Event detection in ecological sensor networks

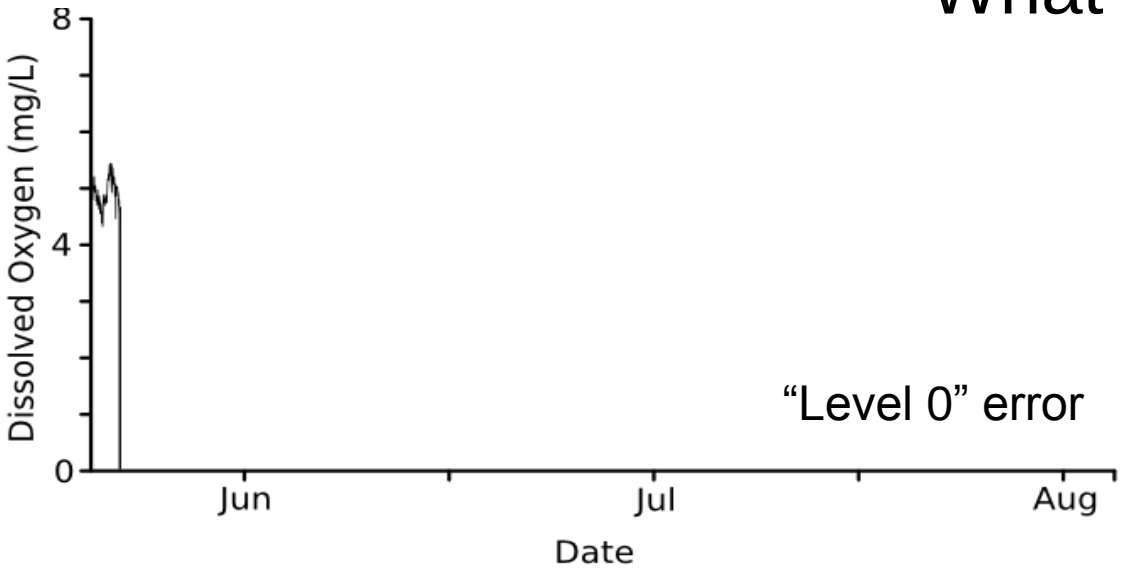
Owen Langman
Center for Limnology
University of Wisconsin - Madison

GLEON 7
Sept. 29, 2008
Norrtälje, Sweden

Outline

- Detecting sensor malfunctions
  - Motivation
  - Surprise Theory: how does it work?
  - Performance on real-world data
  - Implementation status
- Detecting ecological events
  - Integrating Surprise Theory with dynamic Bayesian networks

What are we trying to detect?

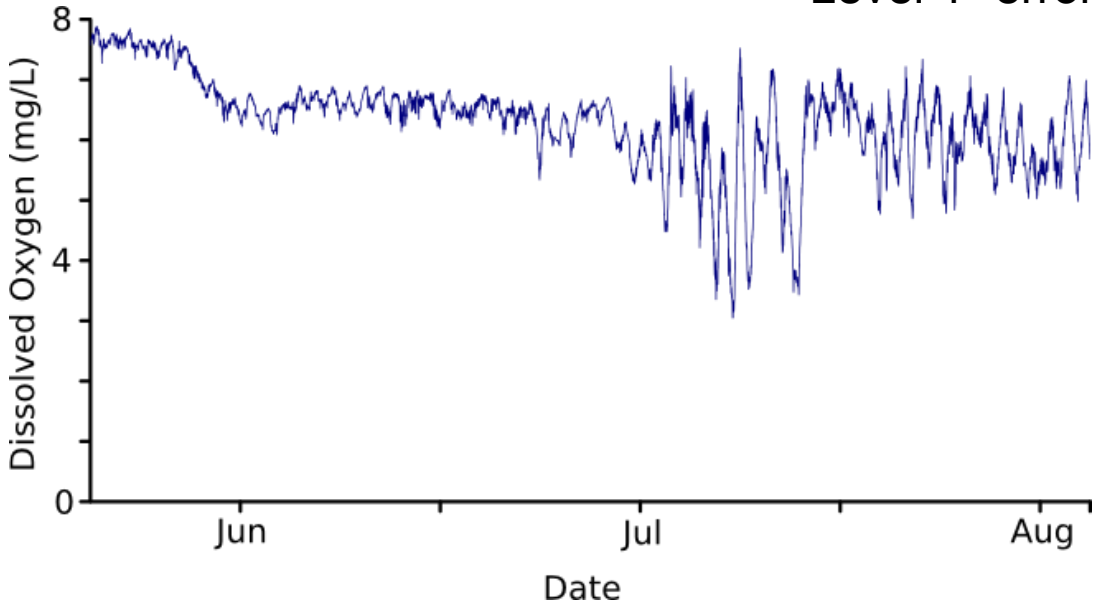


"Level 0" error

- No data returned
- Error value
- Sensor itself generates notification of error

"Level 1" error

This is what we want to detect!



