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/UCBb99MTRwREfP0y5Sp0NYiA

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.

Integrated Environmental Assessment and Management — Volume 13, Number 1—pp. 115–126
Received: 22 February 2016 | Revised: 5 April 2016 | Accepted: 27 May 2016 115

Health & Ecological Risk Assessment

A General Risk-Based Adaptive Management Scheme Incorporating the Bayesian Network Relative Risk Model with the South River, Virginia, as Case Study

Wayne G Landis,† April J Markiewicz,† Kim K Ayre,† Annie F Johns,† Meagan J Harris,† Jonah M Stinson,† and Heather M Summers†*
†Institute of Environmental Toxicology, Huxley College of the Environment, Western Washington University, Bellingham, Washington, USA

Integrated Environmental Assessment and Management — Volume 17, Number 1—pp. 79–94
Received: 12 June 2019 | Returned for Revision: 17 July 2020 | Accepted: 23 September 2020 79

Special Series

The Origin, Development, Application, Lessons Learned, and Future Regarding the Bayesian Network Relative Risk Model for Ecological Risk Assessment

Wayne G Landis†*
†Institute of Environmental Toxicology and Chemistry, Huxley College of the Environment, Western Washington University, Bellingham, Washington, USA

A Bayesian Approach to Integrated Ecological and Human Health Risk Assessment for the South River, Virginia Mercury-Contaminated Site

Meagan J. Harris, Jonah Stinson, and Wayne G. Landis*

Risk Analysis, Vol. 37, No. 7, 2017 DOI: 10.1111/risa.12691

Integrated Environmental Assessment and Management — Volume 17, Number 1—pp. 95–109
Received: 14 May 2020 | Returned for Revision: 15 July 2020 | Accepted: 6 October 2020 95

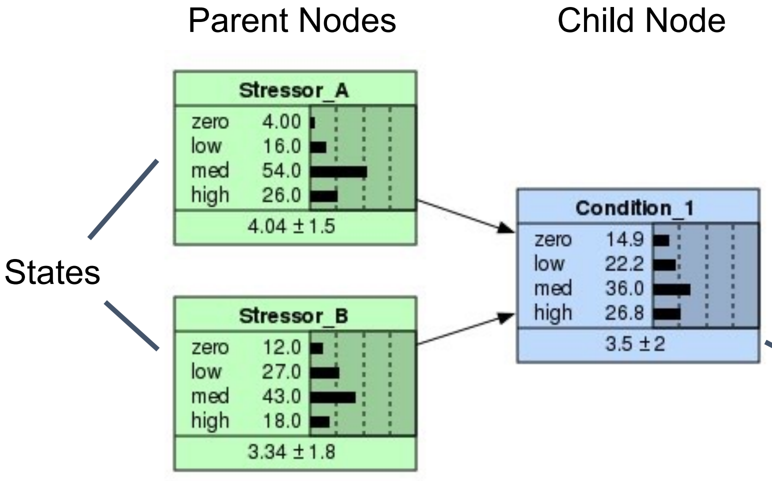
Special Series

Integrating Metapopulation Dynamics into a Bayesian Network Relative Risk Model: Assessing Risk of Pesticides to Chinook Salmon (*Oncorhynchus tshawytscha*) in an Ecological Context

Chelsea J Mitchell,† Eric Lawrence,‡ Valerie R Chu,‡ Meagan J Harris,§ Wayne G Landis,‡ Katherine E von Stackelberg,|| and John D Stark†*
†Washington State University, Puyallup, Washington, USA
‡Institute of Environmental Toxicology, Huxley College of the Environment, Western Washington University, Bellingham, Washington, USA
§Whatcom Conservation District, Lynden, Washington, USA
||Harvard Center for Risk Analysis, Boston, Massachusetts, USA

Bayesian Network Relative Risk Model

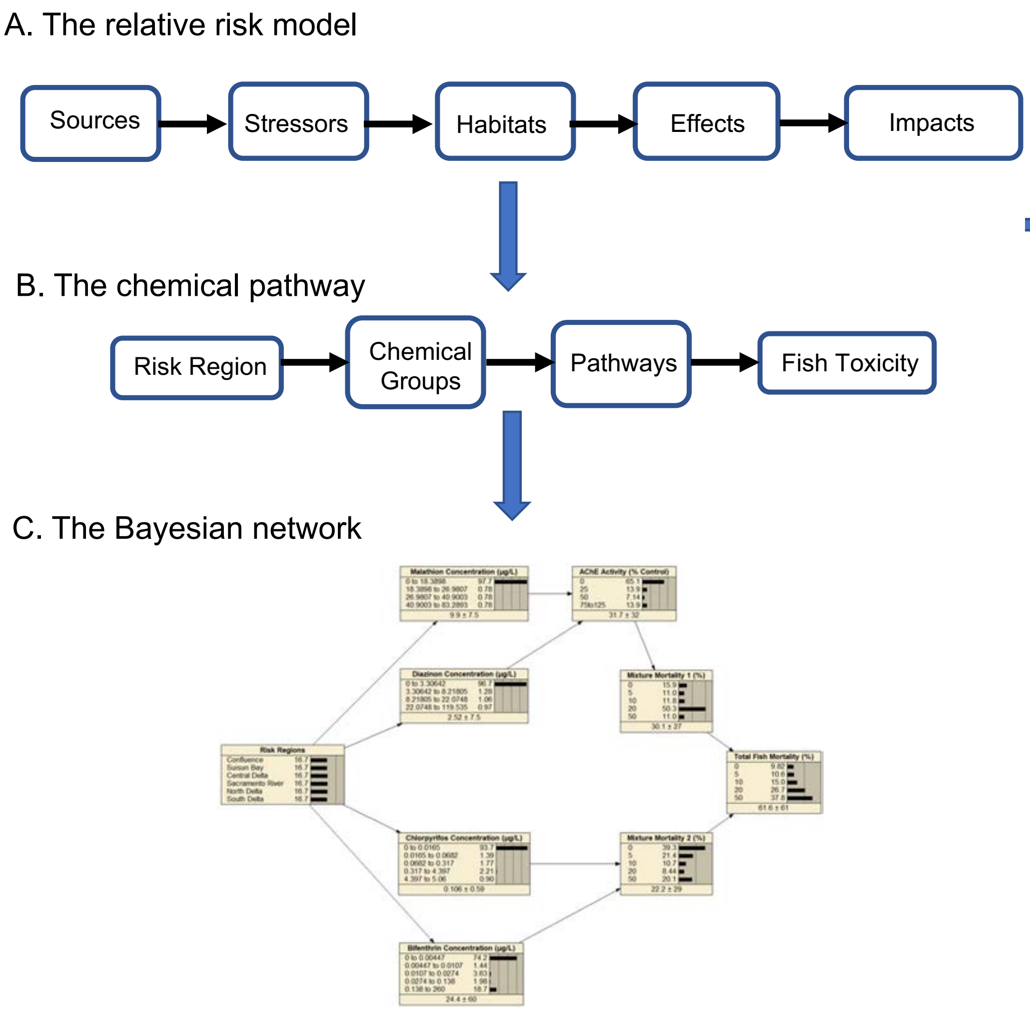
The methods have been published for other sites and a variety of stressors:



Conditional Probability Table

Stressor_A	Stressor_B	zero	low	med	high
zero	zero	100.00	0.000	0.000	0.000
zero	low	90.000	8.000	1.500	0.500
zero	med	75.000	20.000	4.000	1.000
zero	high	60.000	25.000	10.000	5.000
low	zero	75.000	20.000	4.000	1.000
low	low	50.000	35.000	10.000	5.000
low	med	25.000	35.000	30.000	10.000
low	high	10.000	30.000	45.000	15.000
med	zero	25.000	35.000	30.000	10.000
med	low	10.000	30.000	45.000	15.000
med	med	5.000	25.000	50.000	20.000
med	high	1.000	9.000	40.000	50.000
high	zero	15.000	25.000	40.000	20.000
high	low	10.000	15.000	35.000	40.000
high	med	5.000	10.000	30.000	55.000
high	high	1.000	4.000	20.000	75.000

Bayesian Network Relative Risk Model



The process is based on causality as defined by specific pathways and incorporates probability.

