

An Expedition into the Realm of Environmental Monitoring Data:

How Can We Make Sense and Make Use of the Information Contained in the Data?

Dan Wang, Jennifer Teerlink, Nan Singhasemanon, Kean S. Goh
California Department of Pesticide Regulation

ACS 2017



DATA



KNOWLEDGE

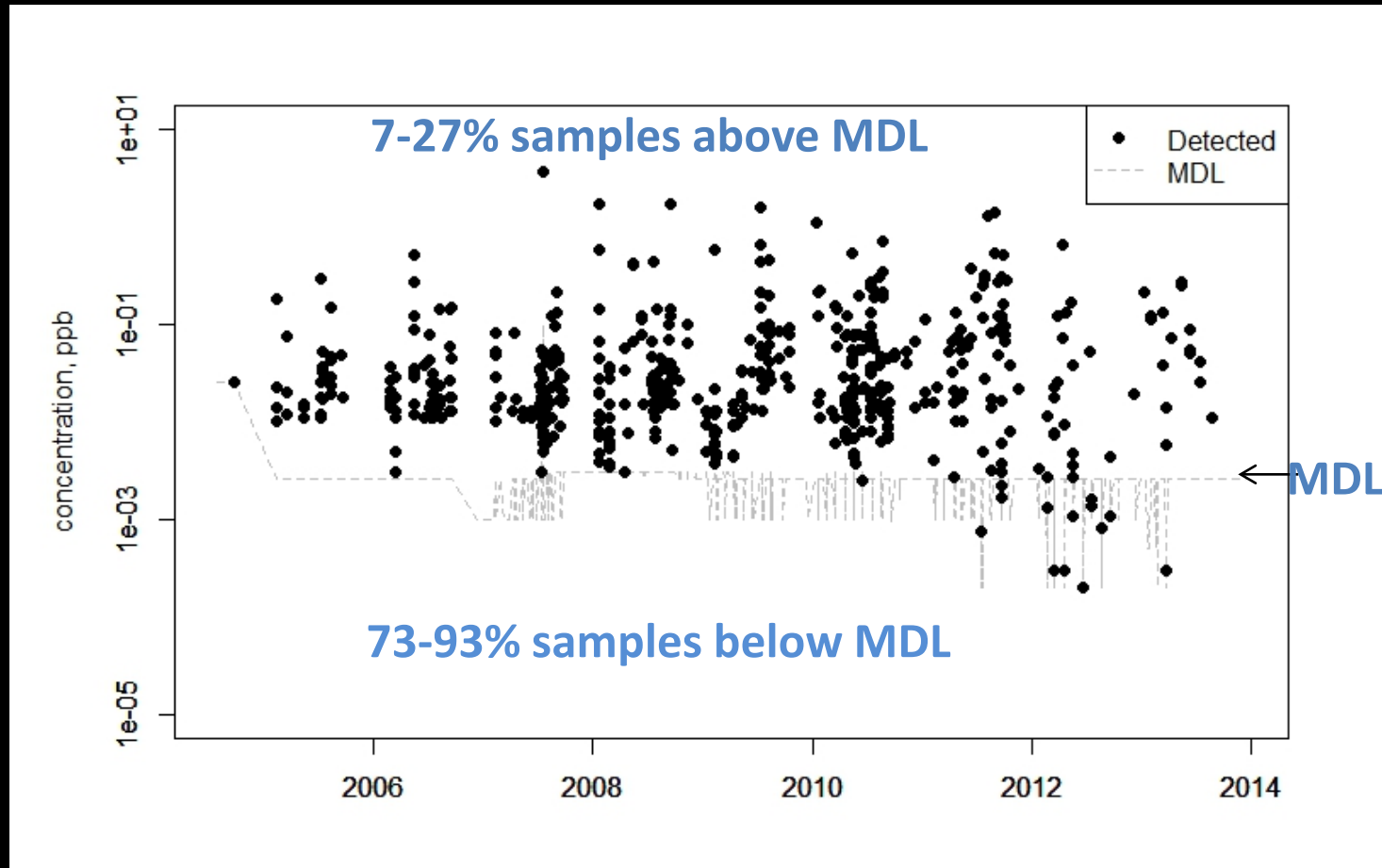


ACTION

Outline

- Challenges in interpreting monitoring data
- Illustration of statistical methods

Challenges in Interpreting Monitoring Data



- censored: vast amount non-detect, detection limit varies over time
- intermittent: irregular and infrequent sampling

