

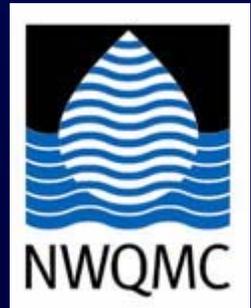
Automated Validation and Grading of Aquatic Time Series Using a Probabilistic Parity Space Method

Peter Hudson, Touraj Farahmand, Ed Quilty

Aquatic Informatics Inc.
Vancouver, Canada

Presented By: Touraj Farahmand

5th National Monitoring Conference, May 7-11, 2006



Aquaticinformatics



Purpose

Introducing an automatic data validation and grading methodology for aquatic time series:

- ✓ *Automatically flag regions of suspect data*
- ✓ *Derive data quality information between site visits and during:*
 - *Extreme events*
 - *Temperature forcing events*
 - *Biological events*
- ✓ *Assign point by point data grades*
- ✓ *Grade the data with any confidence level of your choice*
- ✓ *Do it all in real-time with a fast and efficient algorithm to the Web*



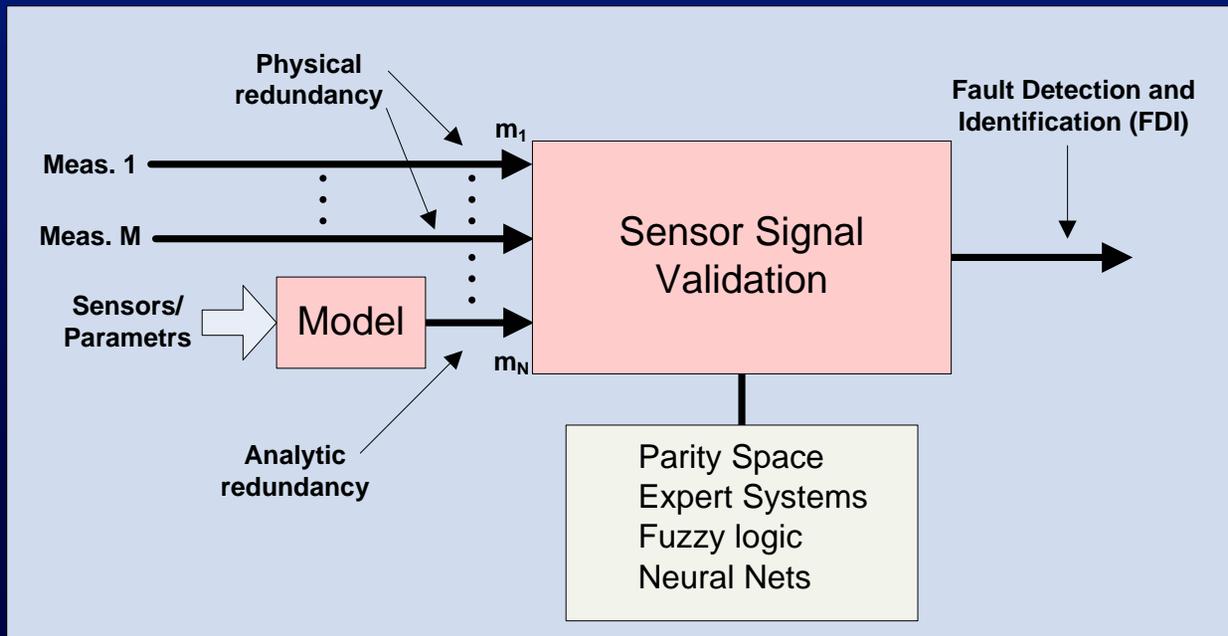
Outline

- 1. Introduction to sensor signal validation*
- 2. Parity Space Fault Detection and Identification (FDI)*
- 3. Parity Space data grading for aquatic data*
- 4. Real world applications*
- 5. Aquarius Time-Series Software: data grading*
- 6. Conclusions and discussions*



1. Introduction to sensor signal validation

- **Sensor signal validation is a method of evaluating the performance of a sensor to provide evidence of the sensor status**
- **Historically, sensor validation is introduced to improve the reliability of complex dynamic systems such as aircrafts, chemical plants and nuclear power plants based on the redundant measurements of the same parameter.**