Temporal

- Real-time observations
- Always-on sampling

Well, our data managers and QA/QC people do the detection now, what is wrong with their eyes scanning through data visualizations?

Spatial

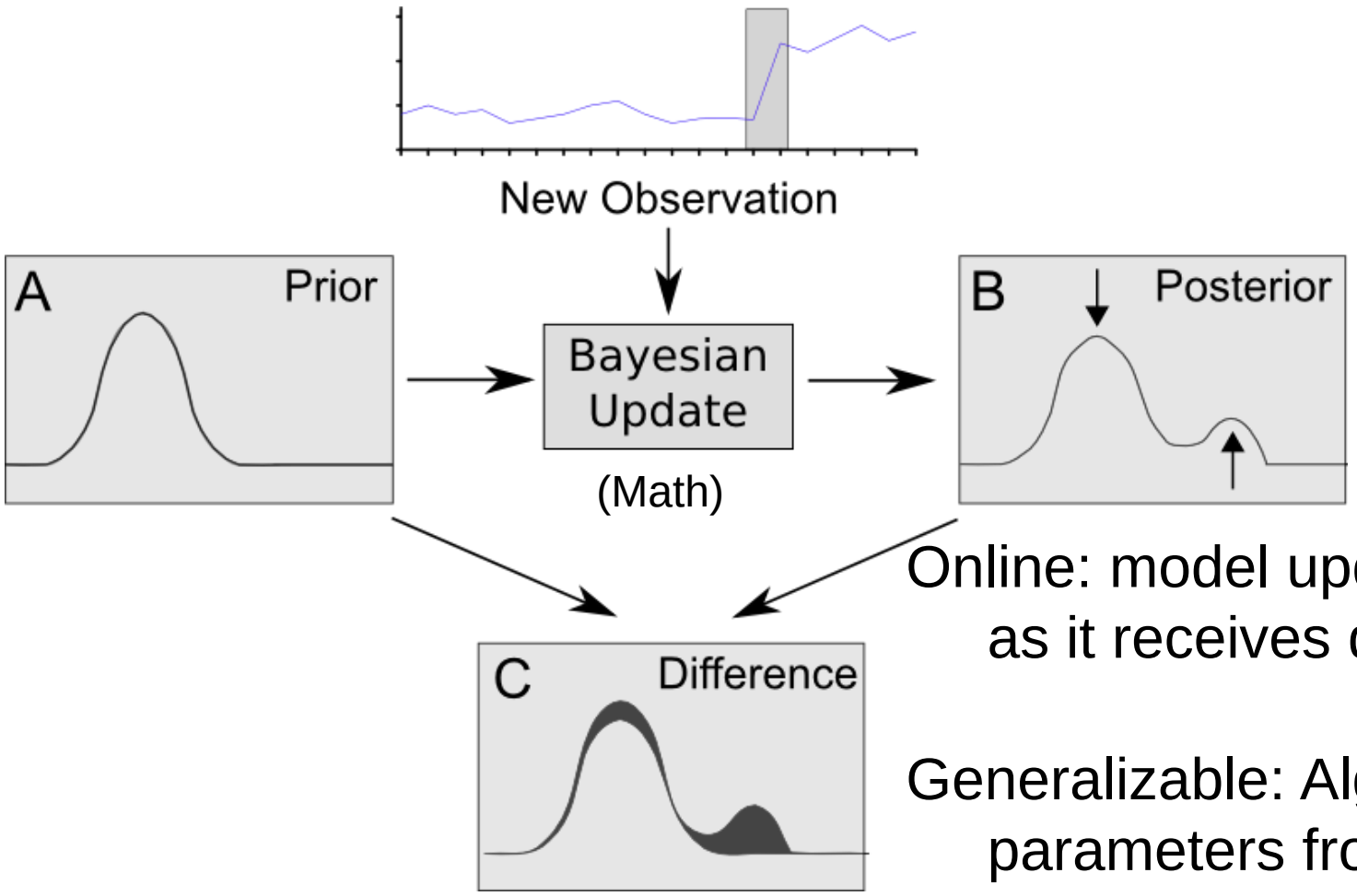
- Multiple platforms per site
- Many sites per organization

