

Impacts of Varying Event Definitions for Event Detection System Performance

Sean A. McKenna¹, David B. Hart¹, Katie Umberg², Terra Haxton², John Hall², Sri Panguluri³

¹Sandia National Laboratories, Albuquerque, NM

²U.S. Environmental Protection Agency, Cincinnati, OH

³Shaw Environmental, Cincinnati, OH

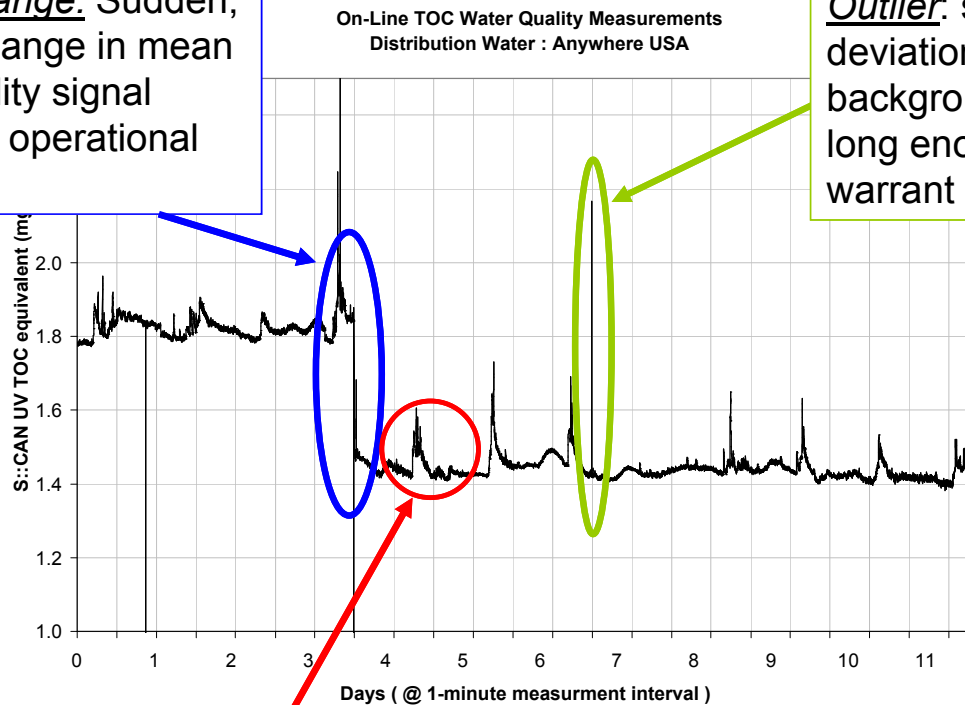
contact: samcken@sandia.gov

Outline

- Motivation: what are we (not) hoping to find
 - Event strength vs. Longevity
- How things work now
 - Fusion across sensors, aggregation in time
- Terminology
 - Time steps vs groups of time steps – what is what?
 - Role of BED in identifying events
- Other options for fusion across sensors?
 - Use simulated events to examine other options
- Conclusions

Examples of Anomalous WQ

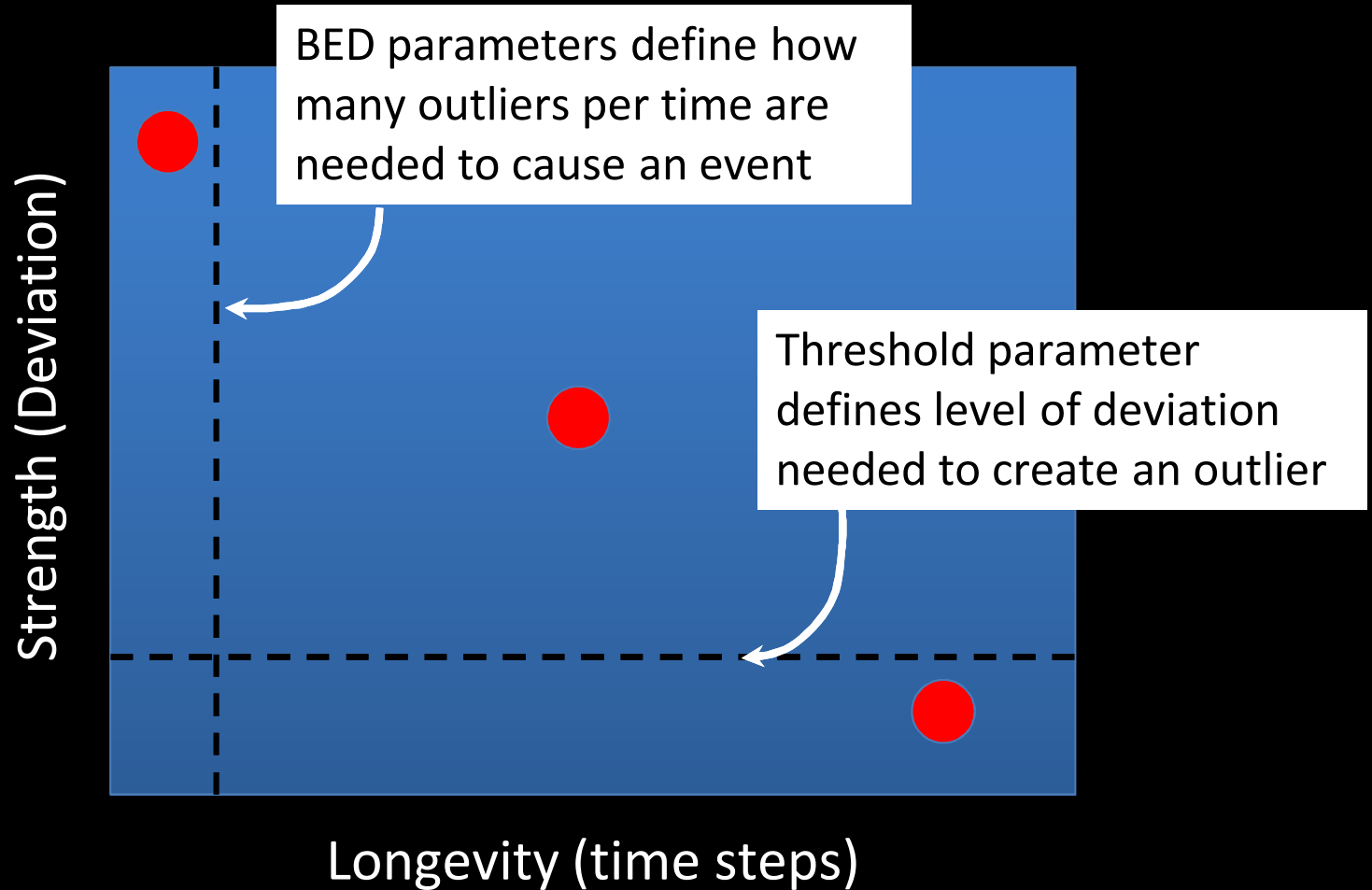
Baseline Change: Sudden, persistent change in mean of water quality signal (often due to operational changes)