2. Parity Space Fault Detection & Identification

- **Mathematical model for the redundant measurement systems:**

$$\begin{bmatrix} m_1(t) \\ m_2(t) \\ \vdots \\ m_N(t) \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_N \end{bmatrix} \bullet x(t) + \begin{bmatrix} e_1(t) \\ e_2(t) \\ \vdots \\ e_N(t) \end{bmatrix}$$

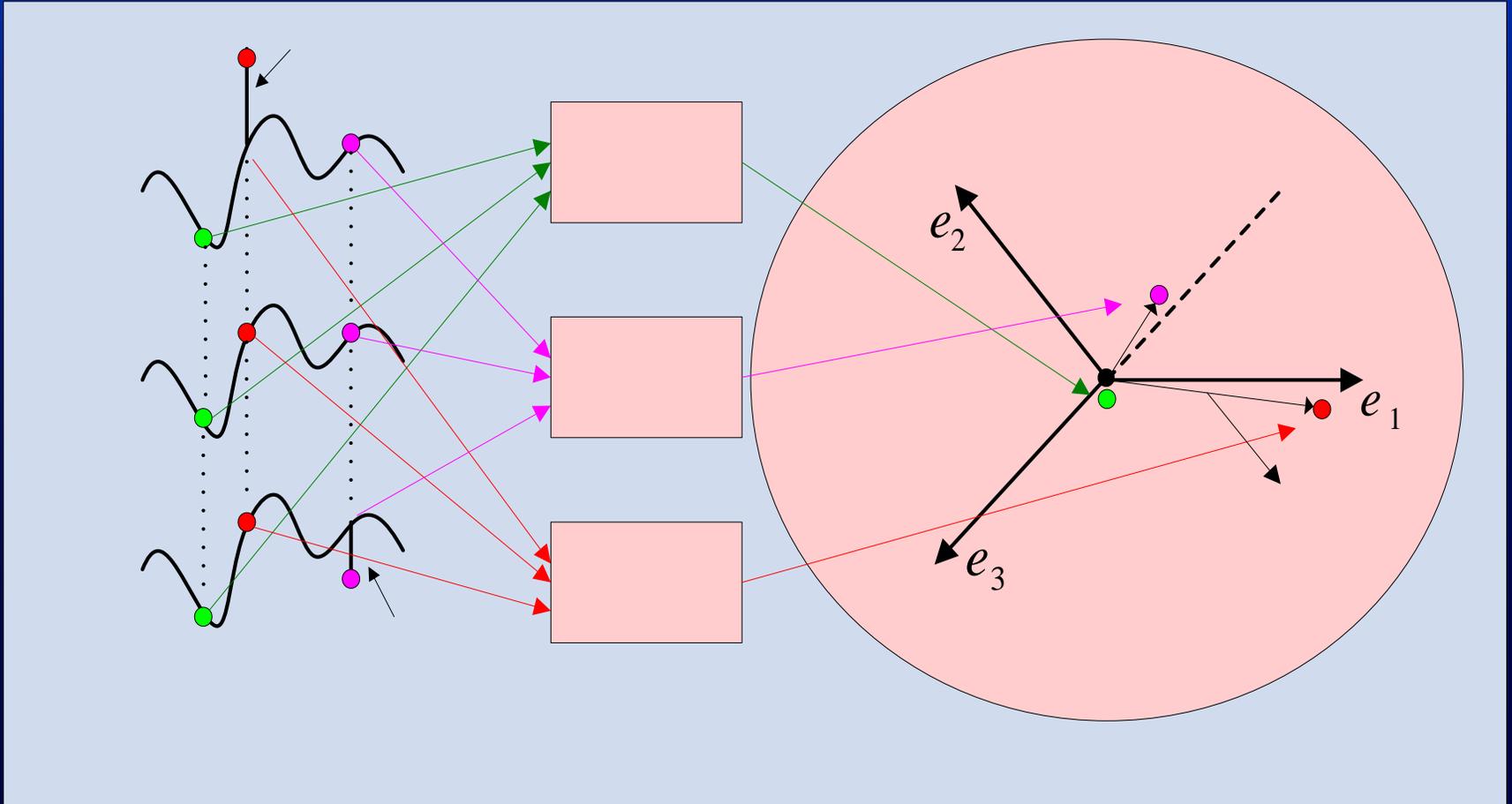
$$m(t) = H \cdot x(t) + e(t)$$

- **The objective is to obtain linear function $f(\cdot)$ such that:**

$$f(m) = f(H \cdot x + e) = f(\underbrace{H \cdot x}_0) + f(e) = f(e)$$

...Parity Space Fault Detection & Identification

- Parity Space mappings for threefold redundancy



3. Parity Space data grading for aquatic data



Example watershed with stage gauges on Tributary A, Tributary B, and the Main Stem, and a nearby meteorological station as indicated.