*

*Image source: NTL LTER webpage: http://www.limnology.wisc.edu

"Excessive" and Diverse Data

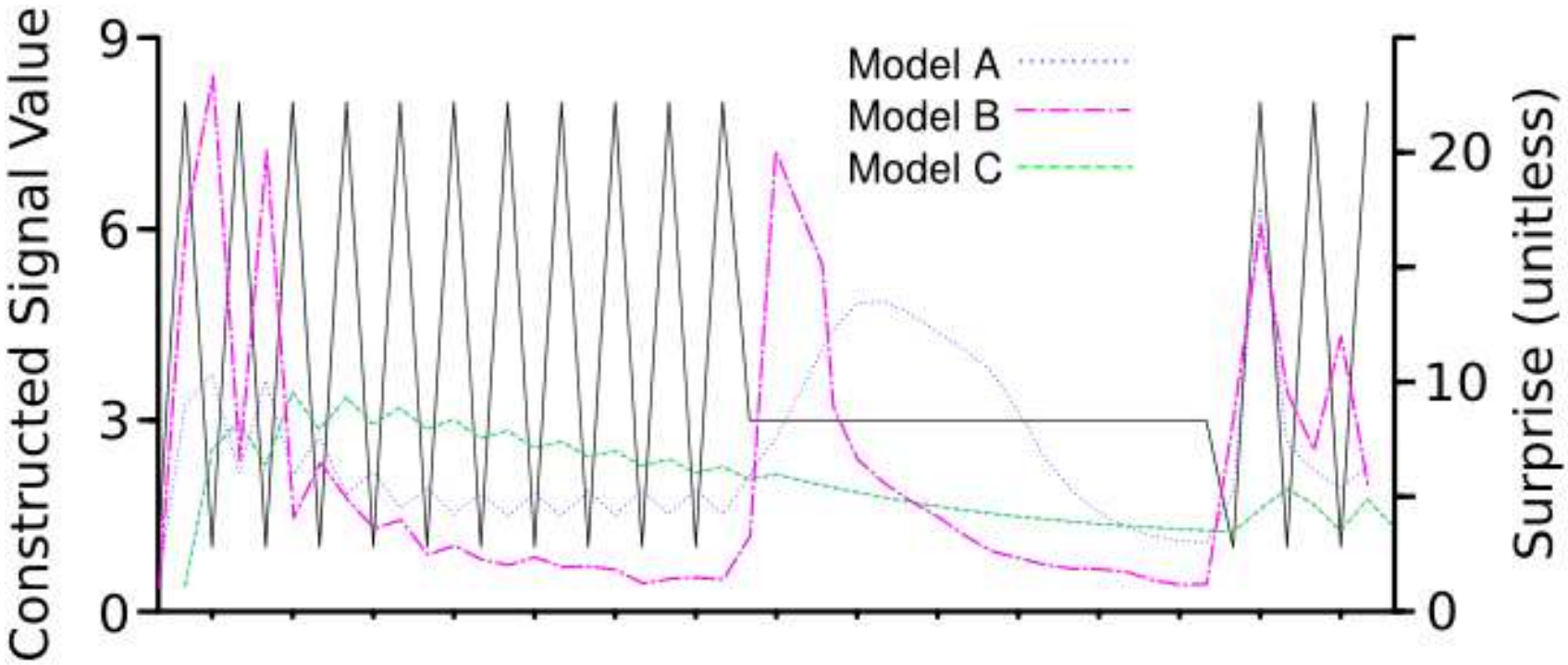
- ~20 million data for 2007 alone
- Many different types of sensors