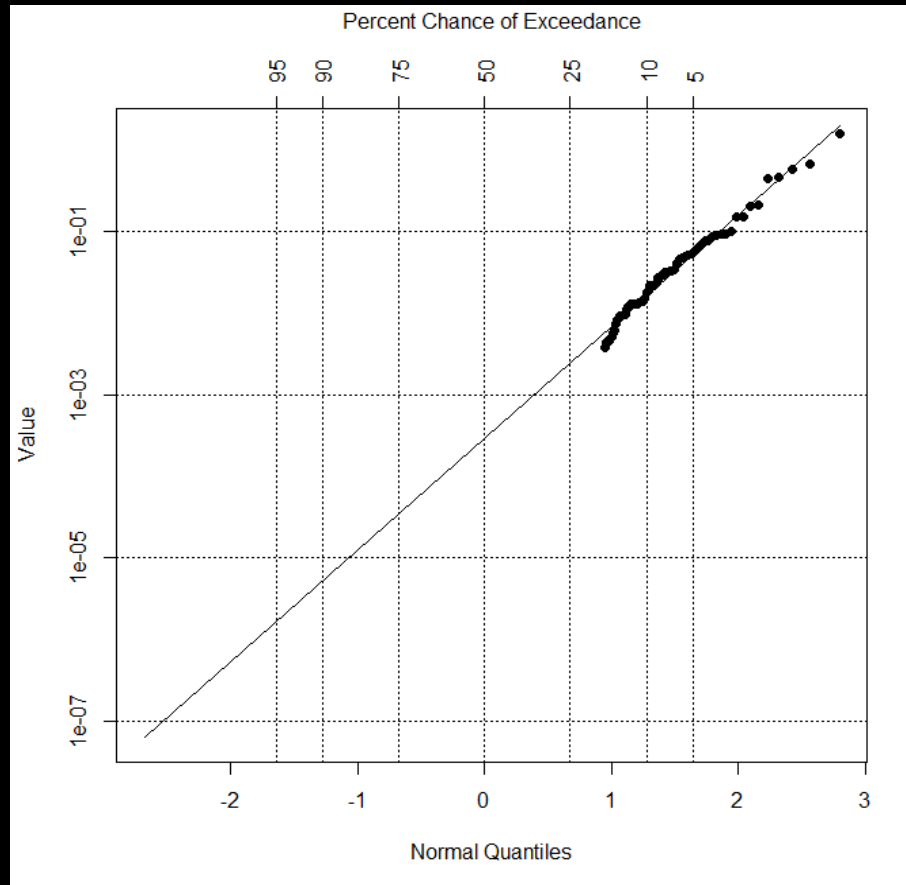
Statistical Methods for Censored Data

Two R packages:

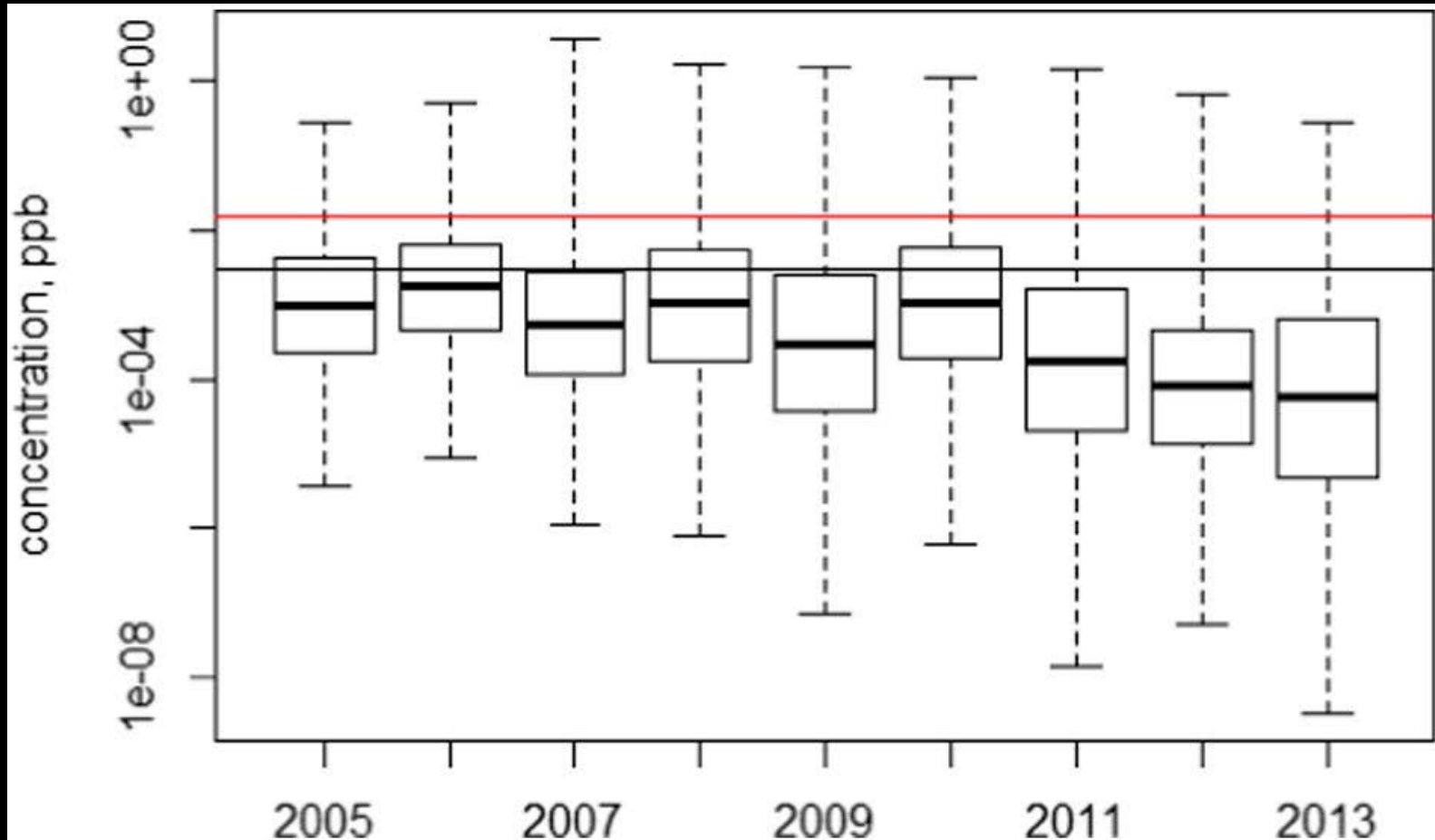
- NADA: Nondetects and Data Analysis for Environmental Data
- COIN: Conditional Inference Procedures in a Permutation Test Framework