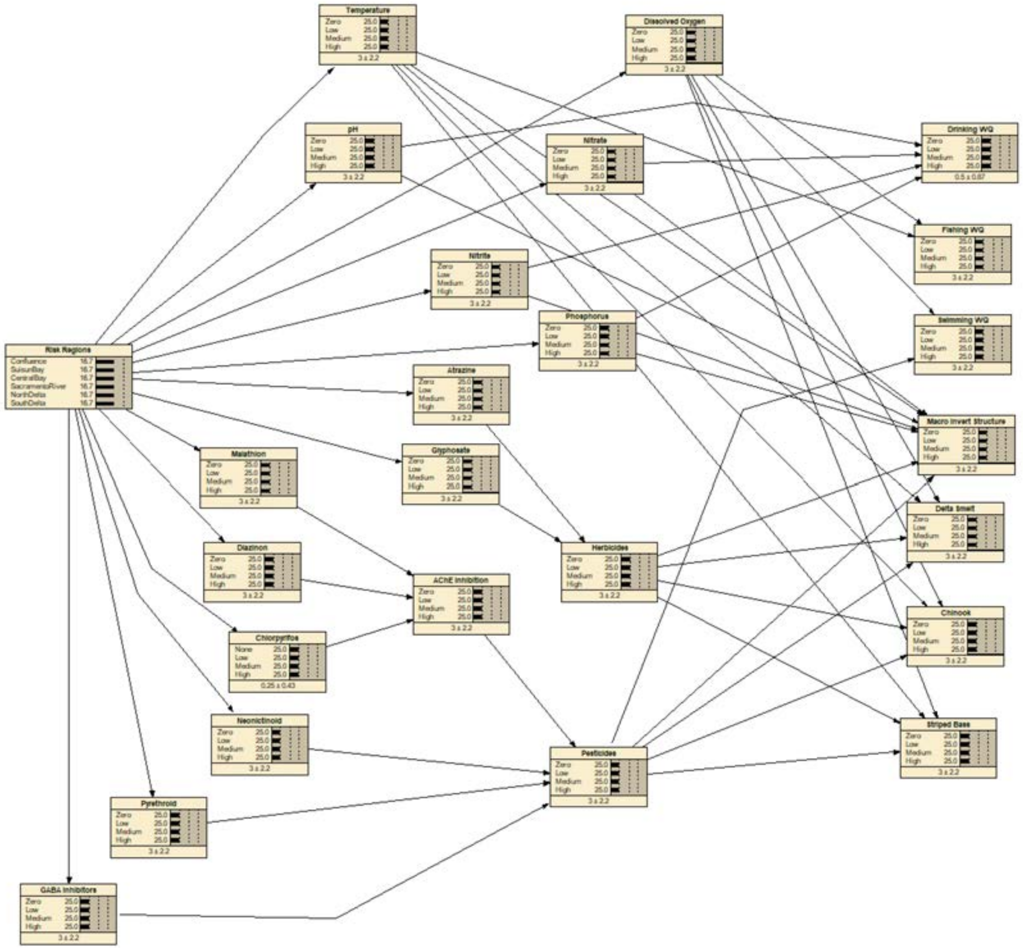
The linkages between nodes are justified by information related to cause and effect, not mere association.

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.

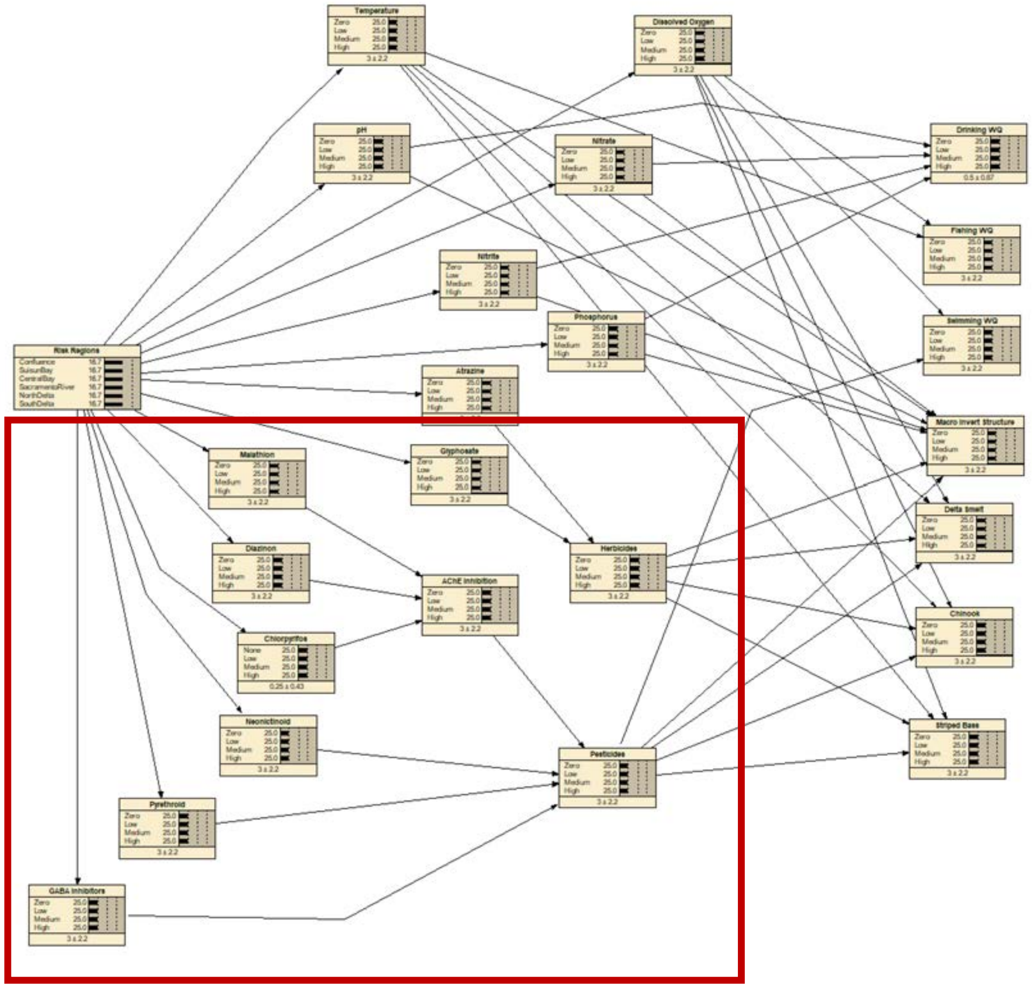
Bayesian Network Relative Risk Model for the USFE