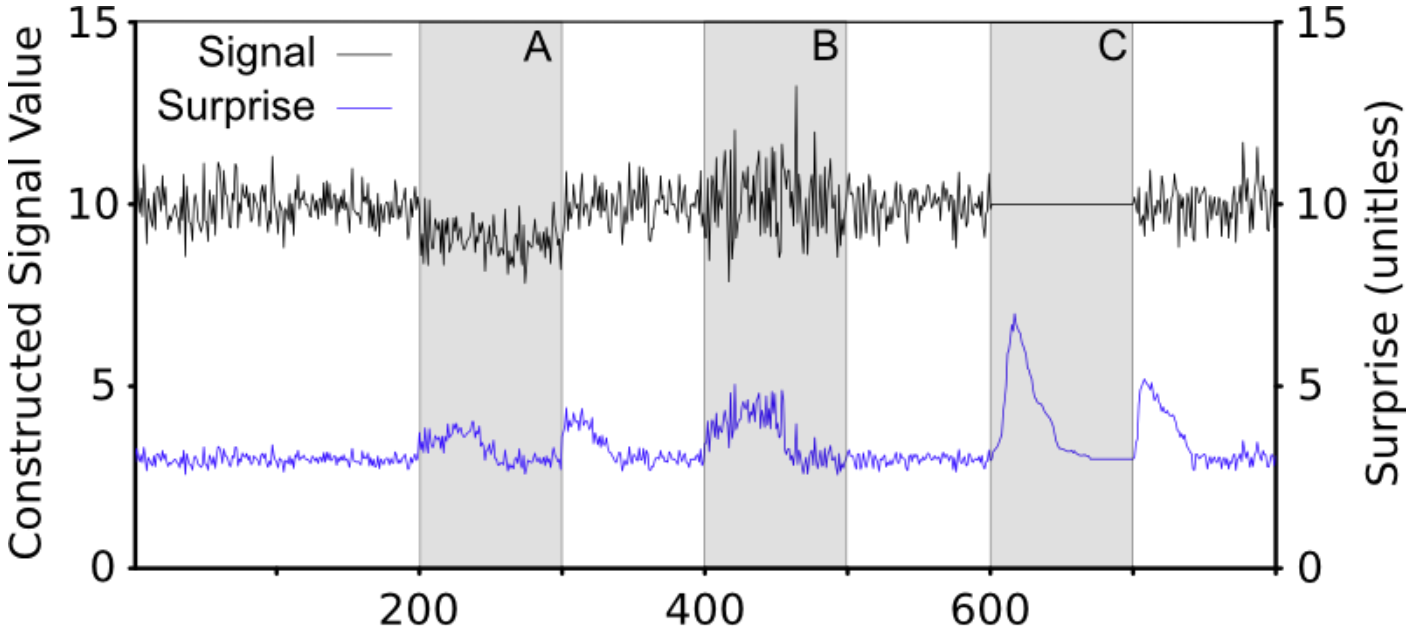
Online: model updates immediately as it receives data

Generalizable: Algorithm learns its parameters from the data and it can describe a large number of sensor traces