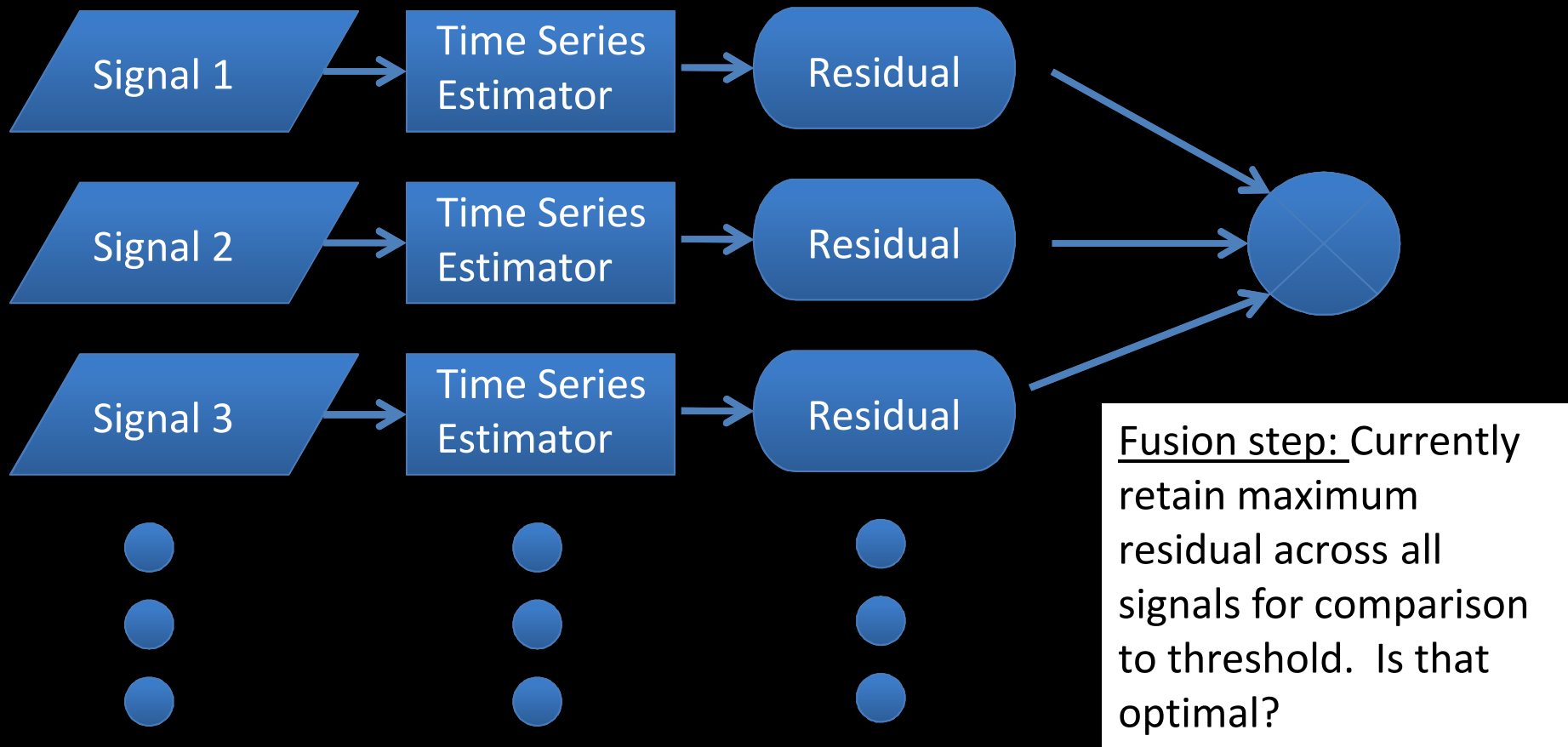
Outlier: significant deviation from background that is not long enough to warrant an alarm

Event: Multiple outliers within a specified period of time

Defining What is Important

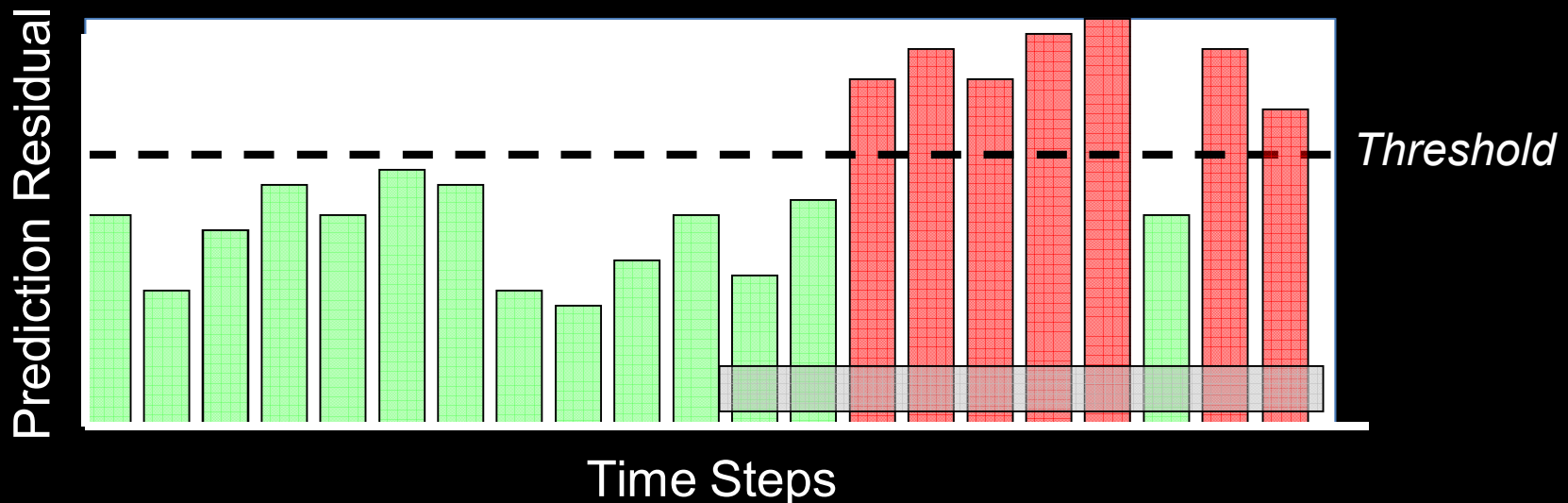


Current Process: Step 1 (cross signals)



Fusion of water quality estimates for the current time step

Current Process: Step 2 (cross time)



A single time step that exceeds the threshold is an “alert timestep” or “*outlier timestep*” (red bars above), but not all outliers cause an event