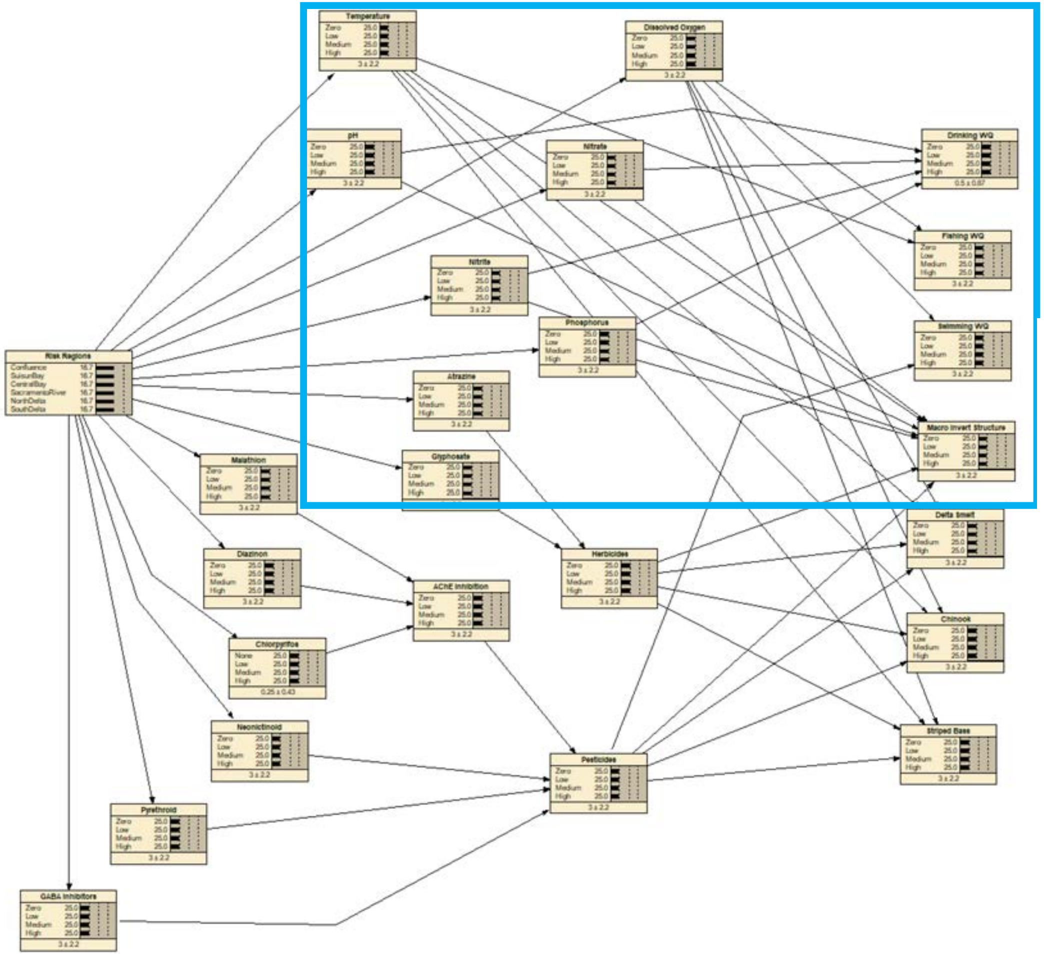
Here is the overall Bayesian network for a variety of endpoints.

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

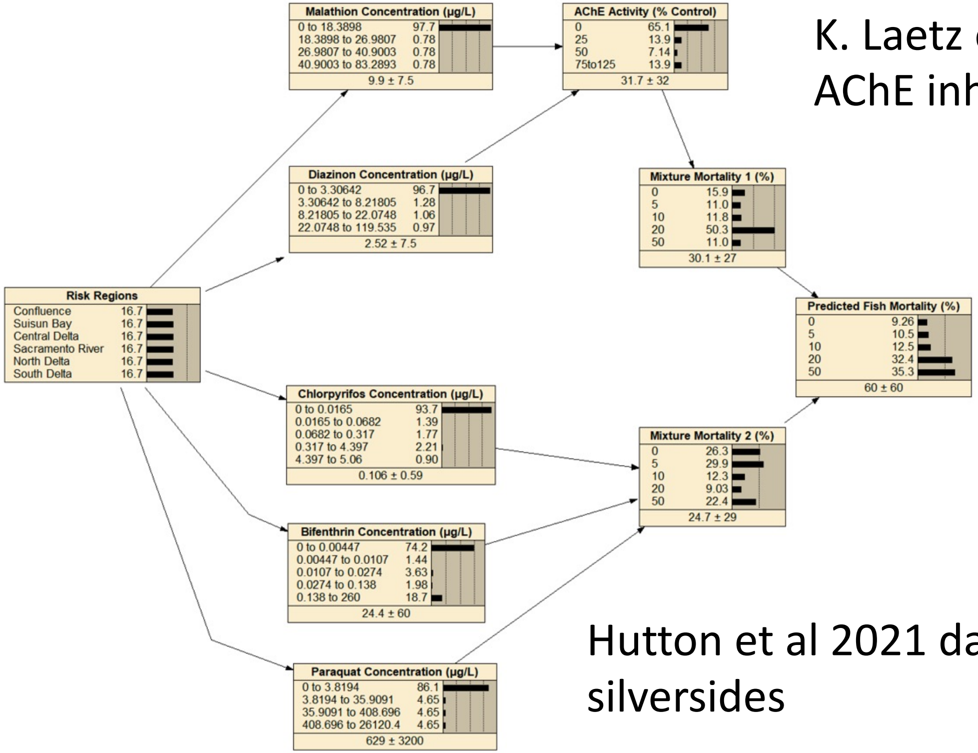


Bayesian Network Relative Risk Model for the USFE

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

Hutton et al 2021 data
silversides

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.

The dose-response curves were used to generate the categorizations for each pesticide node.

Values (µg/L)	Type
0.0165	EC5
0.0682	EC10
0.317	EC20
4.397	EC50
5.06	Highest record concentration from field data

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 https://cpb-us-e1.wpmucdn.com/wp.wwu.edu/dist/1/2430/files/2021/10/WWU-DPR-technical-memo-10-2021.pdf

<u>Mixture Models-Dose Response Model Averaging</u>. 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 = Concentration of chemical i in a mixture.

ECx_i = Effective concentration for x level of effect for chemical i.

Mortality2 Table (in Bayes net EES__MOD_Draft_BN_Delta_10112)