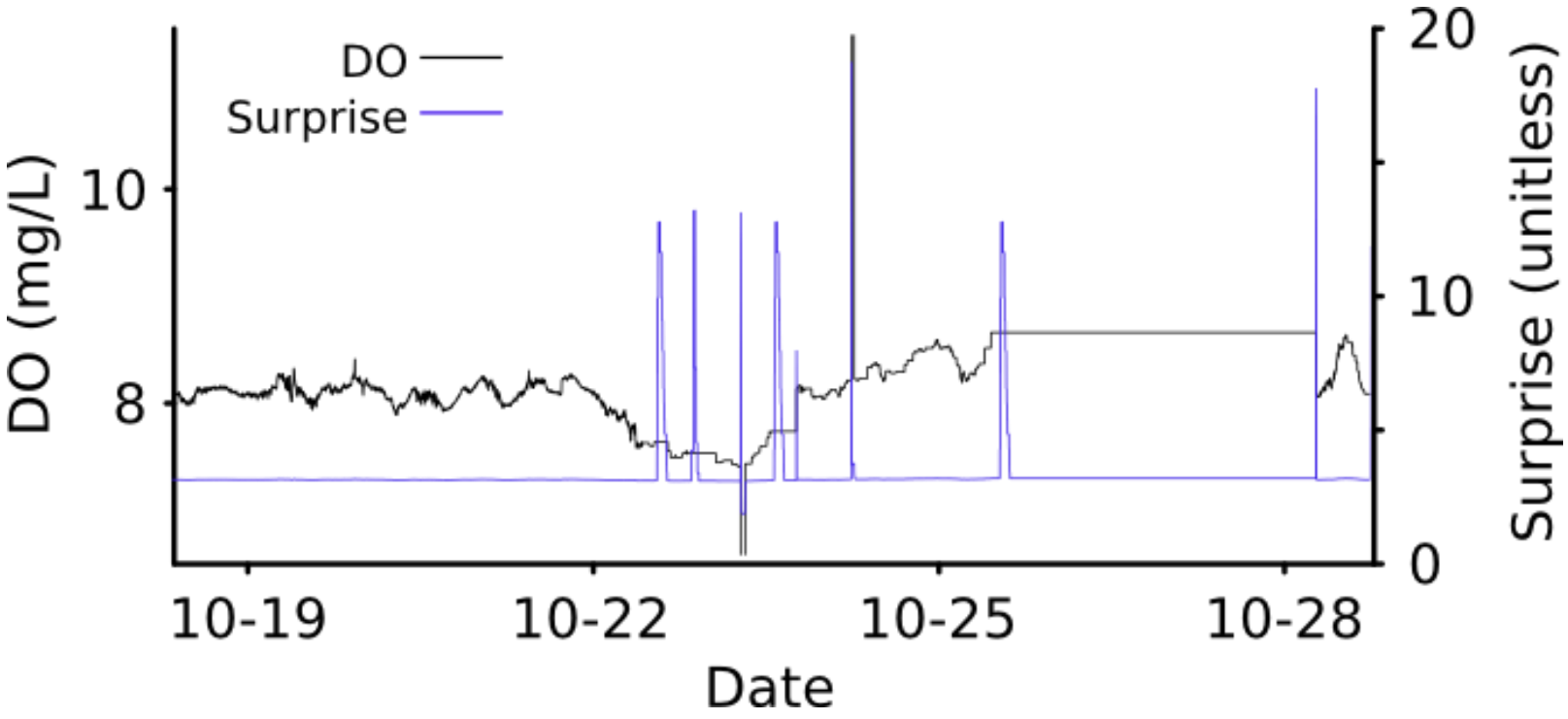
State driven: Computationally simple enough to put near the sensor



Model A: Gamma PDF, moderate sensitivity
Model B: Skew-elliptical (bimodal) PDF, moderate sensitivity
Model C: Gamma PDF, very low sensitivity