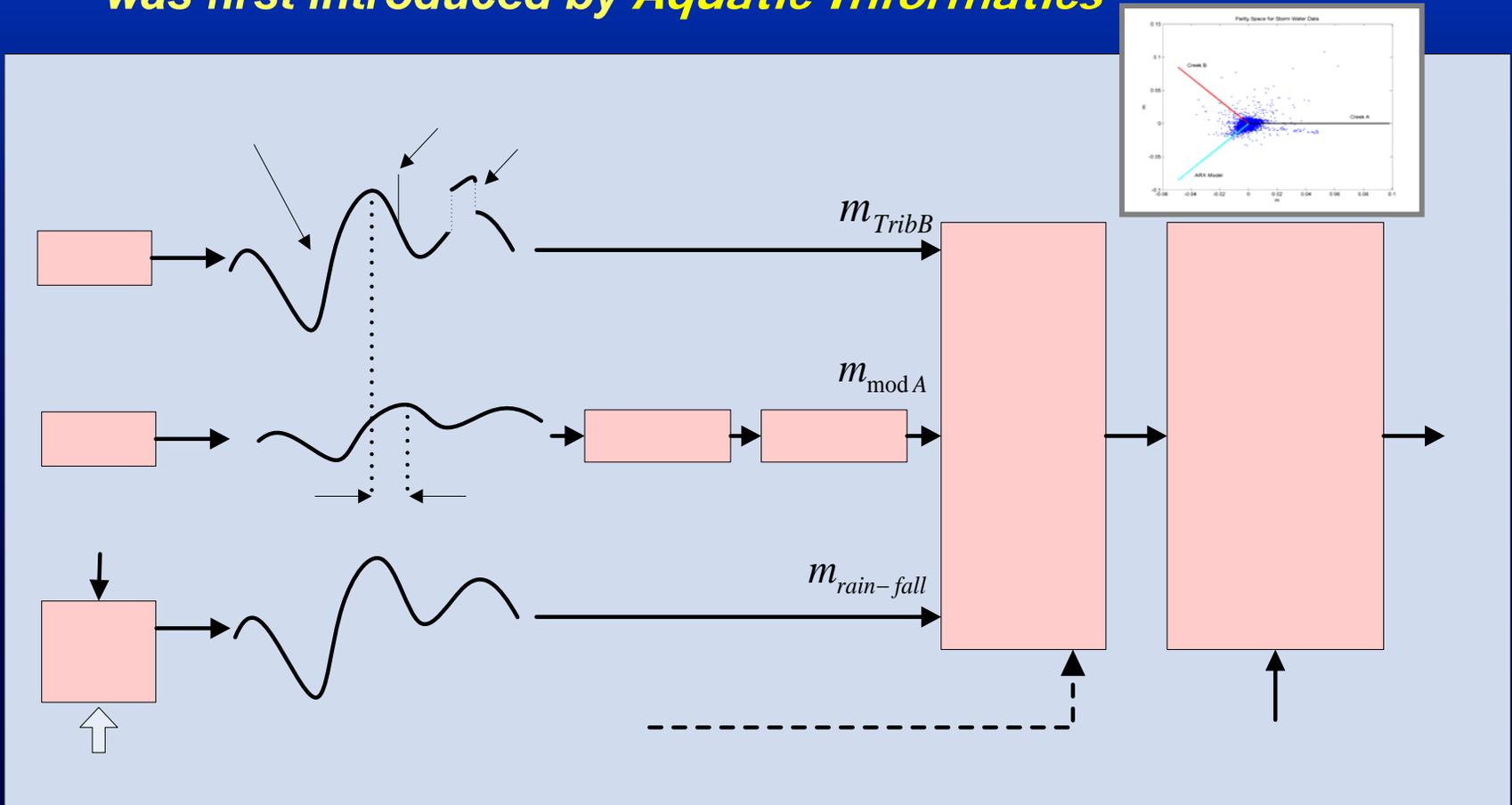


Aquaticinformatics

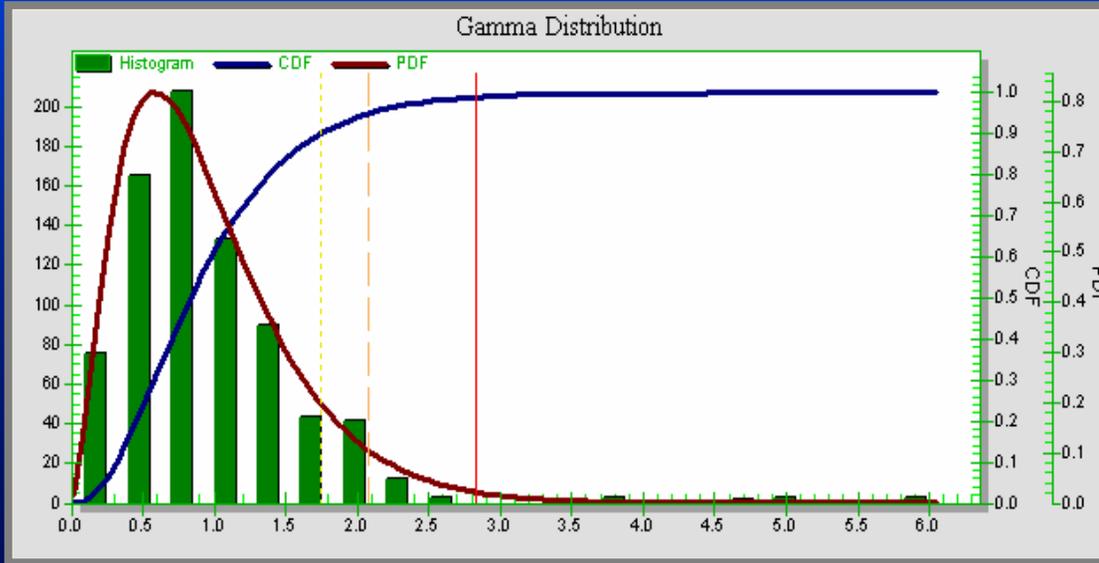


...Parity Space data grading for aquatic data

- Grading aquatic data with probabilistic Parity Space method was first introduced by *Aquatic Informatics*



...Parity Space data grading for aquatic data



Data Grade	Error Confidence (percentile)
Excellent	$p < 80\%$
Good	$80\% \leq p < 90\%$
Fair	$90\% \leq p < 95\%$
Poor	$p \geq 95\%$

Hypothetical data grading system for use with a gamma distribution of parity vector lengths

USGS Site Visit data grading model for water quality sensors (taken from Wagner et al., 2000). Values refer to the discrepancy between insitu readings and site visit measurements.

Measured Physical Property	Rating			
	Excellent	Good	Fair	Poor
Temperature	$\leq \pm 0.2^\circ\text{C}$	$> \pm 0.2$ to 0.5°C	$> \pm 0.5$ to 0.8°C	$> 0.8^\circ\text{C}$
Specific Conductance	$\leq \pm 3\%$	$> \pm 3$ to 10%	$> \pm 10$ to 15%	$> \pm 15\%$
Dissolved Oxygen	$\leq \pm 0.3$ mg/L	$> \pm 0.3$ to 0.5 mg/L	$> \pm 0.5$ to 0.8 mg/L	$> \pm 0.8$ mg/L
pH	$\leq \pm 0.2$ unit	$> \pm 0.2$ to 0.5 unit	$> \pm 0.5$ to 0.8 unit	$> \pm 0.8$ unit
Turbidity	$\leq \pm 5\%$	$> \pm 5$ to 10%	$> \pm 10$ to 15%	$> \pm 15\%$