Node: Mortality2

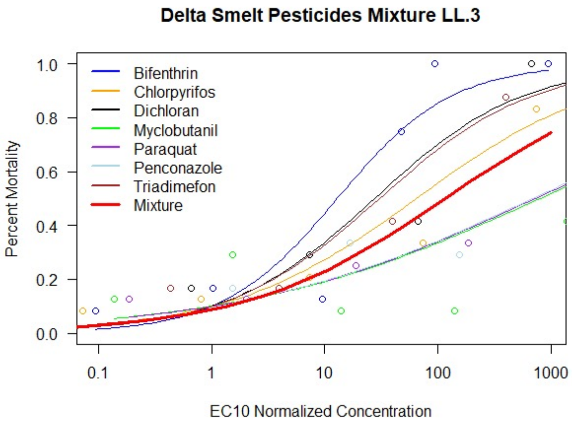
Chance | % Probability

Apply | OK | Reset | Close

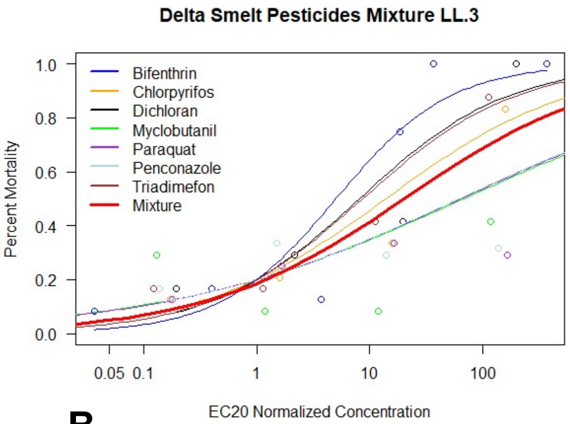
Chlorpyrifos	Bifenthrin	0	5	10	20	50
0 to 0.0165	0 to 0.00447	53.333	26.667	6.667	6.667	6.667
0 to 0.0165	0.00447 to 0.0107	6.667	26.667	53.333	6.667	6.667
0 to 0.0165	0.0107 to 0.0274	6.667	6.667	73.333	6.667	6.667
0 to 0.0165	0.0274 to 0.138	6.667	6.667	13.333	66.667	6.667
0 to 0.0165	0.138 to 260	6.667	6.667	6.667	6.667	73.333
0.0165 to 0.0682	0 to 0.00447	13.333	53.333	20	6.667	6.667
0.0165 to 0.0682	0.00447 to 0.0107	6.667	13.333	66.667	6.667	6.667
0.0165 to 0.0682	0.0107 to 0.0274	6.667	6.667	60	20	6.667
0.0165 to 0.0682	0.0274 to 0.138	6.667	6.667	6.667	73.333	6.667
0165 to 0.0682	0.138 to 260	6.667	6.667	6.667	6.667	73.333
2 to 0.317	0 to 0.00447	6.667	6.667	73.333	6.667	6.667
0682 to 0.317	0.00447 to 0.0107	6.667	6.667	73.333	6.667	6.667
0.0682 to 0.317	0.0107 to 0.0274	6.667	6.667	20	60	6.667
0.0682 to 0.317	0.0274 to 0.138	6.667	6.667	6.667	73.333	6.667
0.0682 to 0.317	0.138 to 260	6.667	6.667	6.667	6.667	73.333
0.317 to 4.397	0 to 0.00447	6.667	6.667	6.667	46.667	33.333
0.317 to 4.397	0.00447 to 0.0107	6.667	6.667	6.667	40	40
0.317 to 4.397	0.0107 to 0.0274	6.667	6.667	6.667	26.667	53.333
0.317 to 4.397	0.0274 to 0.138	6.667	6.667	6.667	26.667	53.333
0.317 to 4.397	0.138 to 260	6.667	6.667	6.667	6.667	73.333
4.397 to 5.06	0 to 0.00447	6.667	6.667	6.667	6.667	73.333
4.397 to 5.06	0.00447 to 0.0107	6.667	6.667	6.667	6.667	73.333
4.397 to 5.06	0.0107 to 0.0274	6.667	6.667	6.667	6.667	73.333
4.397 to 5.06	0.0274 to 0.138	6.667	6.667	6.667	6.667	73.333
4.397 to 5.06	0.138 to 260	6.667	6.667	6.667	6.667	73.333

Comparison of the exposure-response model curves for several chemicals using EC10 (A), EC20 (B) and EC50 (C)