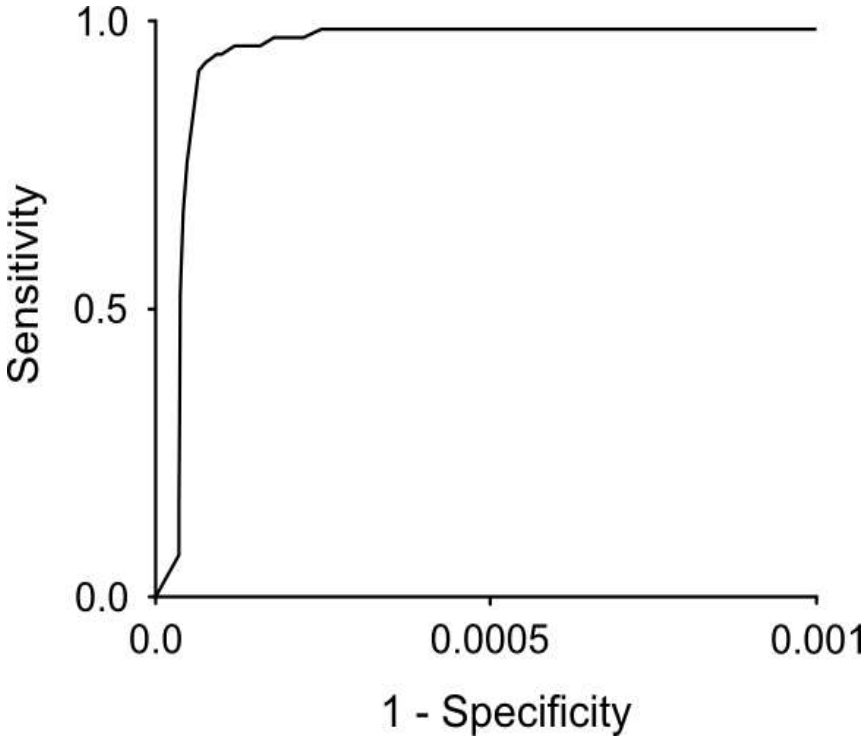


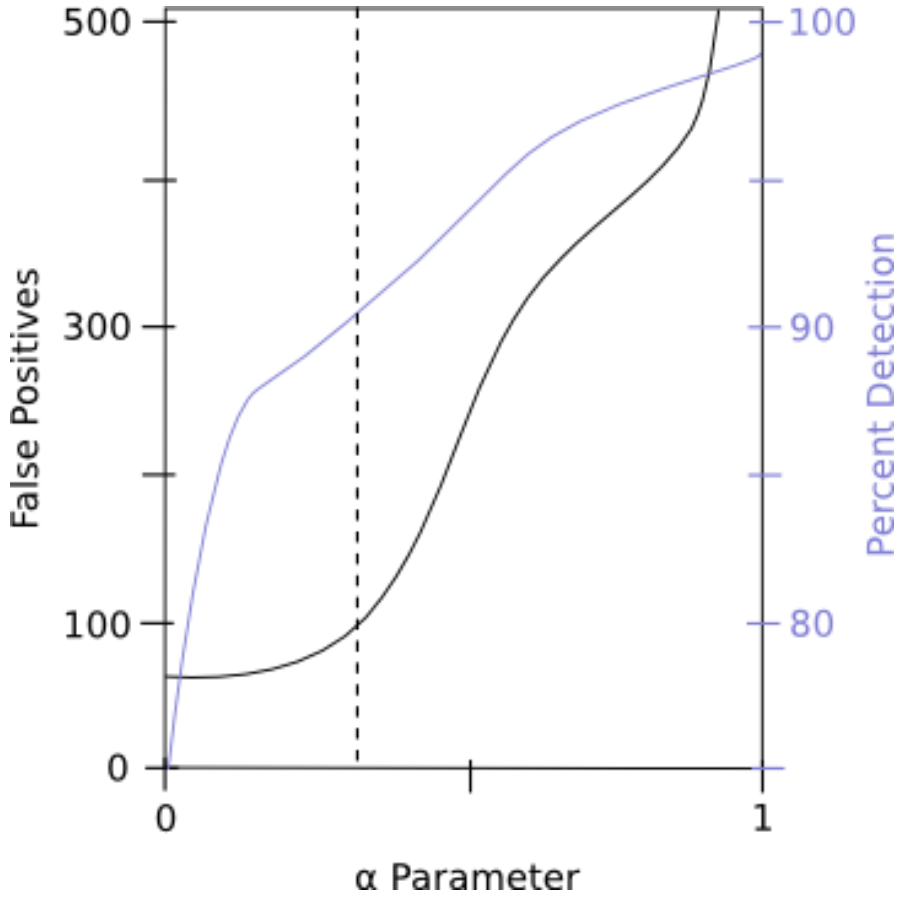
N(μ = 10, σ = 0.5)
A) Decrease in the mean to μ = 9
B) Increase in the variance to σ = 1
C) Decrease of the variance to σ = 0



Earlier detection?