Check Normality

Q-Q plot: quantile of the data against normal quantile

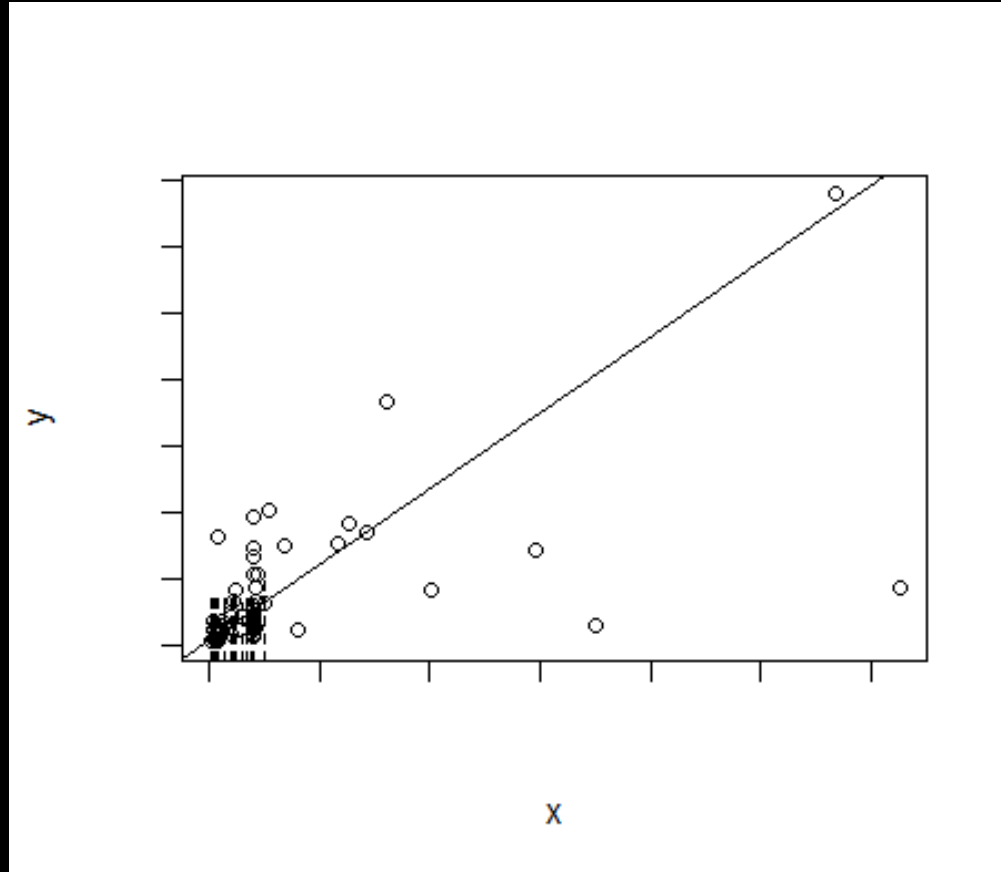


Censored Boxplot Visualization



black line: highest MDL; red line: water quality criteria

Censored Linear Regression

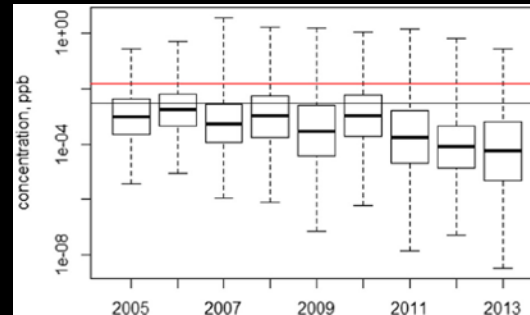
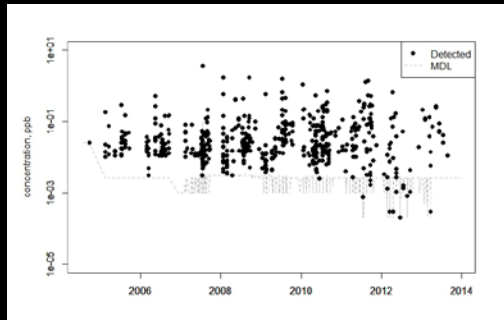


