
<td>zero</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>med</td>
<td>zero</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>high</td>
<td>zero</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

![Probability Table]
The process is based on causality as defined by specific pathways and incorporates probability.

The linkages between nodes are justified by information related to cause and effect, not mere association.
The study area—the Upper San Francisco Bay Estuary

Diverse system with multiple stressors, urban to agriculture, to parklands.

We divide the study area into regions-risk regions.
Bayesian Network Relative Risk Model for the USFE

Here is the overall Bayesian network for a variety of endpoints.

The model structure is the same but uses data specific to each of the regions.
Bayesian Network
Relative Risk Model for the USFE

Pesticides and fish toxicity related endpoints
Bayesian Network
Relative Risk Model for the USFE

Water quality and specifically Macroinvertebrate community structure

We are focusing on these for this presentation.
Bayesian Network Relative Risk Model Fish Toxicity-

K. Laetz et al 2009 data
AChE inhibition

Mixture BN model for Fish Mortality. Five different pesticides are incorporated. The mixture mortality nodes incorporate the mixture additive equations to estimate the toxicity. The concentration distributions are taken from measured values for each of the risk regions.

Hutton et al 2021 data
silversides
The dose-response curves were used to generate the categorizations for each pesticide node.

<table>
<thead>
<tr>
<th>Values (µg/L)</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0165</td>
<td>EC5</td>
</tr>
<tr>
<td>0.0682</td>
<td>EC10</td>
</tr>
<tr>
<td>0.317</td>
<td>EC20</td>
</tr>
<tr>
<td>4.397</td>
<td>EC50</td>
</tr>
<tr>
<td>5.06</td>
<td>Highest record concentration from field data</td>
</tr>
</tbody>
</table>
Estimating mixture toxicity steps

• For each mixture component, fit a log logistic 3 parameter model to the available toxicity data.

• For each mixture component, calculate the ECx.

• For each mixture component, normalize the concentrations of the toxicity data by the ECx.

• For each mixture component, fit a log logistic 3 parameter model to the ECx normalized data.

• Take the geometric mean of the three-log logistic 3 parameter model parameters for the ECx normalized models.

• Use the geometric means in the log logistic 3 parameter model to create the mixture equation.
Building the conditional probability tables for mixture interactions....Mixture Mortality Node for Chlorpyrifos and Bifenthrin. E. Lawrence


Mixture Models - Dose Response Model

Averaging. The concentration addition (CA) model normalizes concentrations within a mixture by an ECx value, or a concentration that corresponds to a level of toxic effect. These normalized concentrations, also called toxic units, represent the relative potencies of the mixture that can then be added together.

\[ \sum_{i=1}^{n} \left( \frac{c_i}{ECx_i} \right) = 1 \]

\[ c_i = \text{Concentration of chemical } i \text{ in a mixture.} \]

\[ ECx_i = \text{Effective concentration for } x \text{ level of effect for chemical } i. \]
Comparison of the exposure-response model curves for several chemicals using EC10 (A), EC20 (B) and EC50 (C) normalized concentrations.
Example of additive exposure-response model with EC20 Normalization
Risk Calculation for the Confluence from measured concentrations

There is an 83 percent probability of the Fish Mortality being equal to or greater than an EC10.
There is an 80.9 percent probability of the Fish Mortality being equal to or greater than an EC10.
Comparative risk for the USFE

<table>
<thead>
<tr>
<th>Risk Region</th>
<th>Probability of Greater than an EC10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confluence</td>
<td>83</td>
</tr>
<tr>
<td>Suisun Bay</td>
<td>79.6</td>
</tr>
<tr>
<td>Central Delta</td>
<td>78.9</td>
</tr>
<tr>
<td>Sacramento River</td>
<td>78.7</td>
</tr>
<tr>
<td>North Delta</td>
<td>80.4</td>
</tr>
<tr>
<td>South Delta</td>
<td>80.9</td>
</tr>
</tbody>
</table>

*Confluence* is the greatest with the *Sacramento* the lowest. All are greater than a 30 percent probability. Note how close they are for the 5-chemical model.
Bifenthrin was the single most important chemical concentration for each of the risk regions.
Fish Risk summary

The interactions of chemical mixtures can be estimated and built into a Bayesian network.

The fish toxicity node can be used to estimate survivorship for input into a population model, either Leslie matrix or IBM.

The building blocks and tools need to build the rest of the risk assessment BN have been constructed.
Macroinvertebrates

- Commonly used as indicators for aquatic ecosystems.
- Play important role in food webs.

- There are many metrics used to measure macroinvertebrate community structure
  - Taxa abundance
  - Taxa richness
  - Order specific metrics such as EPT taxa richness (Ephemeroptera, Plecoptera, and Trichoptera).

- California Stream Condition Index (CSCI) – Index that uses multiple macroinvertebrate indices along with other environmental factors to estimate stream impairment (Mazor et al. 2014).

Data Sources

- California Environmental Data Exchange Network (CEDEN)
  - Benthic Macroinvertebrate Samples
  - Water Quality Samples
  - Contaminants

- Zooplankton data synthesizer (ZoopSynth)
  - Water Column Macroinvertebrate Samples


Combining Data

Joining benthic macroinvertebrate samples with water quality samples from CEDEN database

- Benthic samples do not include simultaneous water quality samples.