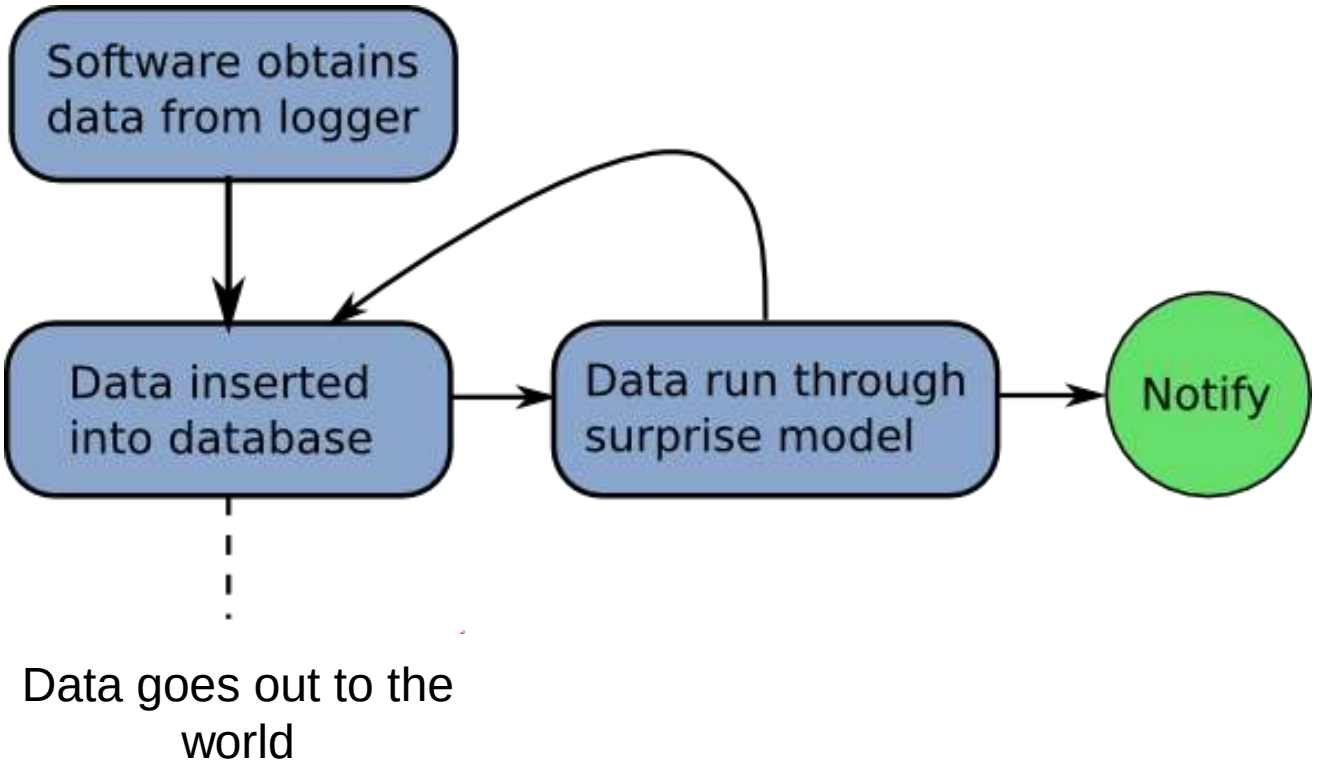
Performance

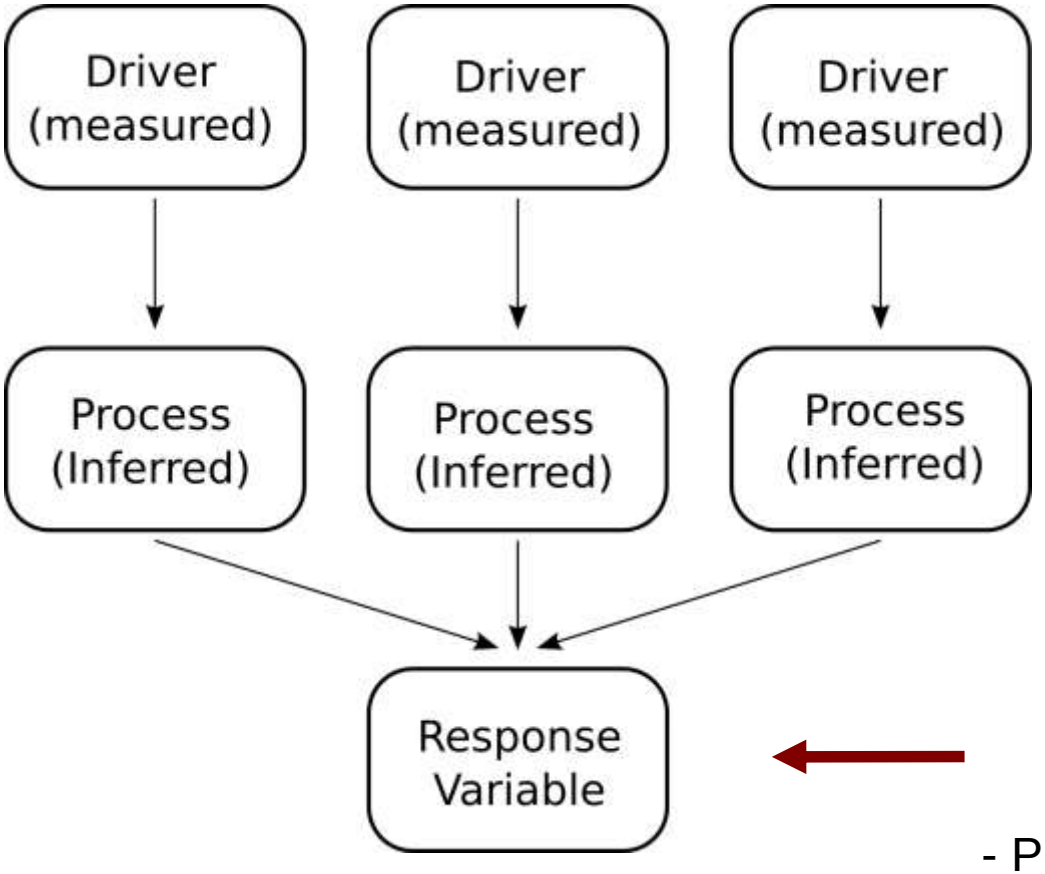




- 2 years of expert classified data
- Somewhat diverse collection of sensors (DO, temperature profile, PAR, wind, etc)