y: censored concentration; x: uncensored continuous variable, e.g., date;
vertical dash lines: censored data points, height represents detection limit

Censored Linear Regression

Parametric: Maximum-Likelihood-Estimated linear line, log-normality

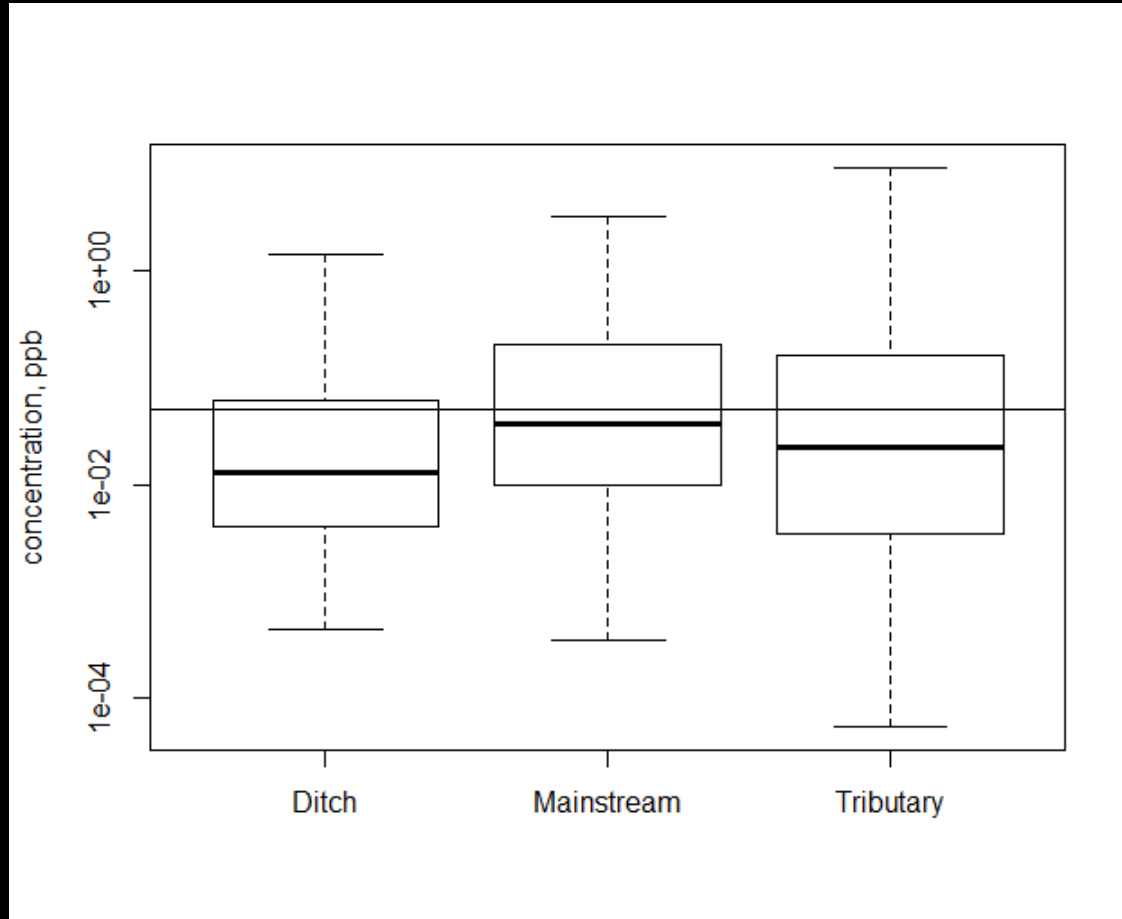
Semi-parametric: Mann-Kendal trend (ATS line), log-normal transformation