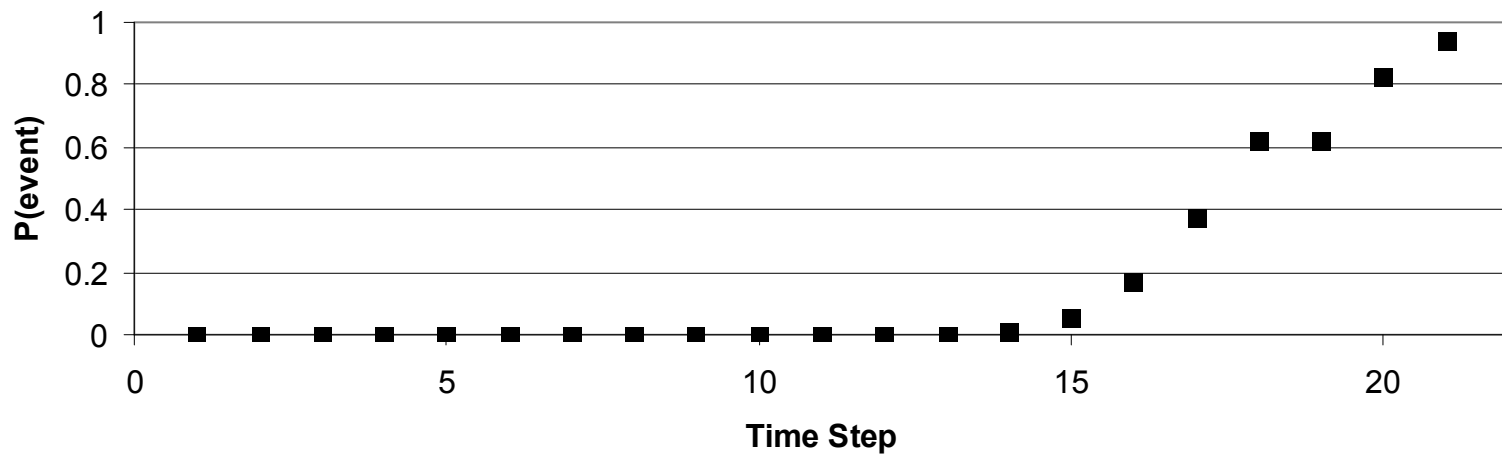
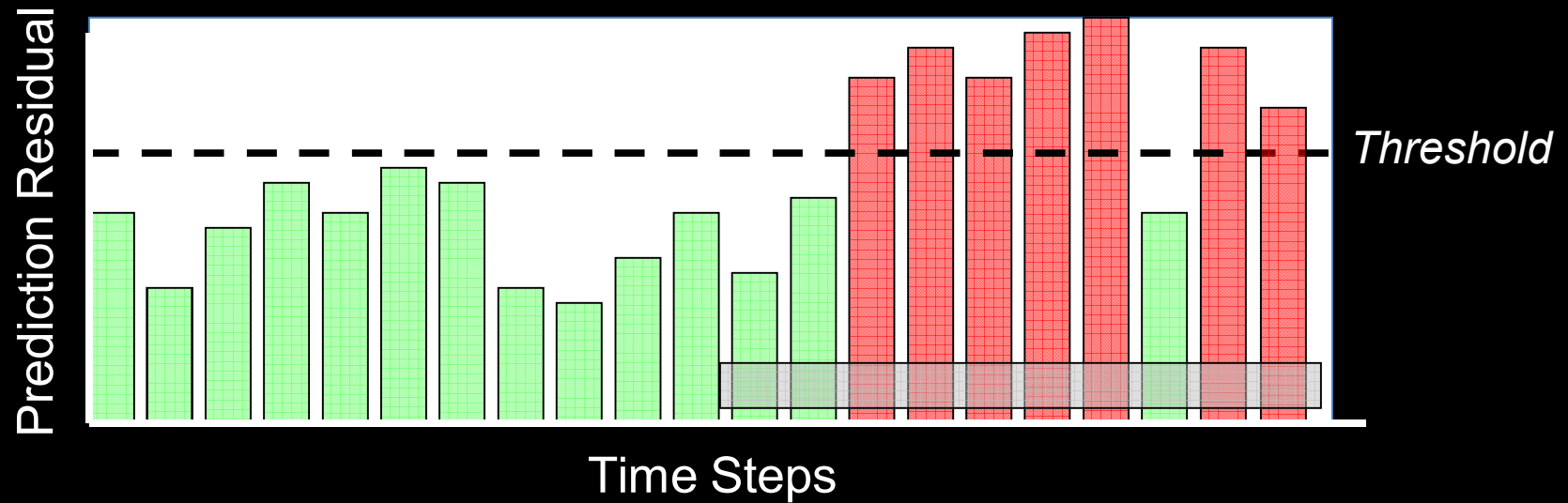
The BED uses a moving window of length *BED Window Size* (10 time steps in example - grey bar) to determine probability that X outliers could occur in *BED Window Size* time steps

Use cumulative distribution function (cdf) of binomial distribution (b)

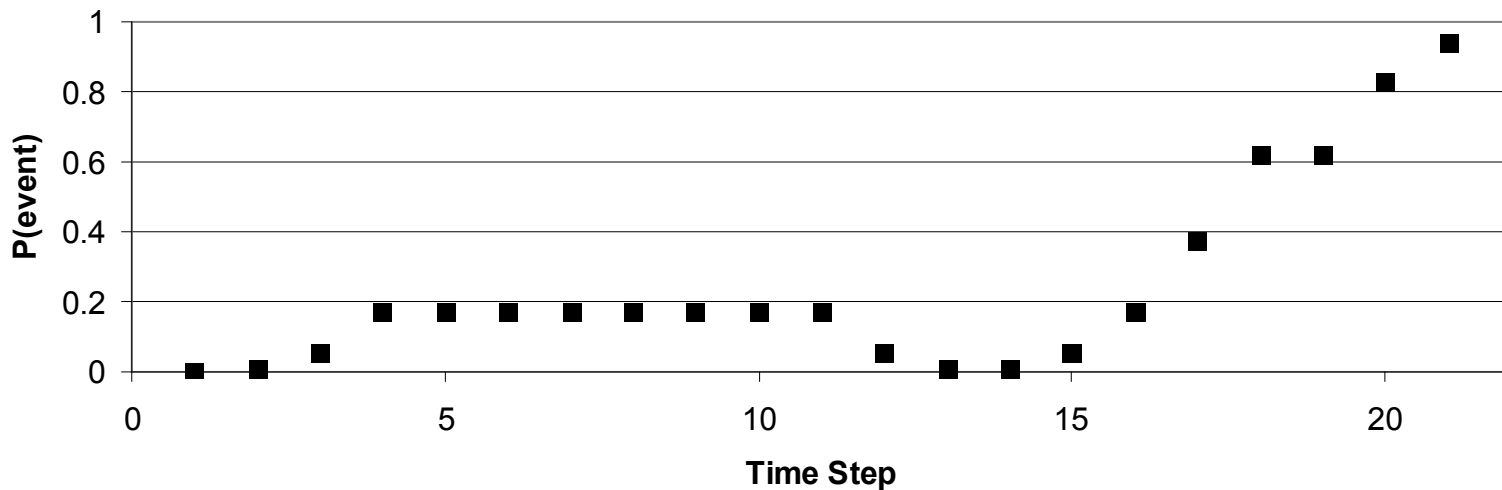
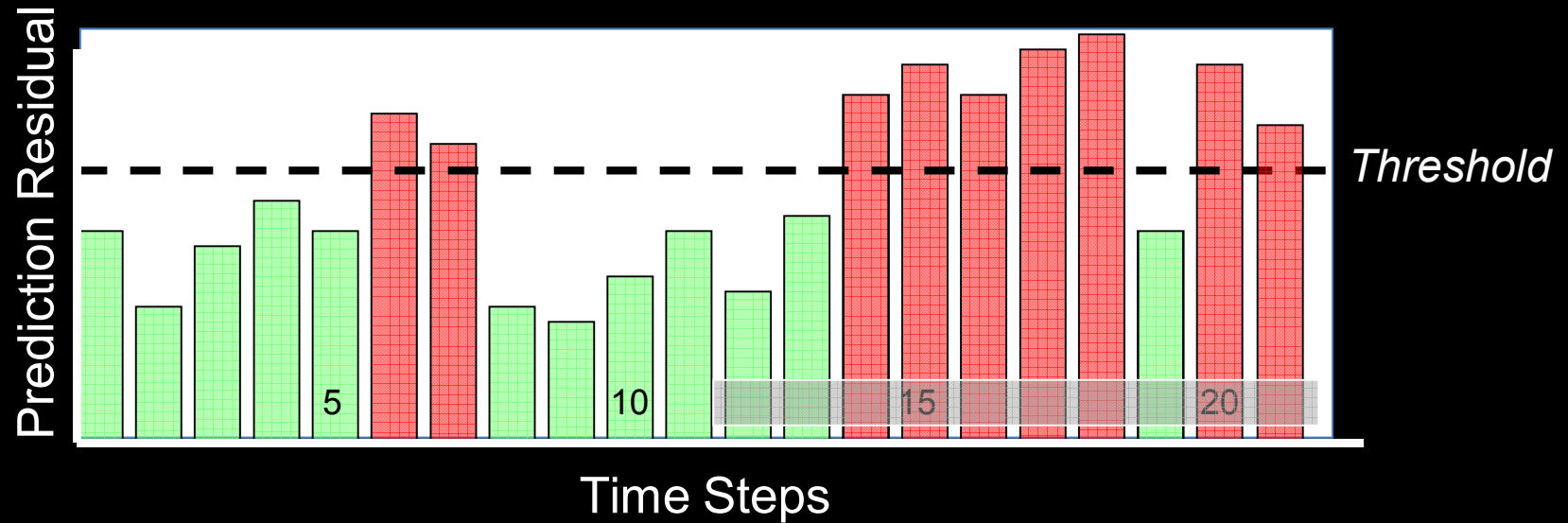
$$P(\text{event}) = b(X; \text{BED Window Size}, \text{Prob. of Outlier}) = b(7; 10, 0.50) = 0.945$$