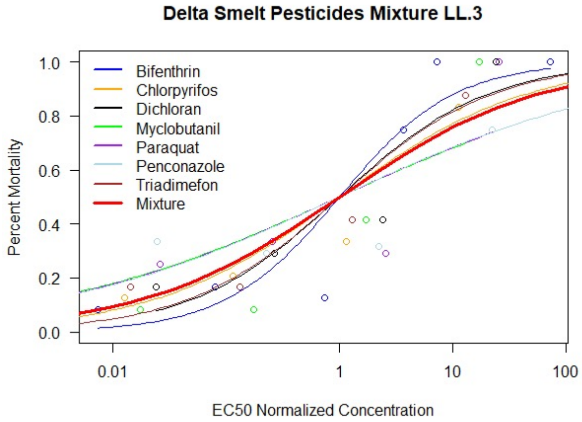
normalized concentrations



A

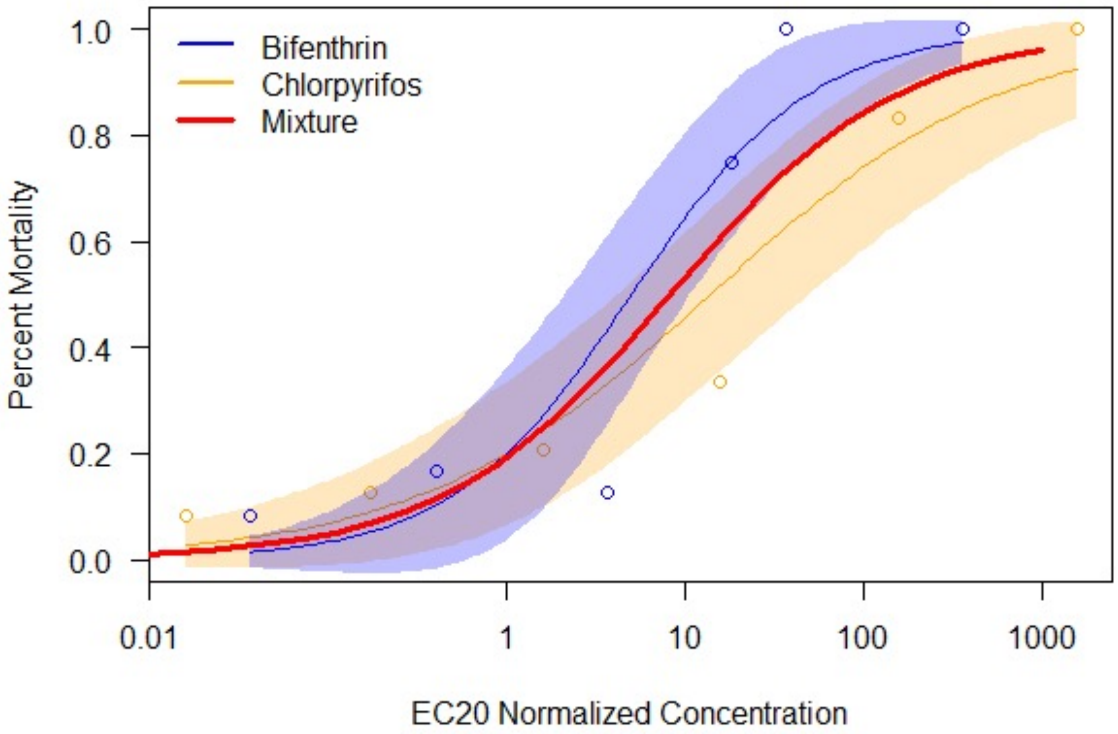


B

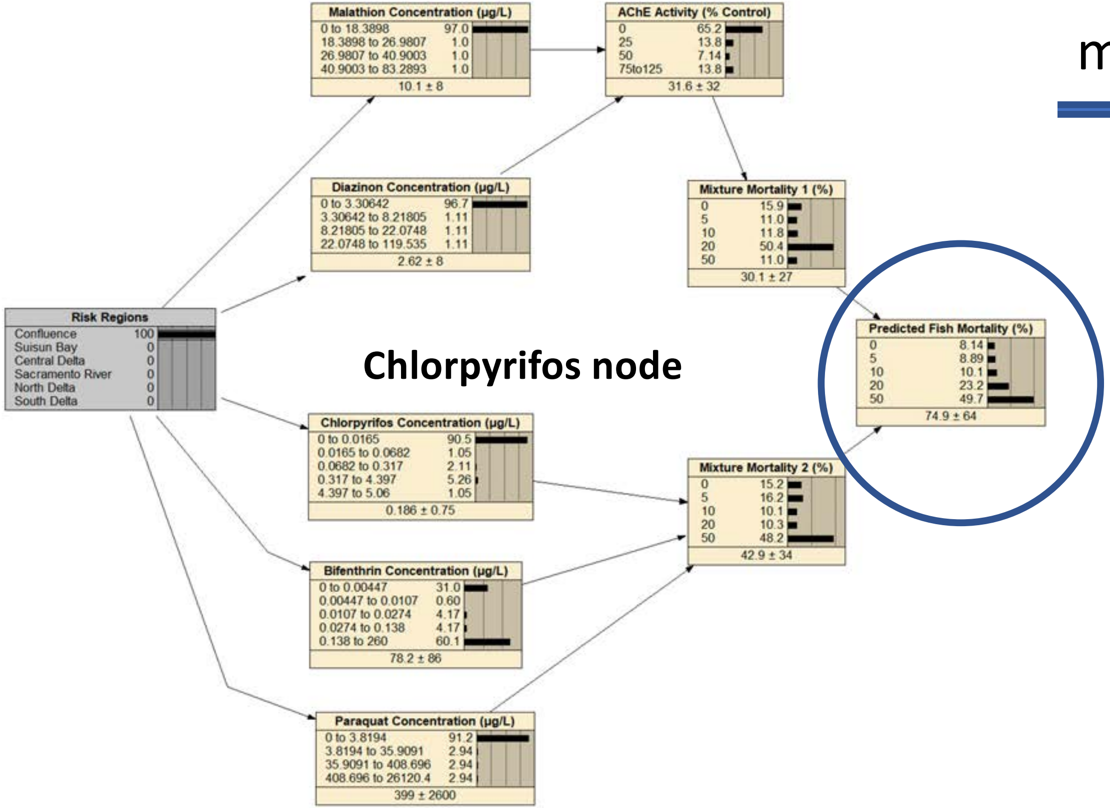


C

Example of additive exposure-response model with EC20 Normalization

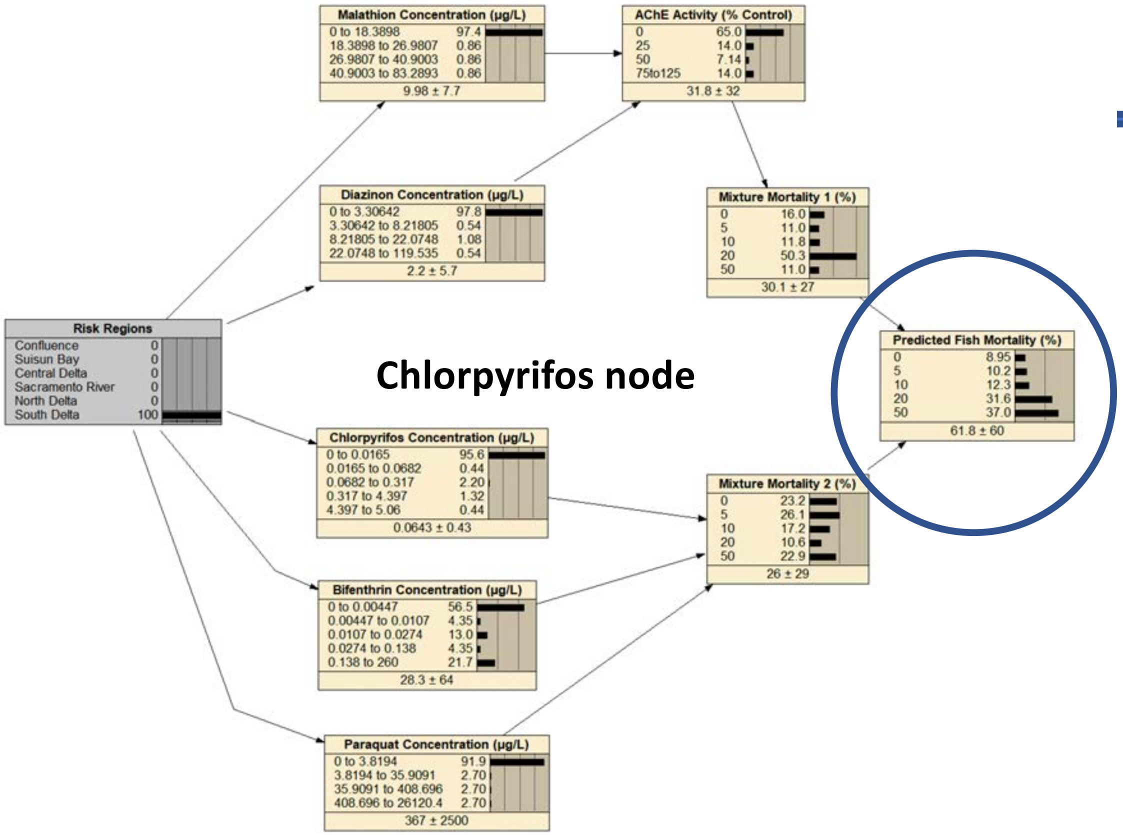


Risk Calculation for the Confluence from measured concentrations



There is an 83 percent probability of the Fish Mortality being equal to or greater than an EC10.

Risk Calculation for the South Delta from measured concentrations

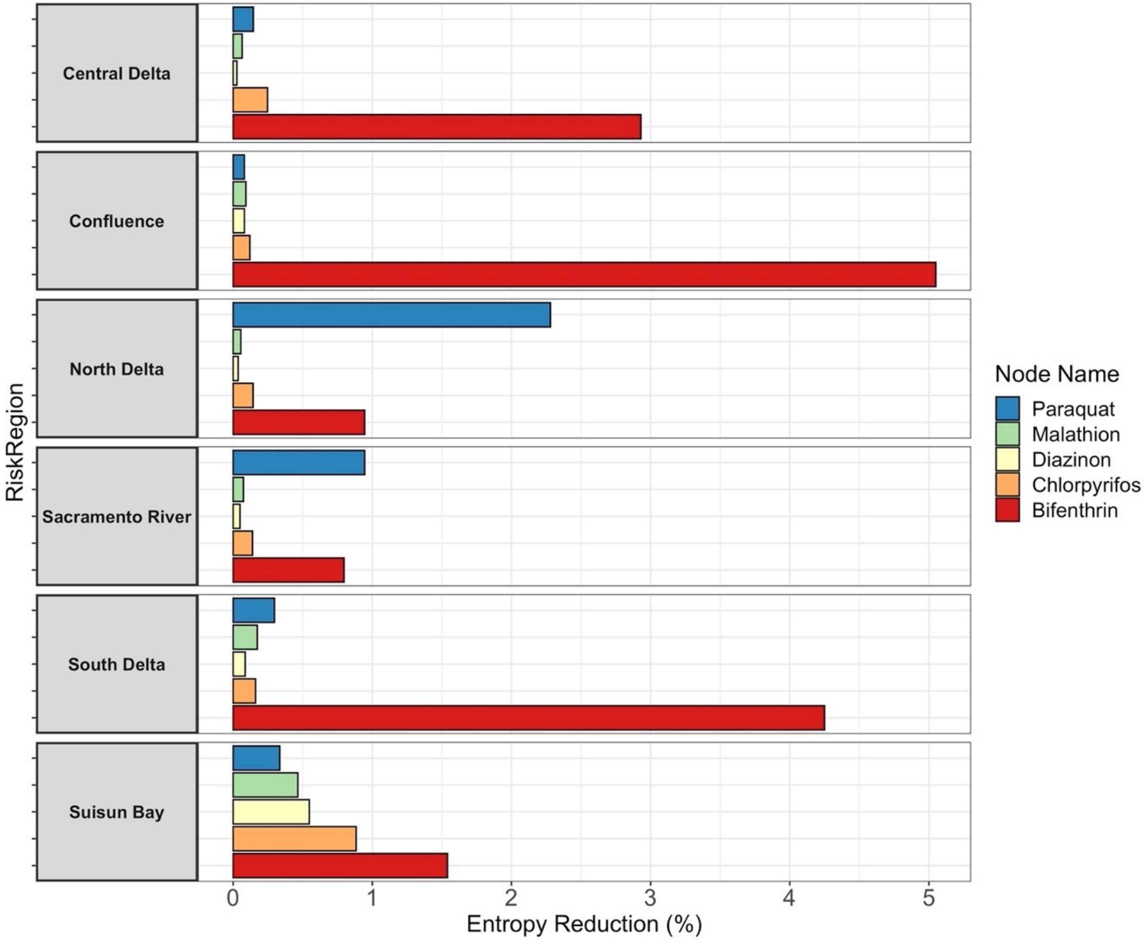


There is an 80.9 percent probability of the Fish Mortality being equal to or greater than an EC10.