Chlorpyrifos, Central Valley Coalition Monitoring	Linear Regression Slope (p-value)		
	Simple Linear Regression	MLE Regression	M-K trend
2005-2013 data	9.69×10^{-5} (0.222)	-1.34×10^{-3} ($\ll 0.05$)	-8.41×10^{-4} ($\ll 0.05$)
2008-2013 data	2.72×10^{-5} (0.851)	-2.77×10^{-3} ($\ll 0.05$)	-2.28×10^{-3} ($\ll 0.05$)
2010-2013 data	-3.64×10^{-4} (0.234)	-4.55×10^{-3} ($\ll 0.05$)	-3.63×10^{-3} ($\ll 0.05$)

Compare Group Difference

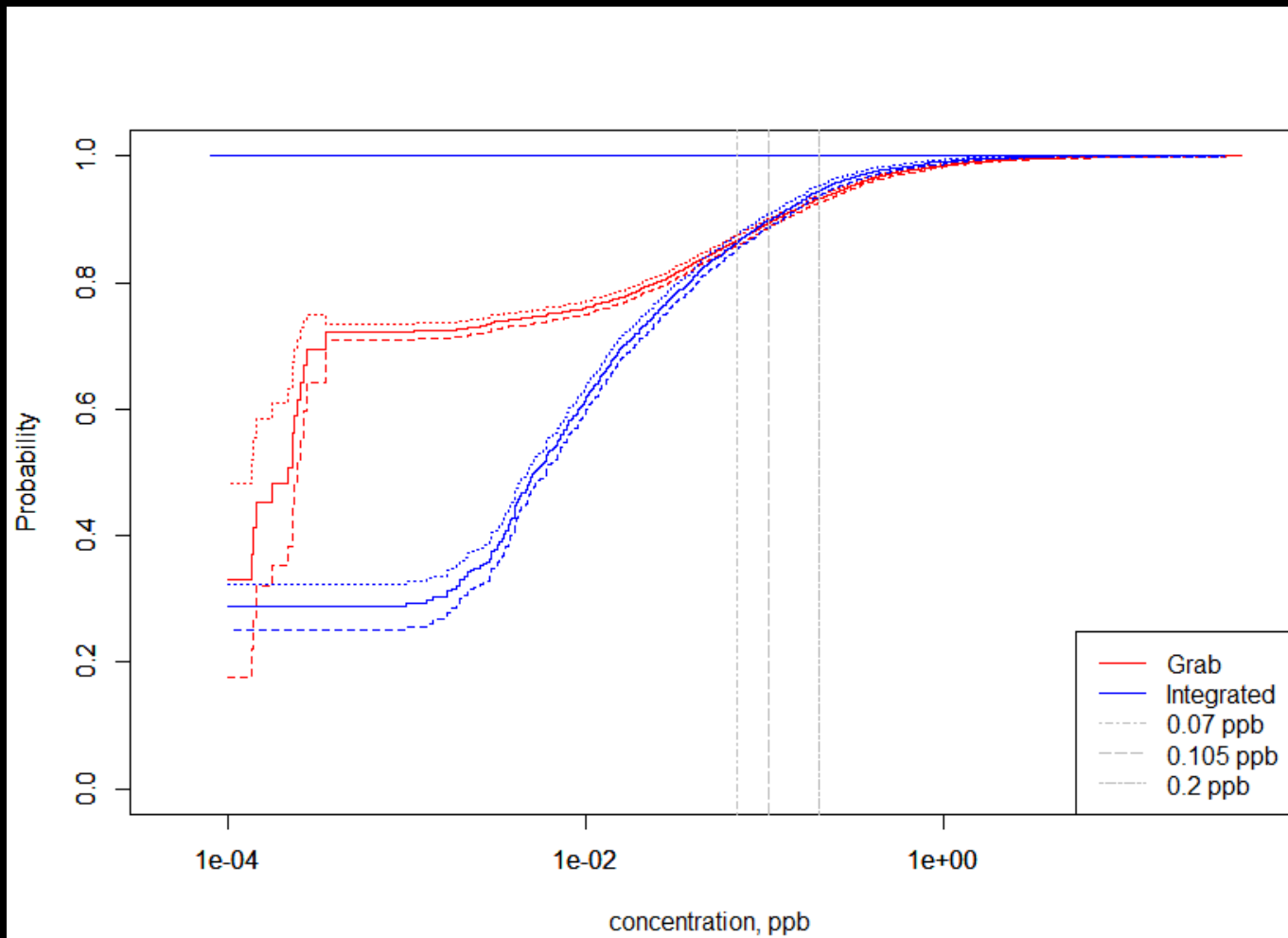
Censored linear regression with group ID, Like ANOVA