Compare result (0.945) to *Event Threshold* to determine if an event is occurring
Fusion of water quality estimates over the past N time steps

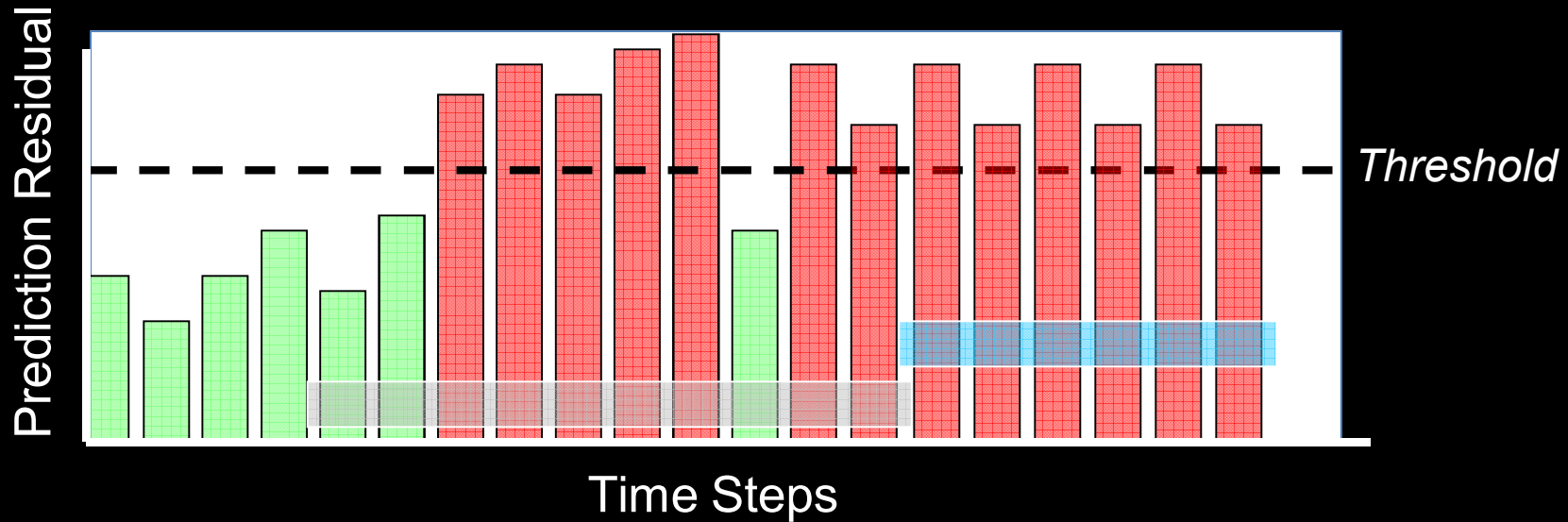
Example Calculation



Another Example Calculation



Definitions: BED

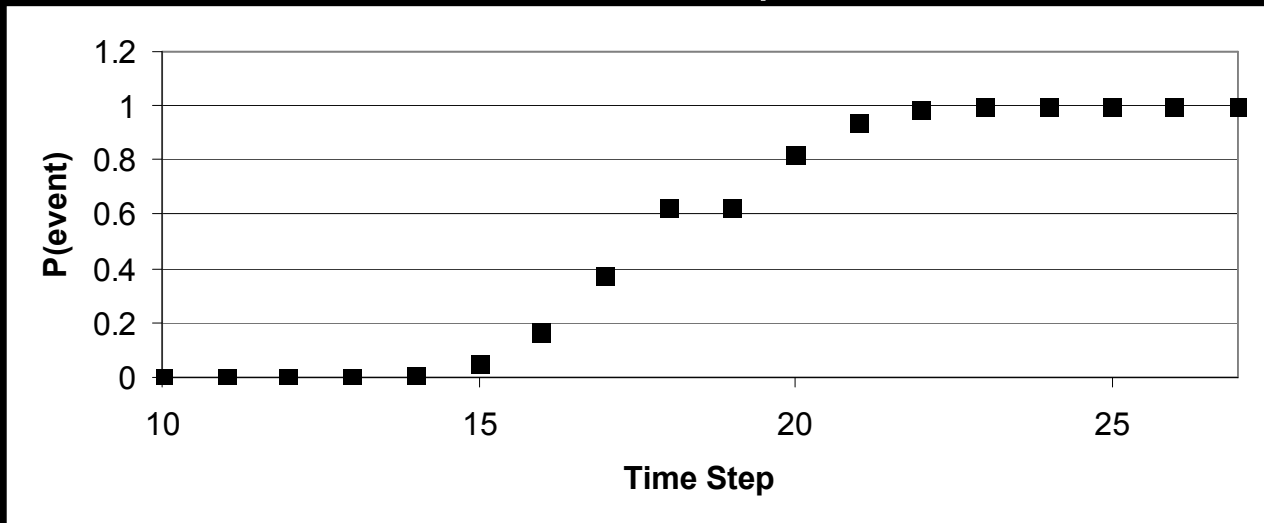
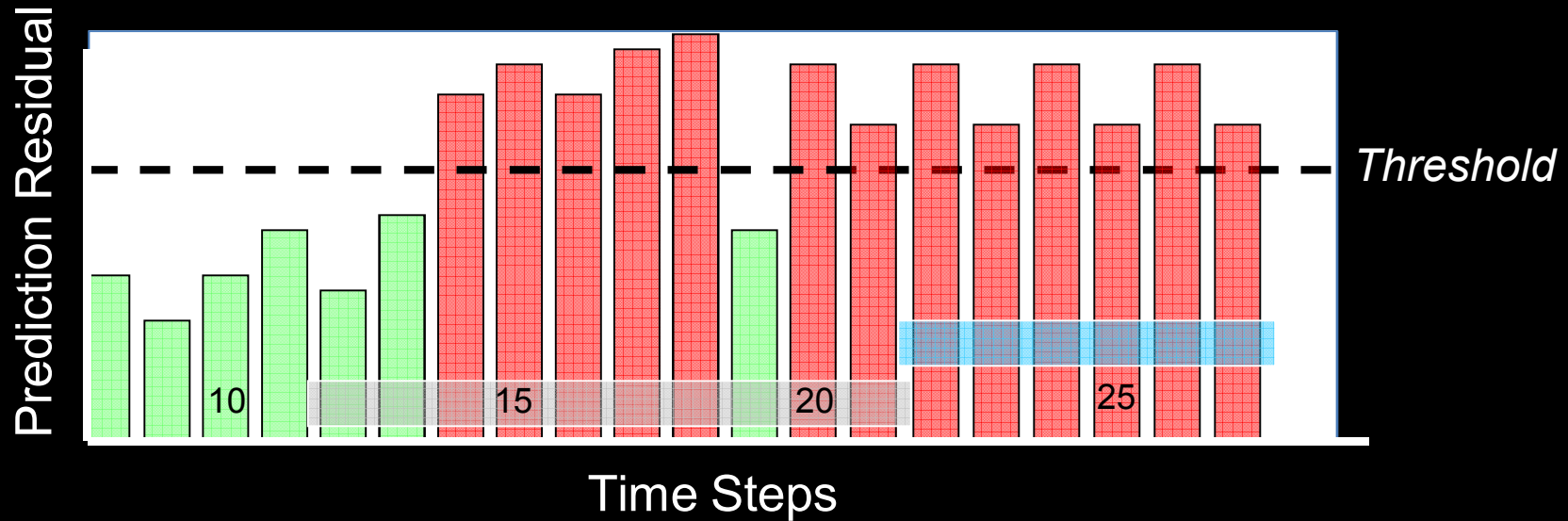