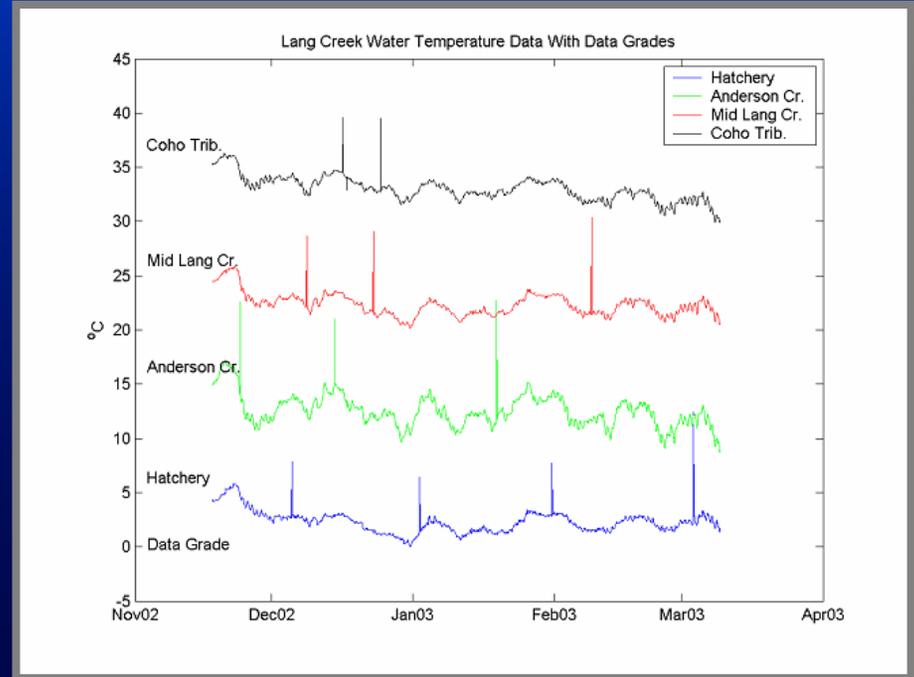
Aquaticinformatics



4. Real world Applications



➤ Lang Creek watershed



Winter 2002/2003 temperatures