Comparative risk for the USFE

Risk Region	Probability of Greater than an EC10
Confluence	83
Suisun Bay	79.6
Central Delta	78.9
Sacramento River	78.7
North Delta	80.4
South Delta	80.9

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.



Sensitivity Analysis using mutual information and

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).

Mazor RD, Rehn AC, Ode PR, Engeln M, Schiff KC, Stein ED, Gillett DJ, Herbst DB, Hawkins CP. 2014. Bioassessment in complex environments: designing an index for consistent meaning in different settings.

Data Sources

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

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

CEDEN. 2020. California Environmental Data Exchange Network Advanced Query Tool. Available online at https://ceden.waterboards.ca.gov/AdvancedQueryTool. Accessed Sep 9th, 2020.

ZoopSynth. 2020. Zooplankton data synthesizer: Version 1.1.1. Available online at https://deltascience.shinyapps.io/ZoopSynth/. Accessed Sep 9th, 2020.

Combining Data

Joining benthic macroinvertebrate samples with water quality samples from CEDEN database

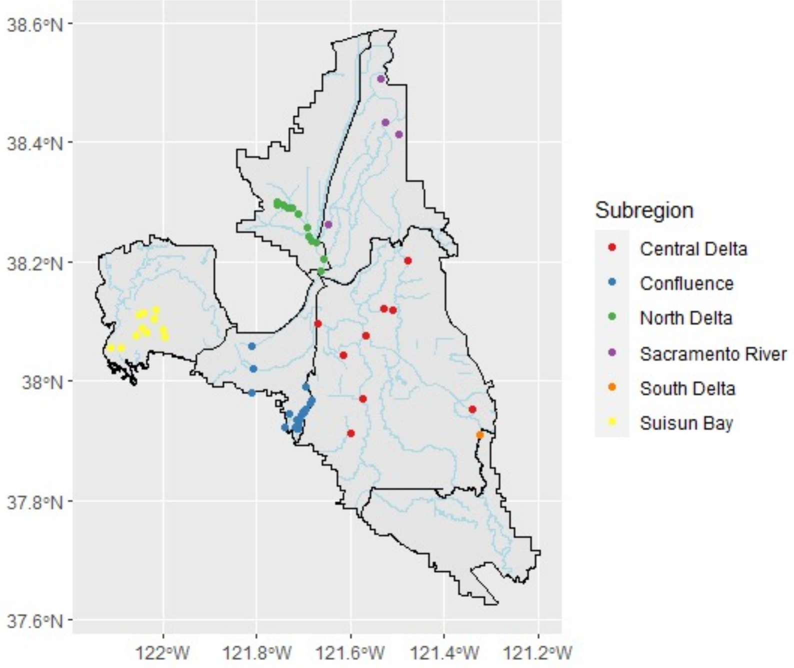
- Benthic samples do not include simultaneous water quality samples.

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

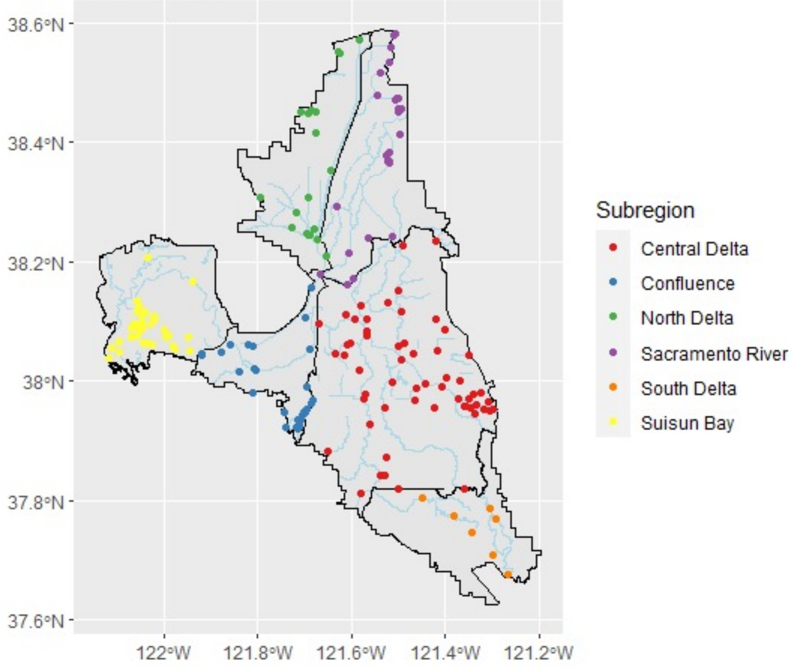
Pebesma E, Bivand R, Racine E, Sumner M, Cook I, Keitt T, Lovelace R, Wickham H, Ooms J, Muller K, Pedersen TL, Baston D. 2021. Package ‘sf’. CRAN. Available online at https://cran.r-project.org/web/packages/sf/sf.pdf. Accessed Jan 16th, 2021.

Combining Data

CEDEN Benthic Samples



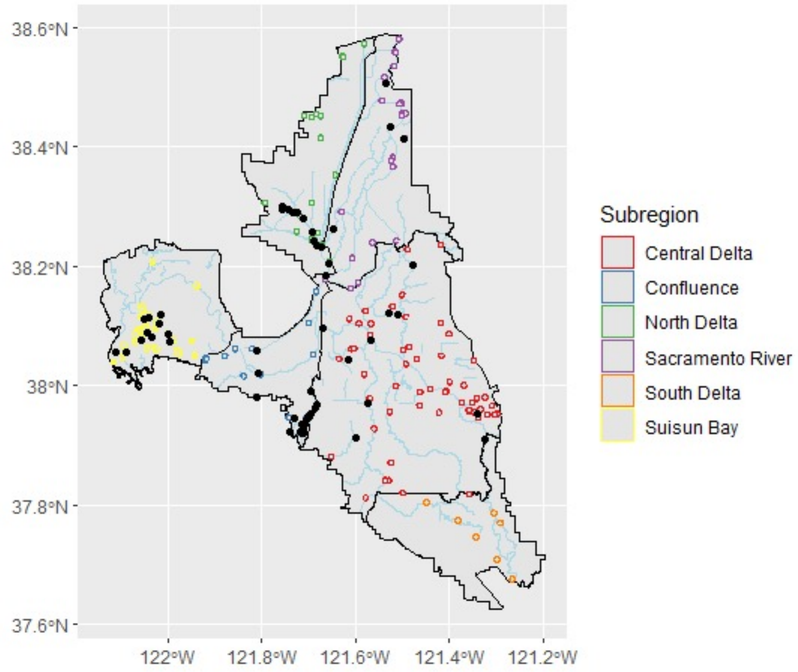
CEDEN Water Quality Samples



Pebesma E, Bivand R, Racine E, Sumner M, Cook I, Keitt T, Lovelace R, Wickham H, Ooms J, Muller K, Pedersen TL, Baston D. 2021. Package 'sf'. CRAN. Available online at https://cran.r-project.org/web/packages/sf/sf.pdf. Accessed Jan 16th, 2021.