y: censored concentration; x: group indicator, e.g., sampling sites category

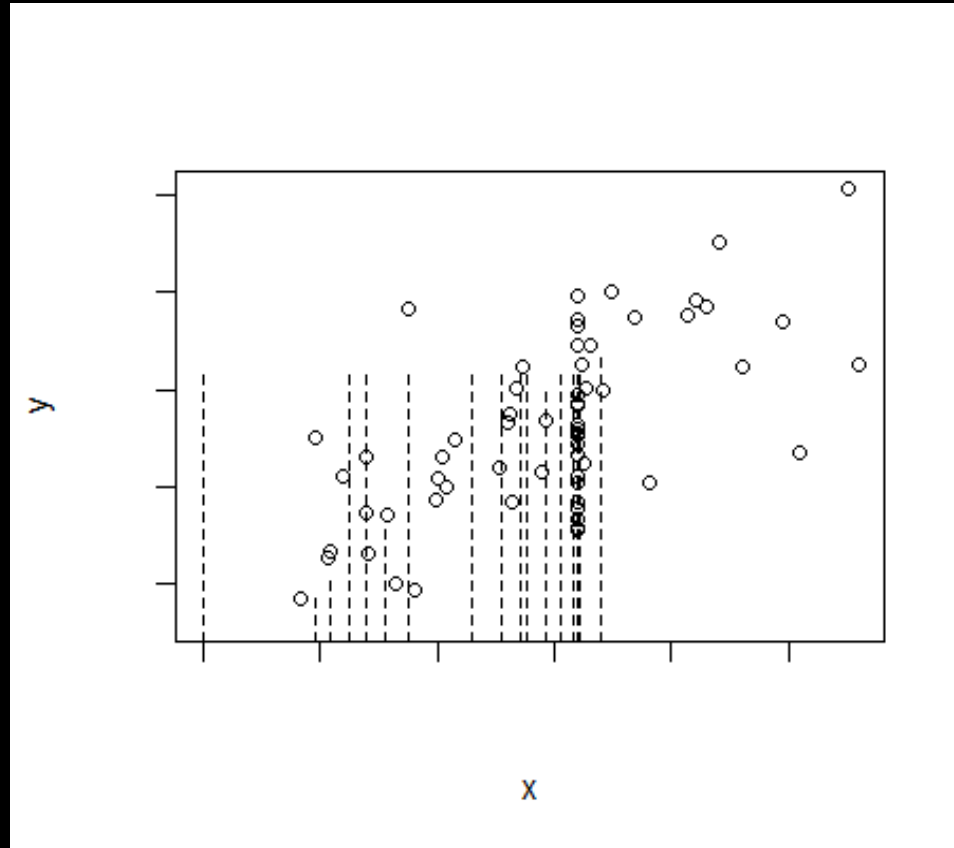
Compare Group Difference

Permutation comparison of cumulative distribution function (CDF)