logger data from four stations in the Lang Creek watershed.

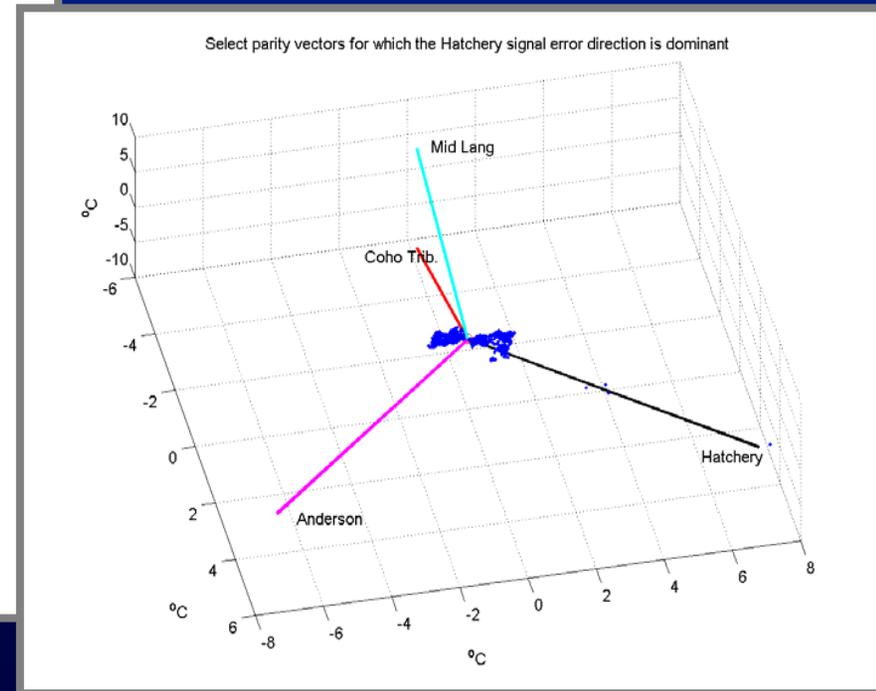
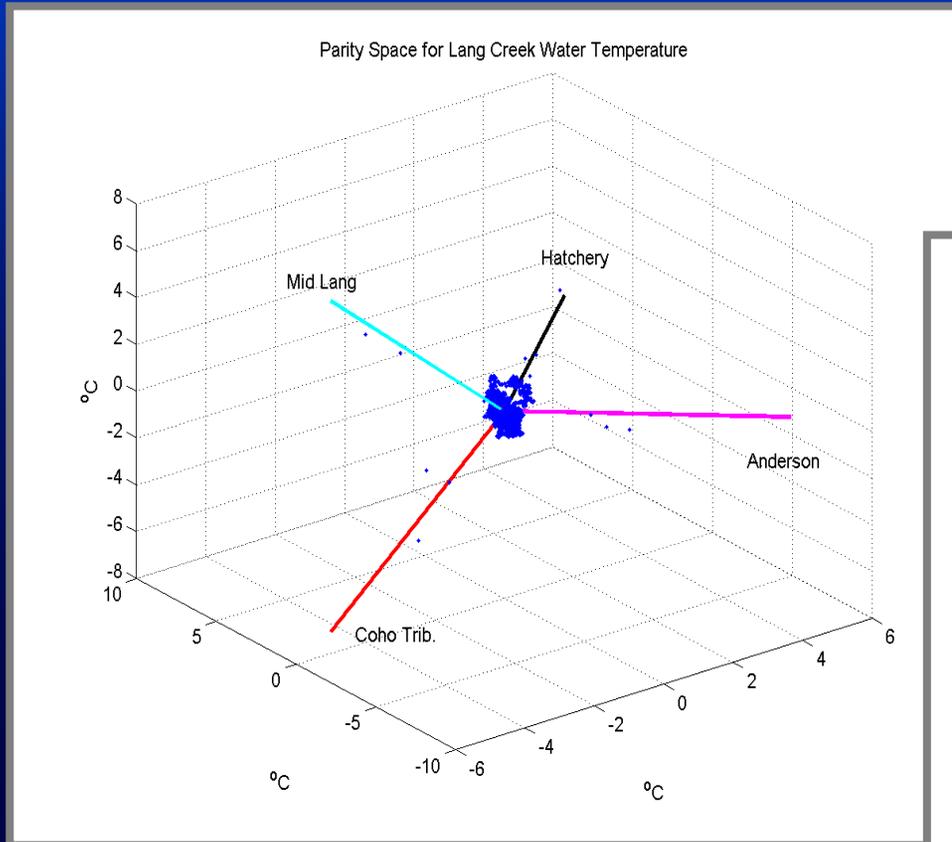