Status: Currently implementing this within the data flow in the Wisconsin buoy network

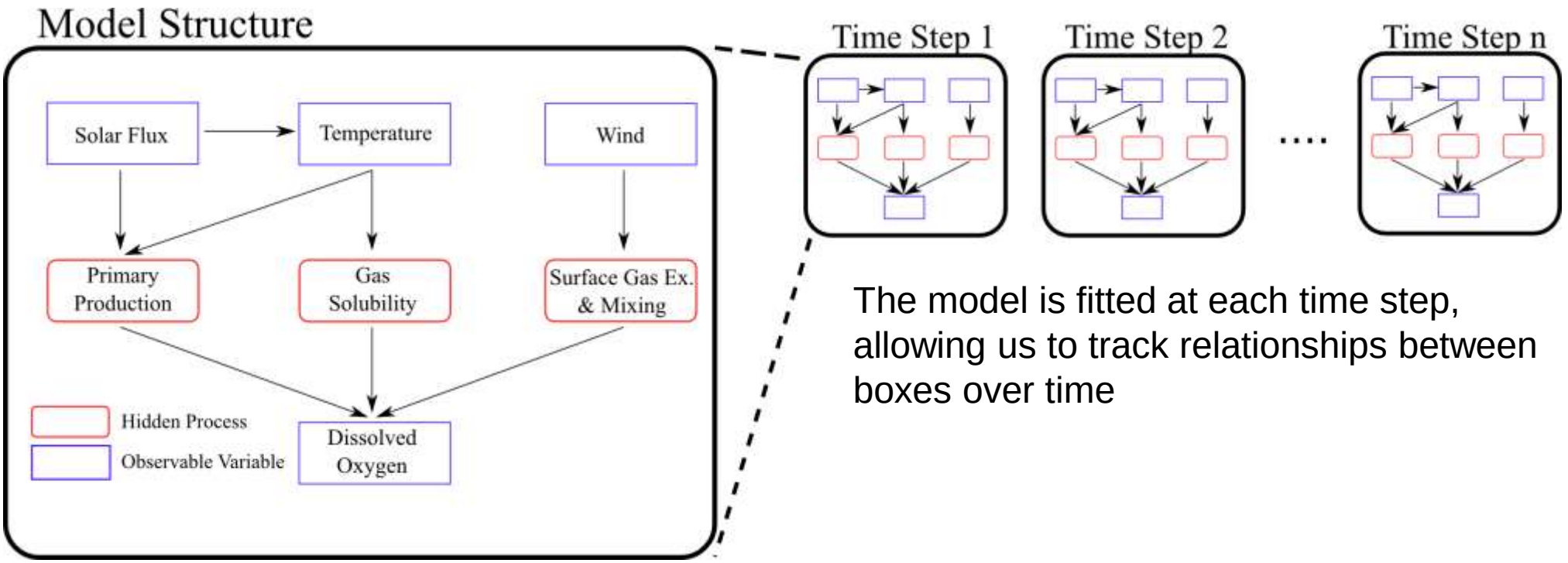




- Sensor malfunctions occur at the measurement nodes

- Ecological events occur at the process nodes

- Processes can have multiple drivers
- Response variables can be affected by multiple processes
- Perhaps if we examine the relationships between drivers-processes-responses, we'll be able to pick up changes in the processes controlling the response variables



The model is fitted at each time step, allowing us to track relationships between boxes over time

Dynamic Bayesian Networks:

- Model structure represents knowledge about the interactions of processes and measurements from our sensors
- Relationships (arrows) represented via probability density functions
- Using Surprise Theory, we can watch for changes in these relationships over time

Still in testing!
Stay tuned for details!

Questions!

Thanks:

Paul Hanson
Steve Carpenter
Luke Winslow
Kenneth Chiu

oclangman@wisc.edu