If the previous slide defines an event (i.e., *Event Threshold* < 0.945), next step is to look for a *Baseline Change* (BLC) Note: Grey bar for moving window is in same place as last slide.

If the number of consecutive outliers after an event is \geq *Event Timeout Window*, then CANARY will signal a BLC.

In example above, *Event Timeout Window* = 6 (shown by blue bar), and a BLC is reached at the final time step shown. Alarm is turned off and BED and Timeout windows are reset.

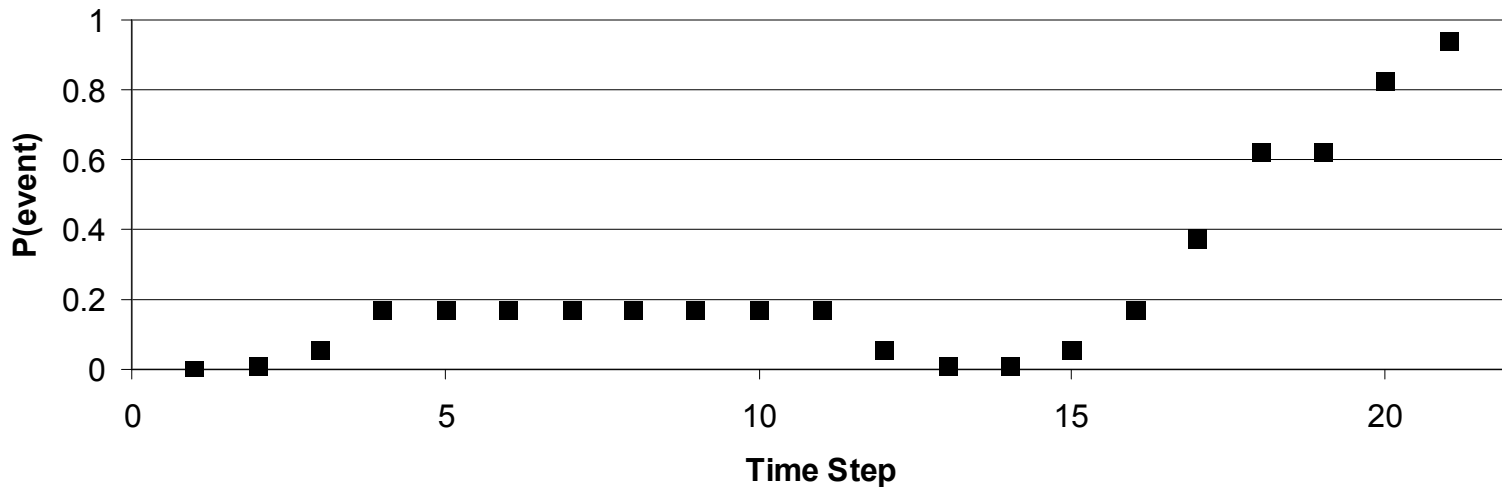
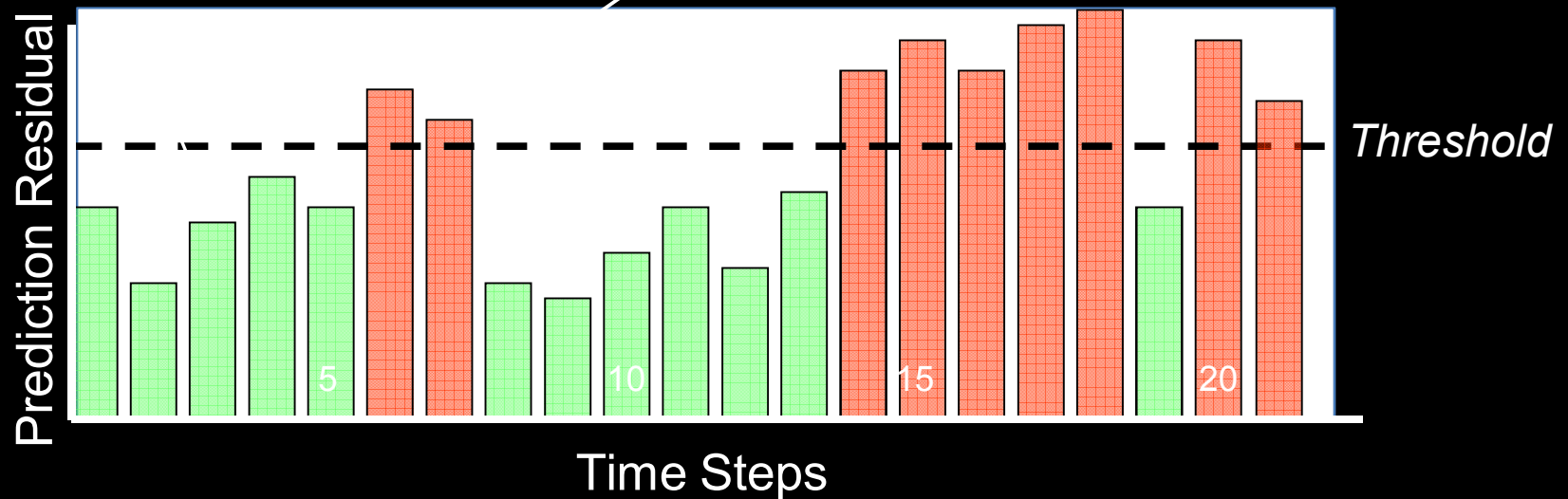
Baseline Change Example