Paired Comparison

Censored linear regression when both x and y as censored data



y and x are paired data

y: censored concentration from group 1; x: censored concentration from group 2

Statistical Methods for Intermittent/Irregular Sampling

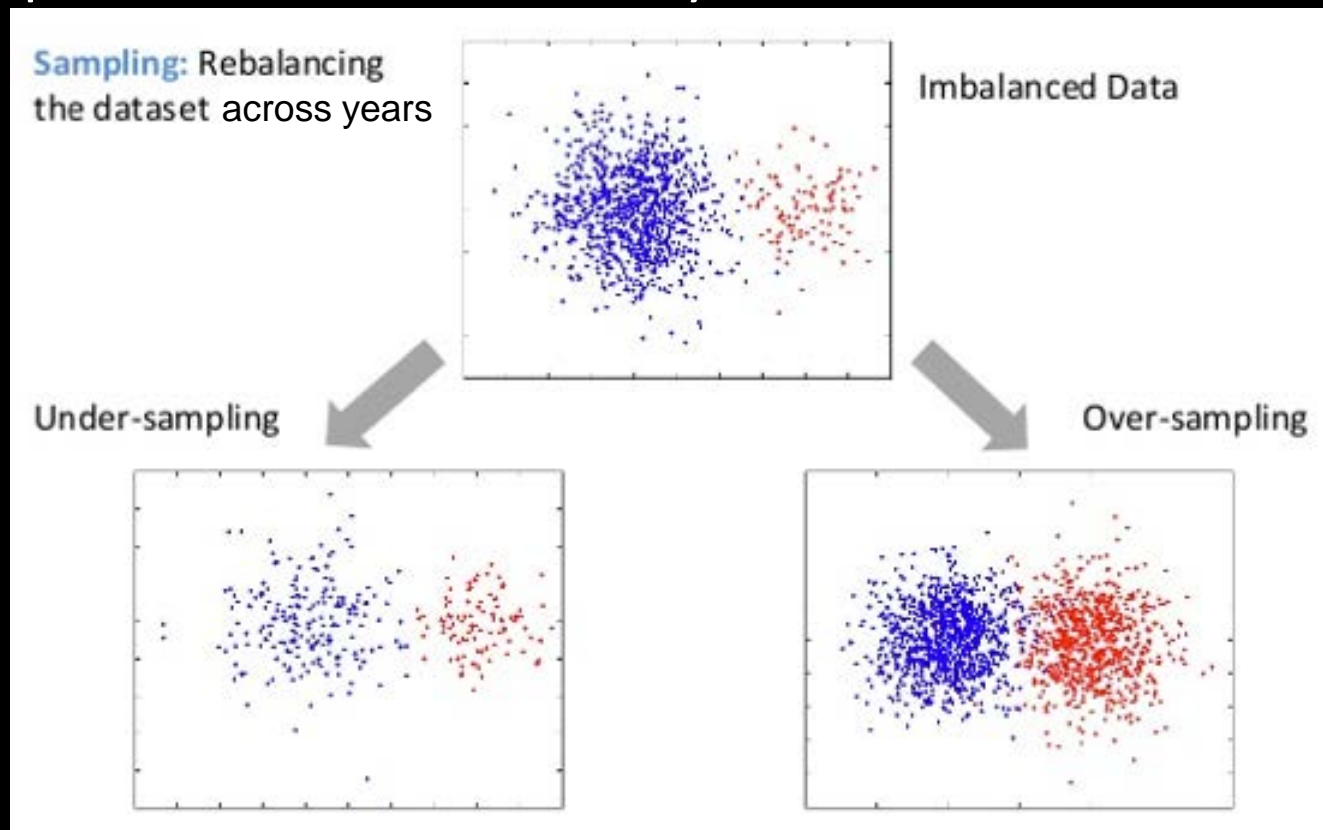
- Statistical Resampling methods, such as bootstrapping, can be used to estimate the uncertainty resulted from insufficient and irregular sampling.
- Bootstrapping: resampling with or without replacement from original dataset

Uneven Sampling Over Time

Water Year	Total	Dormant Season, 4 months	In-season, 8 months
1995	188	130	58
1996	60	15	45
2004	363	85	278
2010	759	206	553
2011	688	204	484

Stratified Bootstrapping

- The resampling datasets were generated by stratified random resampling (bootstrapping): n samples from dormant season, $2n$ samples from in-season each year.



References

Giddings, J. M., W. M. Williams, K. R. Solomon, and J. P. Giesy (2014), Risks to Aquatic Organisms from Use of Chlorpyrifos in the United States, in *Ecological Risk Assessment for Chlorpyrifos in Terrestrial and Aquatic Systems in the United States*, edited, pp. 119-162, Springer.

Hesterberg, T., D. S. Moore, S. Monaghan, A. Clipson, R. Epstein, B. A. Craig, and M. G.P. (2010), Bootstrap Methods and Permutation Tests, in *Introduction to the Practice of Statistics*, edited by D. S. Moore, G. P. McCabe and B. A. Craig, W.H. Freeman, New York.

Hothorn, T., K. Hornik, M. A. van de Wiel, and A. Zeileis (2015), COIN: Conditional Inference Procedures in a Permutation Test Framework (R package).

Lee, L. (2015), NADA: Nondetects And Data Analysis for environmental data (R package), edited.

Maas-Hebner, K. G., M. J. Harte, N. Molina, R. M. Hughes, C. Schreck, and J. A. Yeakley (2015), Combining and aggregating environmental data for status and trend assessments: challenges and approaches, *Environ Monit Assess*, 187(5), 1-16.

Moore, D. R., R. S. Teed, C. D. Greer, K. R. Solomon, and J. P. Giesy (2014), Refined avian risk assessment for chlorpyrifos in the United States, in *Ecological Risk Assessment for Chlorpyrifos in Terrestrial and Aquatic Systems in the United States*, edited, pp. 163-217, Springer.

TenBrook, P. L., A. J. Palumbo, and R. S. Tjeerdema (2009), Methodology for Derivation of Pesticide Water Quality Criteria for the Protection of Aquatic Life. Phase II: Methodology Development and Derivation of Chlorpyrifos Criteria. Chapter 2: Evaluation and Selection of Methods. Report Prepared for the Central Valley Regional Water Quality Control Board.