Combining Data

CEDEN Benthic and Water Quality Samples

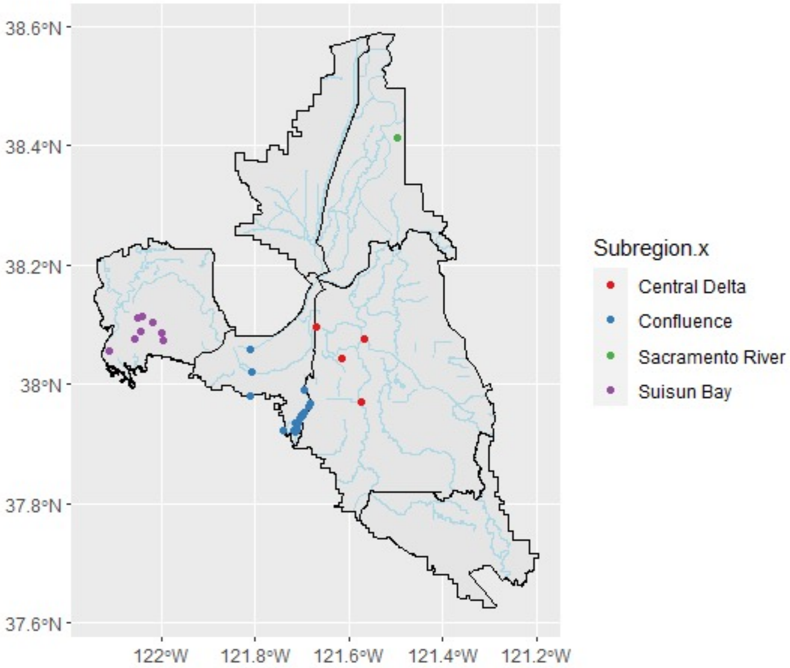


Spatial join of Benthic and Water Quality sample locations.

- Occur on the same day
- Are within 500 meters

Combining Data

Joined CEDEN Samples



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

Multivariate Analysis

- Metric: relative richness by taxa

$$Relative\ Richness = \frac{\#\ taxa}{Total\ taxa\ in\ sample}$$

- 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

Phylum	Order (within Arthropoda)
Mollusca	Mysida
Annelida	Diptera
Chordata	Hymenoptera
Cnidaria	Amphipoda
Nematoda	Thysanoptera
Platyhelminthes	Odonata
Bryozoa	Coleoptera
Nemertea	Trichoptera
	Ephemeroptera
	Hydrachnidia
	Collembola (Class)

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O’Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. 2020. Package ‘vegan’. CRAN. Available online at https://cran.r-project.org/web/packages/vegan/vegan.pdf. Accessed Jan 16th, 2021.

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.

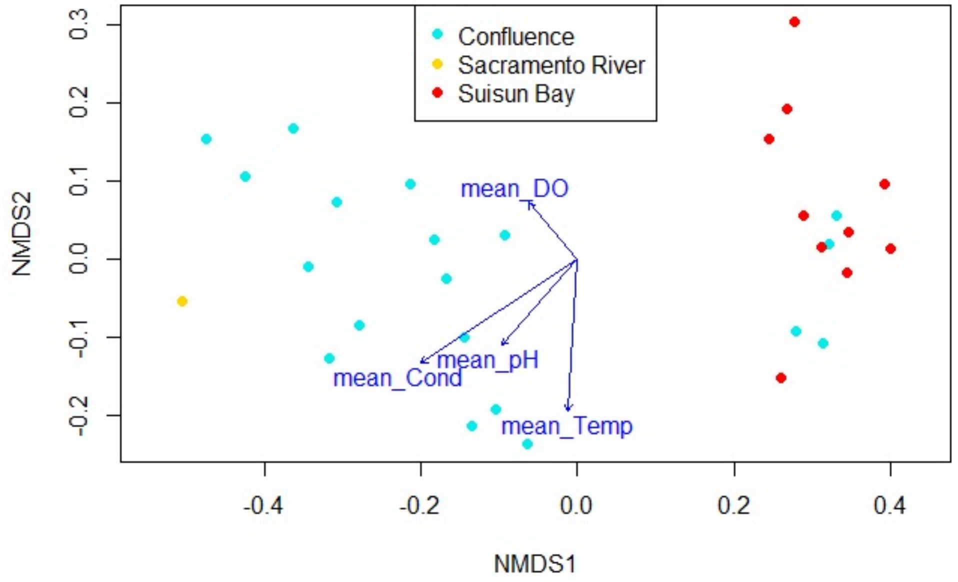
Non-Metric Multidimensional Scaling (NMDS)

- Allows you to specify the number of axes to simplify the data into using an iterative process.

Graham SE, Chariton AA, Landis WG. 2018. Using Bayesian Networks to Predict Risk to Estuary Water Quality and Patterns of Benthic Environmental DNA in Queensland. *Integrated Environmental Assessment and Management* 15(1):93-111.

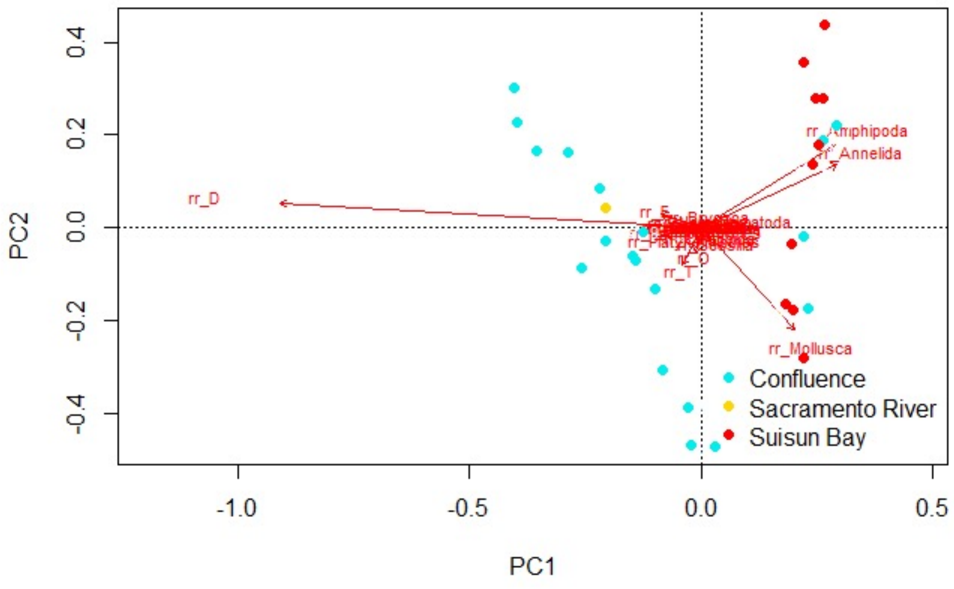
Non-metric Multidimensional Scaling

CEDEN Benthic MI clusters with water quality vectors



- 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.