Terminology: Online

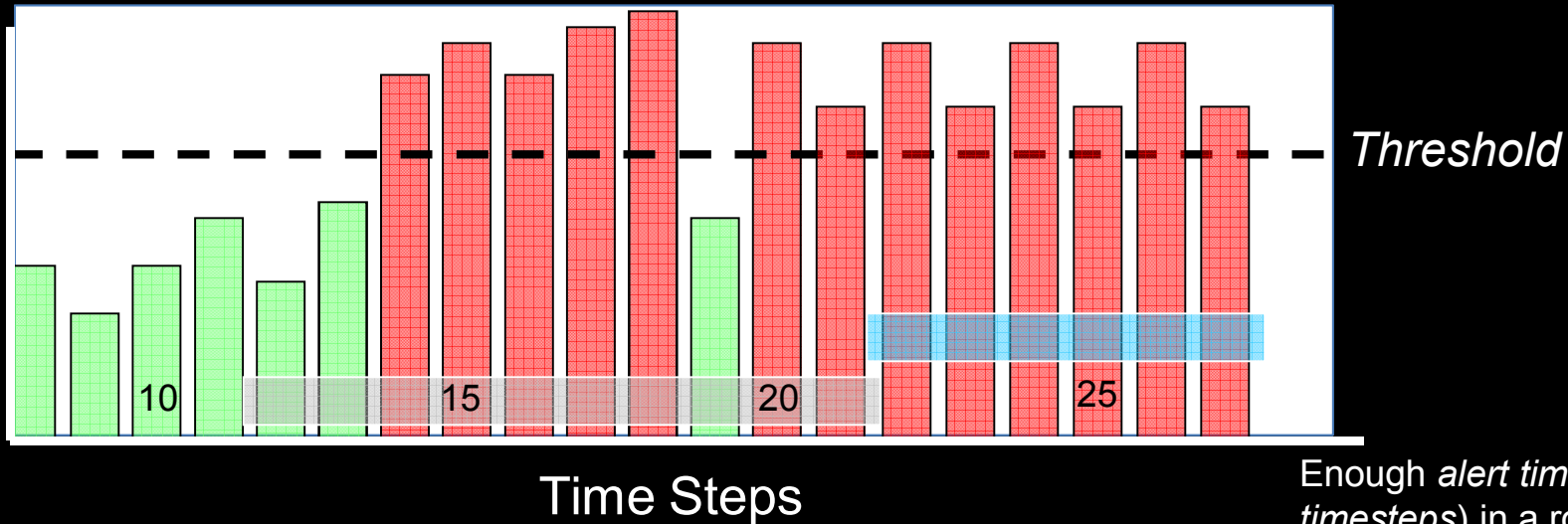
Timestep below the residual threshold =
“normal” or “background” “timestep”

Timestep above the residual threshold =
“alert” or “outlier” “timestep”