Williams, W. M., J. M. Giddings, J. Purdy, K. R. Solomon, and J. P. Giesy (2014), Exposures of aquatic organisms to the organophosphorus insecticide, chlorpyrifos resulting from use in the United States, in *Ecological Risk Assessment for Chlorpyrifos in Terrestrial and Aquatic Systems in the United States*, edited, pp. 77-117, Springer.

Wang, D., N. Singhasemanon, and K. S. Goh (2016), A statistical assessment of pesticide pollution in surface waters using environmental monitoring data: Chlorpyrifos in Central Valley, California, *Science of The Total Environment*, 571, 332-341.

Thank you!

Dan Wang, Ph.D.

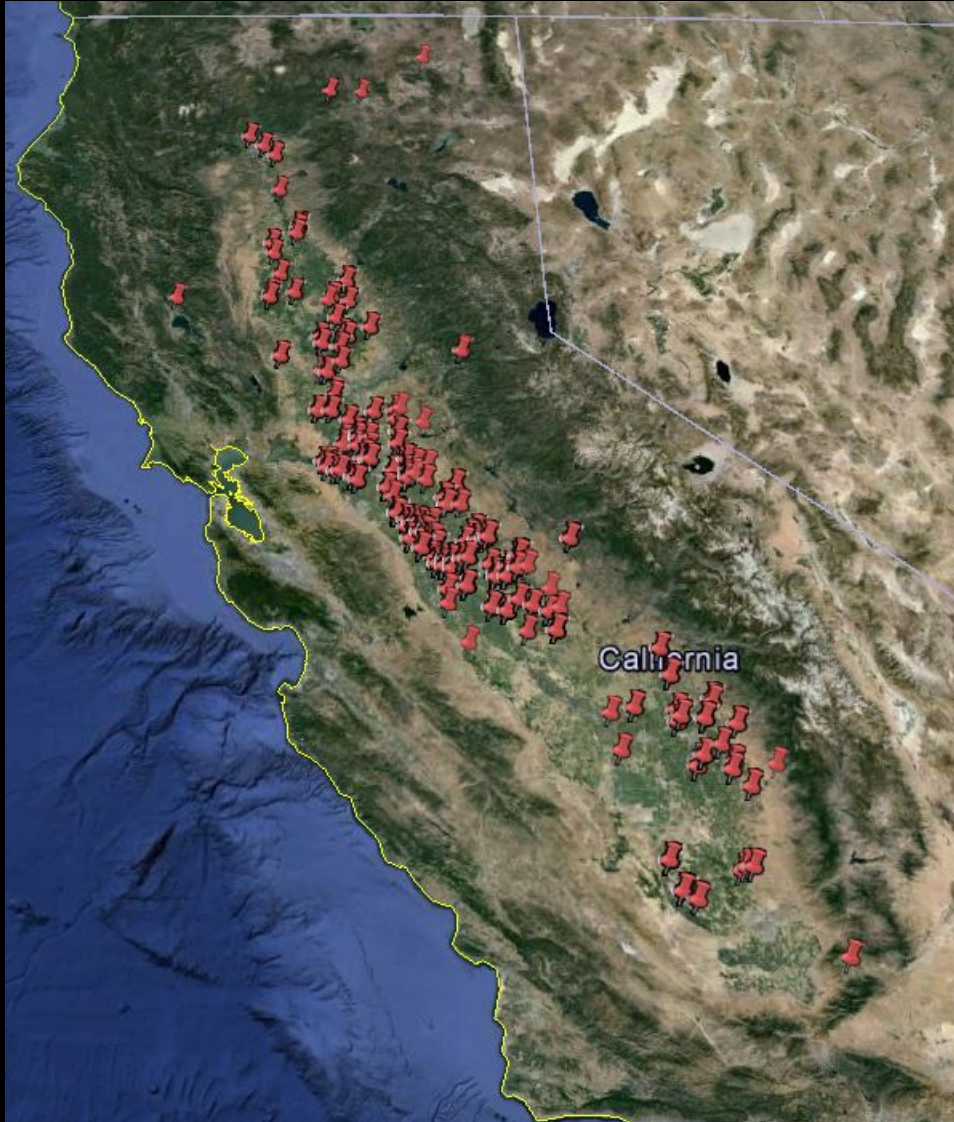
[Email: Dan.Wang@cdpr.ca.gov](mailto:Dan.Wang@cdpr.ca.gov)

Phone: 916-324-4201

Department of Pesticide Regulation

California Environmental Protection Agency

Special Consideration: Scope of the Dataset



- 3496 grab samples
- 175 sites
- 9 years from 2005-2013
- Highly polluted sites, not average condition