- To combine the datasets I conducted a spatial analysis using the “sf” package in R Statistical Software (Pebesma et al. 2021).
Combining Data

CEDEN Benthic Samples

CEDEN Water Quality Samples

Combining Data

CEDEN Benthic and Water Quality Samples

Spatial join of Benthic and Water Quality sample locations.
- Occur on the same day
- Are within 500 meters
Combining Data

Spatial join of Benthic and Water Quality sample locations.
- Occur on the same day
- Are within 500 meters

All Benthic samples: n = 159
Benthic samples after join: n = 64
Metric: relative richness by taxa

\[
Relative\ \text{Richness} = \frac{\#\ \text{taxa}}{\text{Total taxa in sample}}
\]

Multivariate analysis using “vegan” package in R Statistical Software (Oksanen et al. 2020)

Taxa were defined by Phylum or by Order or Class within the Arthropoda Phylum

<table>
<thead>
<tr>
<th>Phylum</th>
<th>Order (within Arthropoda)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mollusca</td>
<td>Mysida</td>
</tr>
<tr>
<td>Annelida</td>
<td>Diptera</td>
</tr>
<tr>
<td>Chordata</td>
<td>Hymenoptera</td>
</tr>
<tr>
<td>Cnidaria</td>
<td>Amphipoda</td>
</tr>
<tr>
<td>Nematoda</td>
<td>Thysanoptera</td>
</tr>
<tr>
<td>Platyhelminthes</td>
<td>Odonata</td>
</tr>
<tr>
<td>Bryozoa</td>
<td>Coleoptera</td>
</tr>
<tr>
<td>Nemertea</td>
<td>Trichoptera</td>
</tr>
<tr>
<td>Nemertea</td>
<td>Ephemeroptera</td>
</tr>
<tr>
<td>Bryozoa</td>
<td>Hydrachnidia</td>
</tr>
<tr>
<td>Collembola (Class)</td>
<td>Collembola (Class)</td>
</tr>
</tbody>
</table>

Multivariate Analysis

Ordination Methods (Graham et al. 2018):

Allow for graphing data with multiple variables where the distance between points represents how similar they are, taking account the information of multiple variables.

**Principal Component Analysis**
- Simplifies multiple variables into a new set of variables (principal components) that attempt to explain most of the variability into just a few axes. Uses linear relationships that maximize variance.

**Non-Metric Multidimensional Scaling (NMDS)**
- Allows you to specify the number of axes to simplify the data into using an iterative process.
Non-metric Multidimensional Scaling

- At least two distinct groupings

- Water quality parameters included as fitted vectors

- Temperature, pH, and Conductivity were significant predictors ($p < 0.05$).

- Water quality parameters limited by what parameters were measured in each sample.

CEDEN Benthic MI clusters with water quality vectors
Principal Components

Similar grouping pattern as NMDS

Main Phylums influencing grouping (potential endpoints):
- Diptera
- Amphipoda
- Annelida
- Mollusca
Community structure and water quality-Next Steps

- Refine methods for joining datasets.
- Include contaminant concentrations as potential influencing factors.
- Include environmental parameters used by CSCI index.
- Develop Macroinvertebrate index or other tools to use as endpoint.
- Build the water quality pathways.
Why are we doing this...to make decisions.

Observe, Orient, Decide and Act—Loop

Unfolding Events → Observe → Orient → Decide → Act → Observe

"OODA.Boyd" by Patrick Edwin Moran - Own work. Licensed under CC BY 3.0 via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:OODA.Boyd.svg#/media/File:OODA.Boyd.svg
Why are we doing this?...to make decisions.

Social goals (economic, cultural, well-being) that correspond to the multiple resources at a site.

Ecological risk assessment: Designed to accommodate the multiple management goals in a site specific manner.

Management and remediation options: Methods are evaluated using risk assessment and evaluated by monitoring that incorporates multiple stressors and multiple endpoints.

Estimates of risk to multiple endpoints across the management region.

Inputs describing the potential outcomes from the remediation options and the data from ongoing monitoring.

Change in Externalities: Alterations in environmental conditions outside the management loop such as climate change, population growth, economics, technology.

Derivation of the endpoints considered in the risk assessment and the criteria to be met in a spatially explicit context.

Constraints due to economic resources, benefits, social concerns, and legislation and the importance of the remediation goals.

Public Engagement and Governance

Research, Engineering, Risk Assessment and Management

Decision making: Can be at different levels including local stakeholders, responsible parties, local, regional, national and international agencies.

Just another OODA loop with the externalities included.
Why are we doing this?...to make decisions.

Calculation to estimated conditions that result in an EC10 in the Confluence.
Final thoughts....

• Work on building the dataset that combined CEDEN and SURF into a single database that can be reliable and repeatable.

• Toxicity not reported to optimize a risk assessment-the Delta Smelt data do not document exposure-response and have few endpoints that directly affect survival and reproduction.

• Chemistry, water quality and invertebrate sampling sites do not correspond in location.

• Next step is to finish building the pathways and running the analyses.
Thanks for your time.....