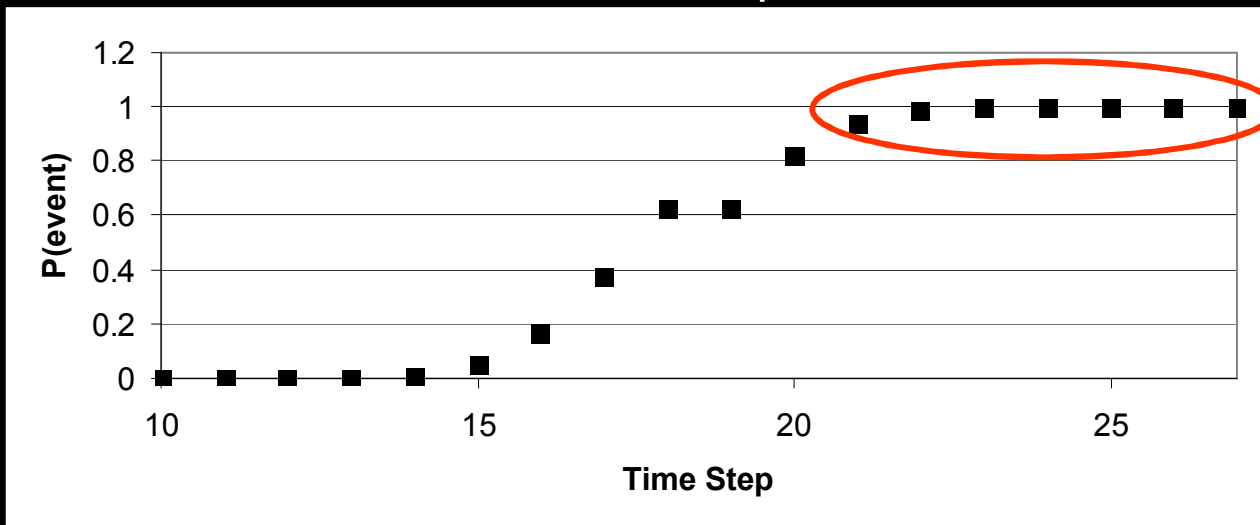
Terminology: Online

Prediction Residual



Threshold

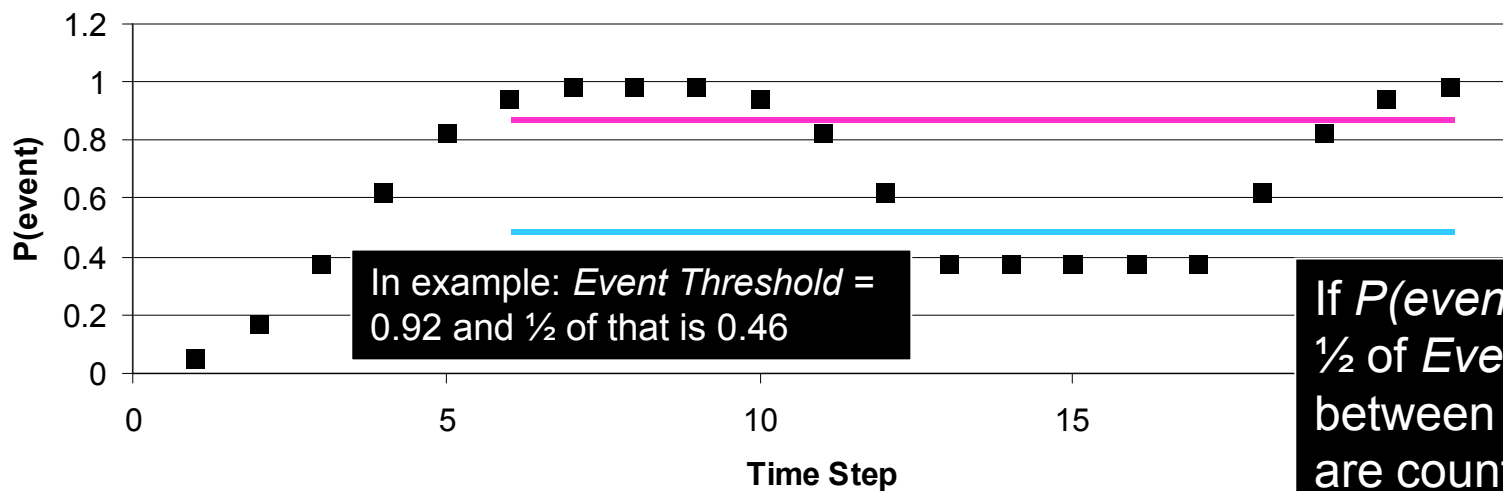
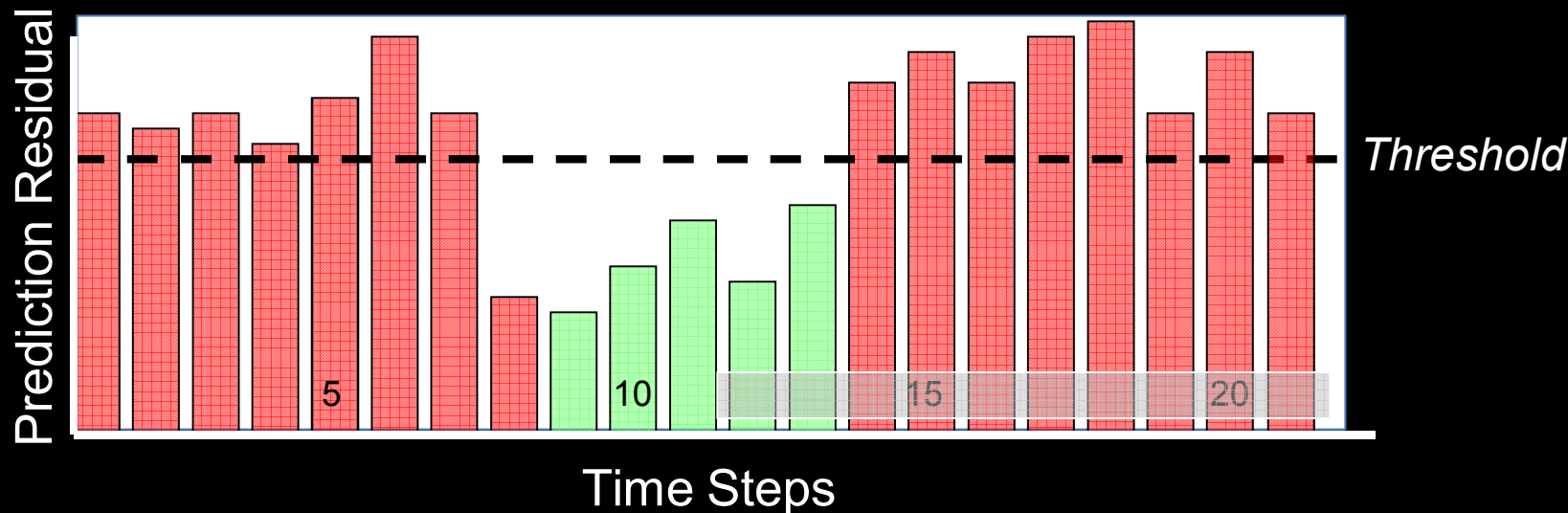
Time Steps



Enough *alert timesteps* (*outlier timesteps*) in a row cause $P(\text{event})$ to exceed the *Event Threshold* (here 0.92) and together, all *alert timesteps* where $P(\text{event})$ exceeds the Event Threshold are called an “*Event*”. An event ends when it reaches a Baseline Change or $P(\text{event})$ drops below the *Event Threshold*.

“*Events*” have previously been called “*Event Clusters*” or “*Event Groups*”

One Event or Two?



In example: *Event Threshold* = 0.92 and $\frac{1}{2}$ of that is 0.46

If $P(\text{event})$ goes below $\frac{1}{2}$ of *Event Threshold*, between events, they are counted as 2 separate events

Possible Options?

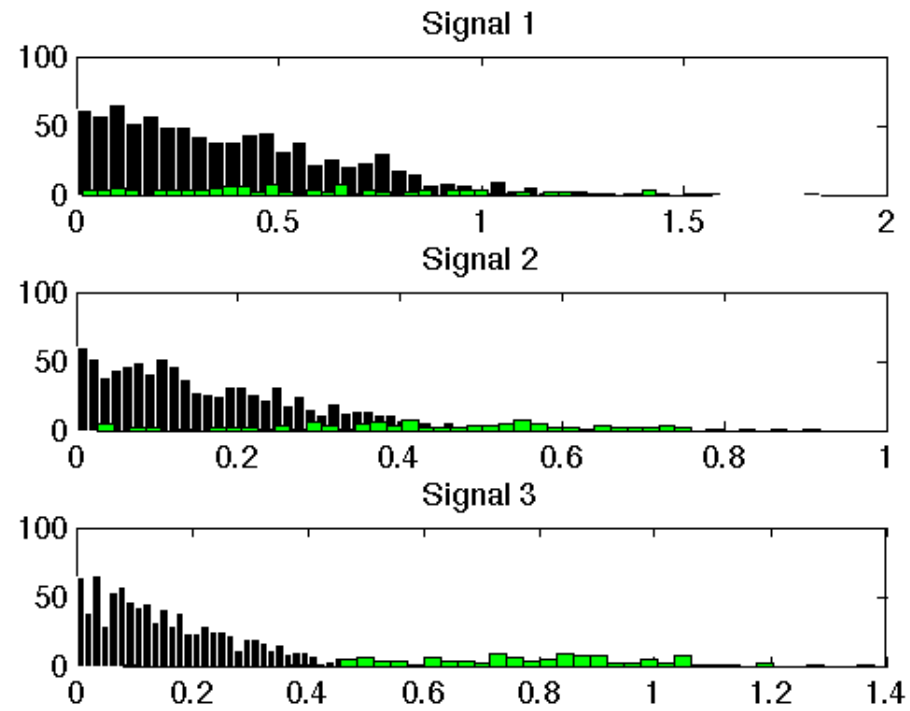
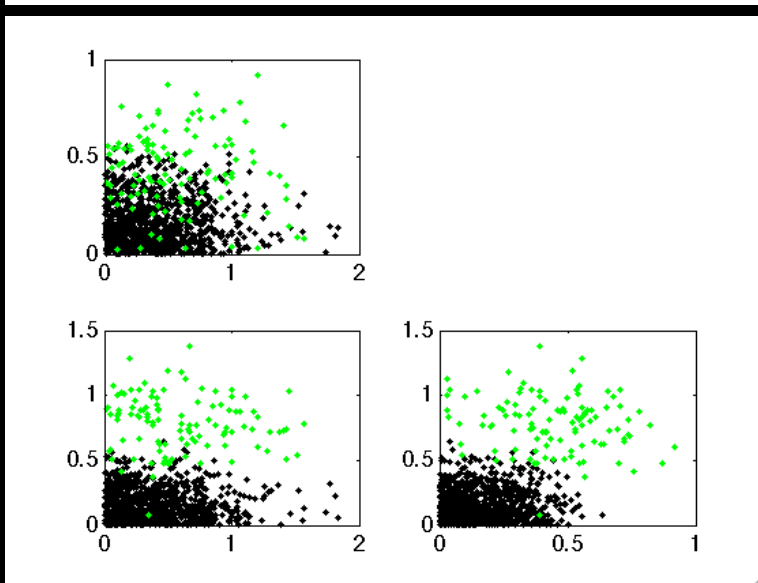
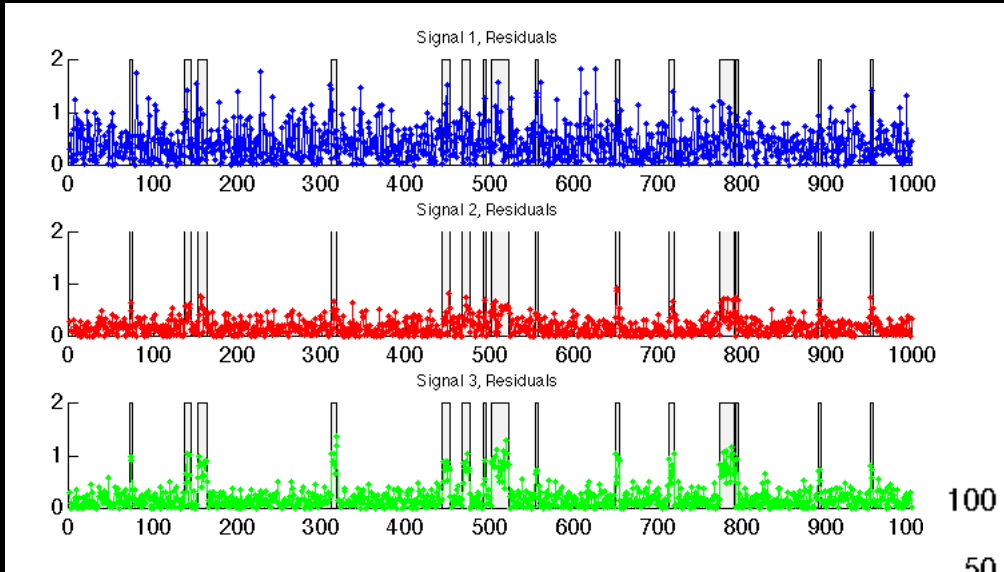
- Is taking the maximum residual at each time across all signals and then aggregating the results across time the best way?
- Possibilities
 - Sum, average, minimum across signals?
 - Keep responses from all signals separate and then for each time step, fuse the $P(\text{event})$ values