Principal Components



Similar grouping pattern as NMDS

Main Phylums influencing grouping (potential endpoints):

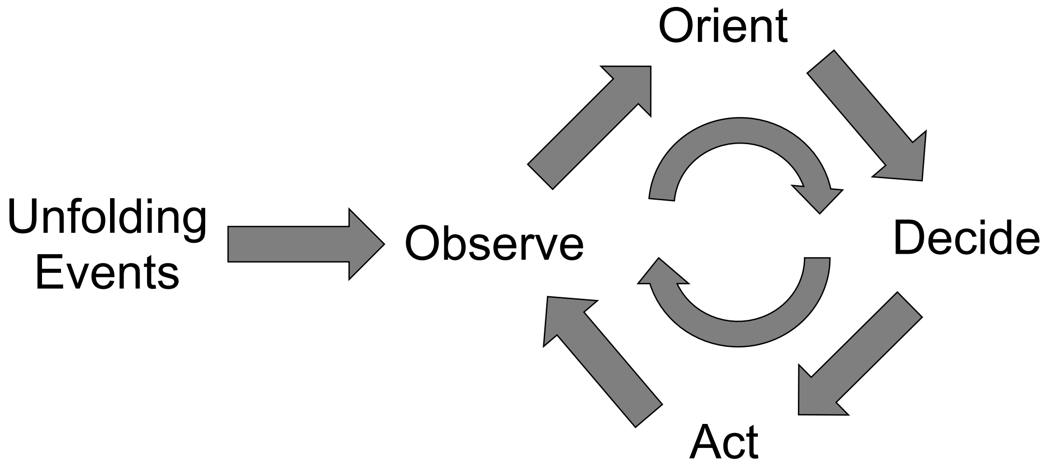
- Diptera
- Amphipoda
- Annelida
- Mollusca

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

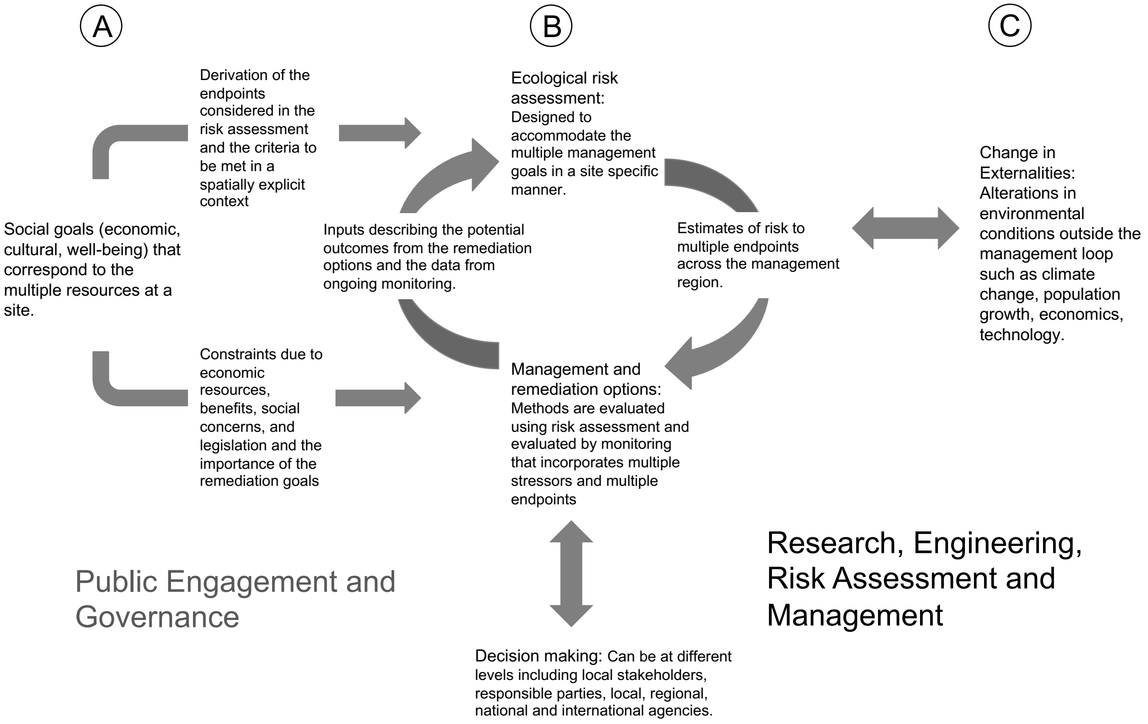
Why are we doing this…to make decisions.

Observe, Orient, Decide and Act—Loop



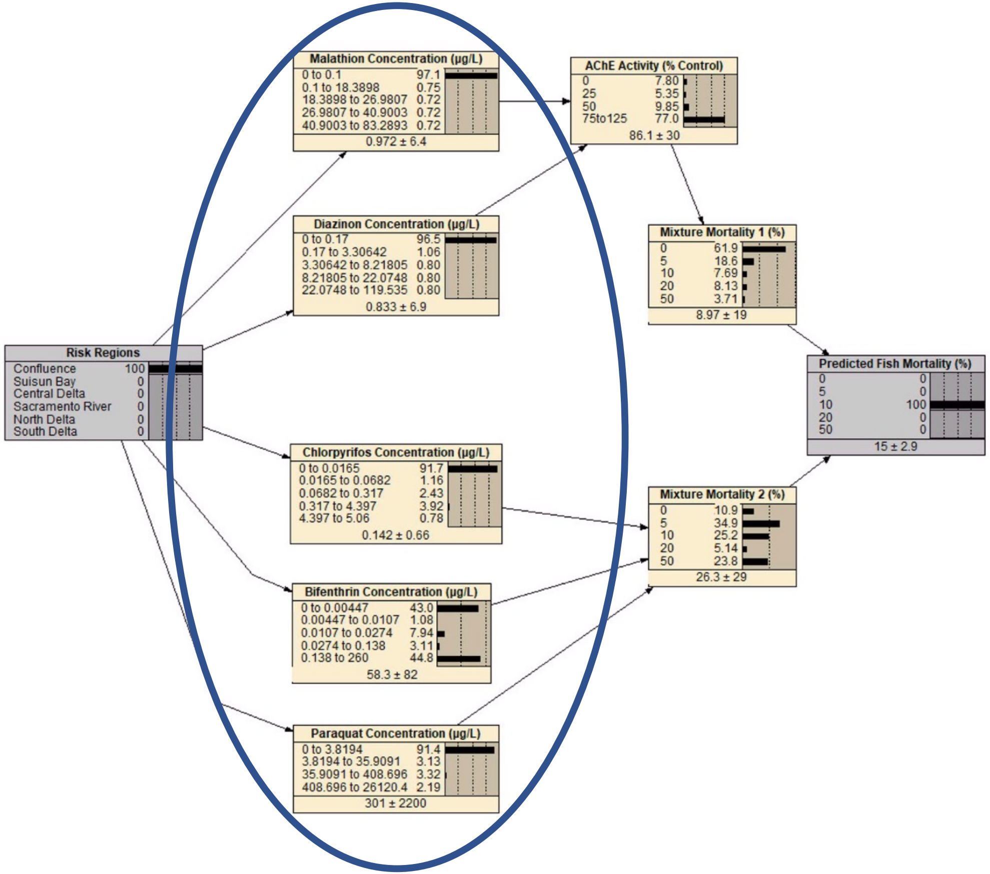
"OODA.Boyd" by Patrick Edwin Moran - Own work. Licensed under CC BY 3.0 via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:OODA.Boyd.svg#/media/File:OODA.Boyd.svg

Why are we doing this?…to make decisions.



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…..