Simulation Study

Simulate residuals from multiple signals drawing from both background and event distributions.

Simulate event locations as a Markov process

Use simulated residuals in fusion study

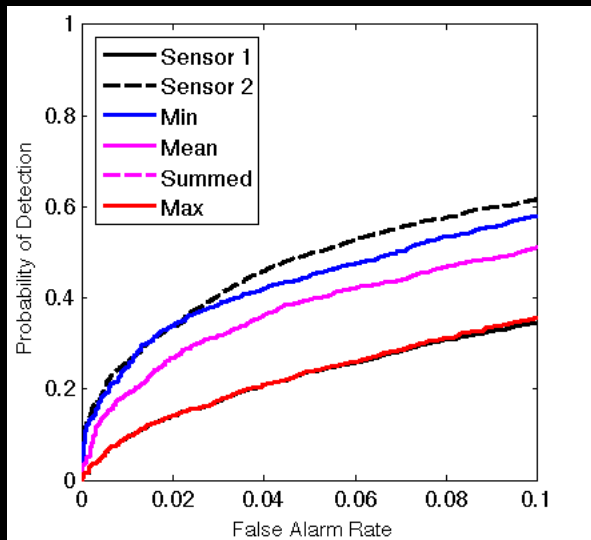


Fusion Options

- Use the residual values themselves
- Linear operations on combinations of residuals (only looking at pairs of signals here)
 - Minimum residual
 - Mean residual
 - Sum of residuals
 - Maximum residual
- Define performance as area under ROC curve for each fusion option and pair of signals

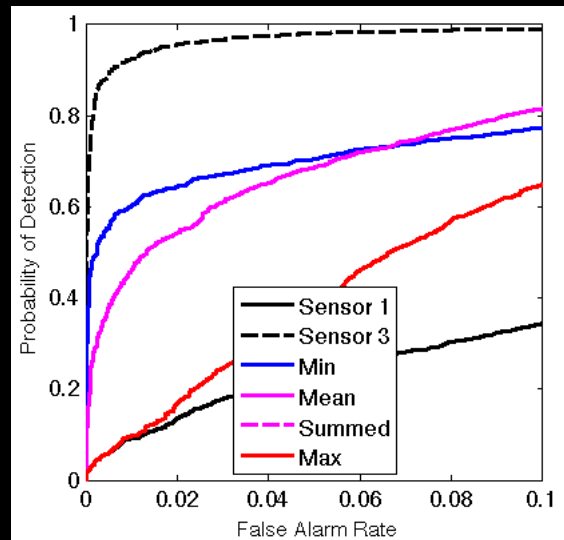
ROC Curve Results

Sensors 1 and 2



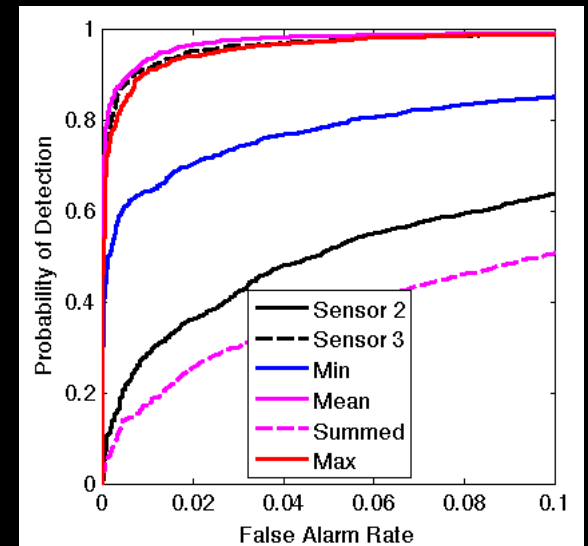
Best Sensor Wins,
Minimum is close

Sensors 1 and 3



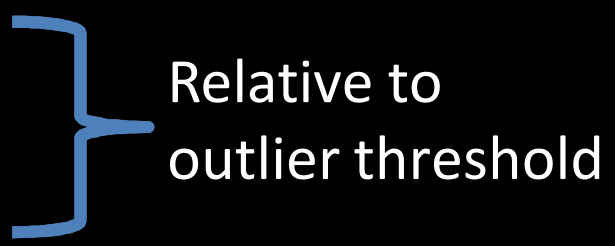
Best Sensor Wins

Sensors 2 and 3



Mean, Max and
Best Sensor are all
similar at the top

Conclusions: Terminology

- Normal/background timestep
 - Alert/outlier timestep
 - Event Threshold
 - Baseline Change
 - Event:
 - all *alert timesteps* where $P(event)$ exceeds the Event Threshold are called an “*Event*”
 - Previously “event groups” or “event clusters”
 - FP, FN, TP, TN – generally reserved for offline analysis (known events)
- 
- Relative to outlier threshold

Conclusions: Fusion

- Maximum performance (ROC Area) is a function of:
 - Information content of signals (differences in background vs. event)
 - Choice of fusion algorithm
 - Not a “one size fits all” solution
- In case of reasonably informative signals, “mean” and “max” fusion options produce best results