Aquaticinformatics



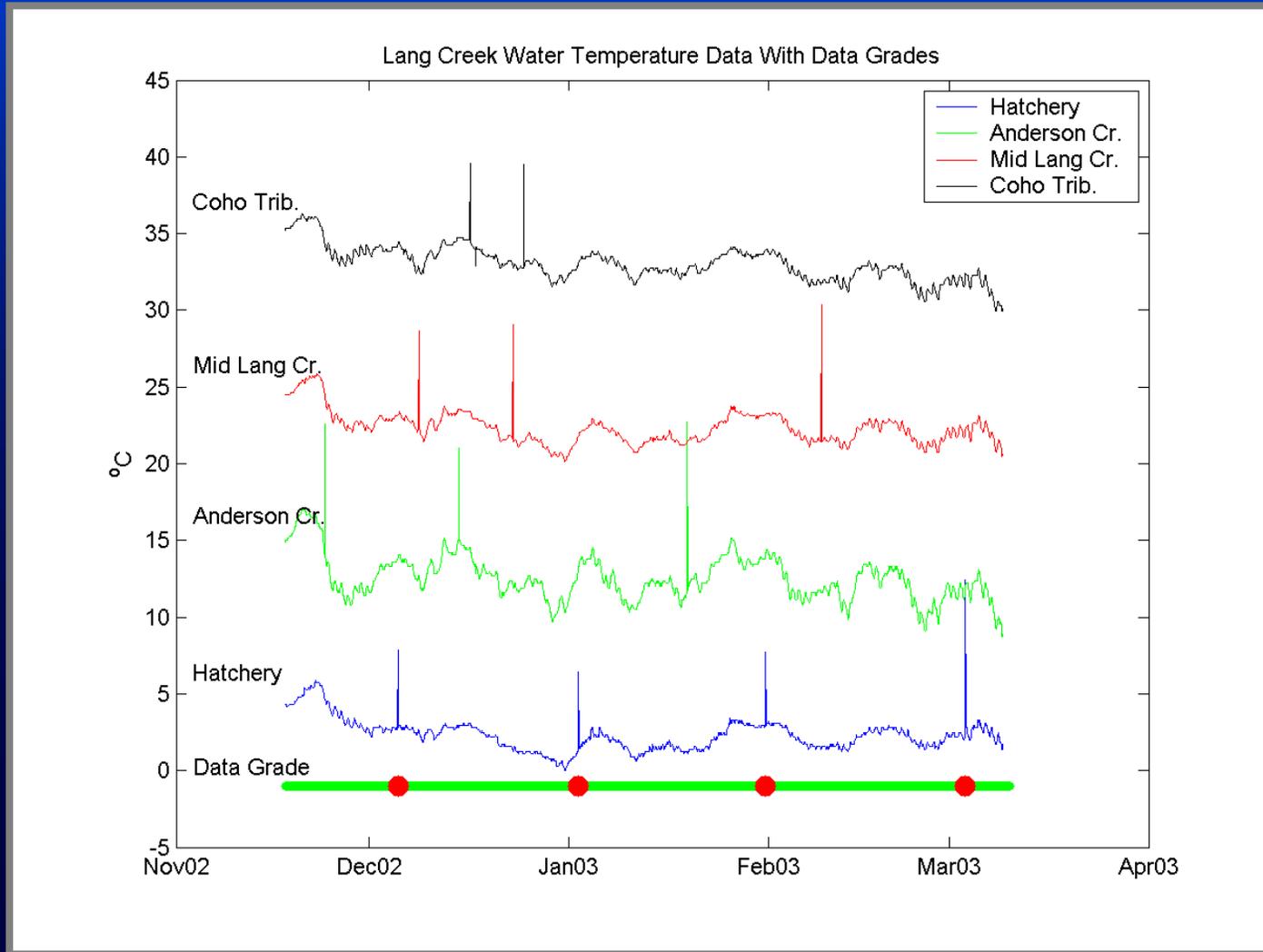
...Real world Applications

➤ Lang Creek watershed - Parity Space plots



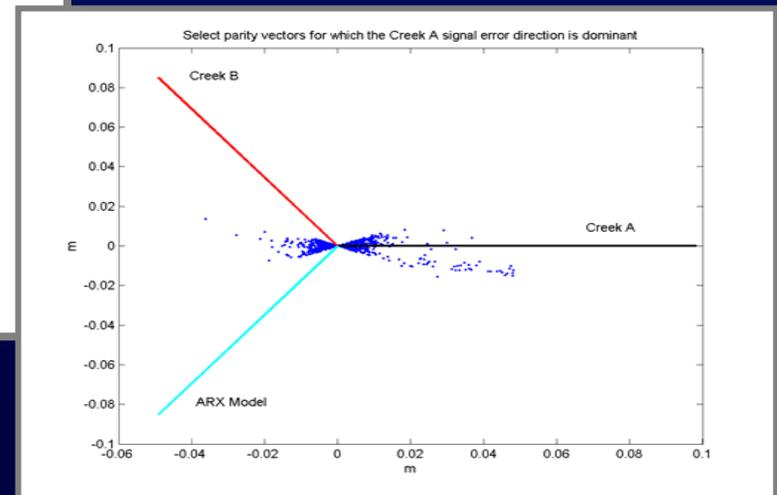
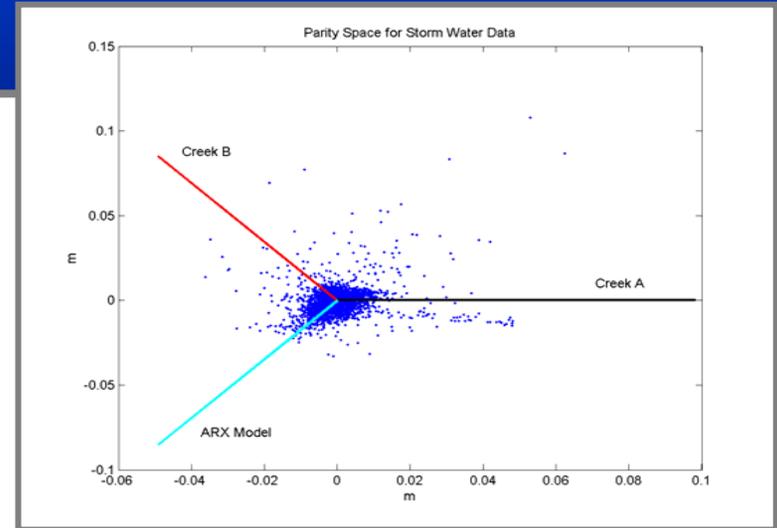
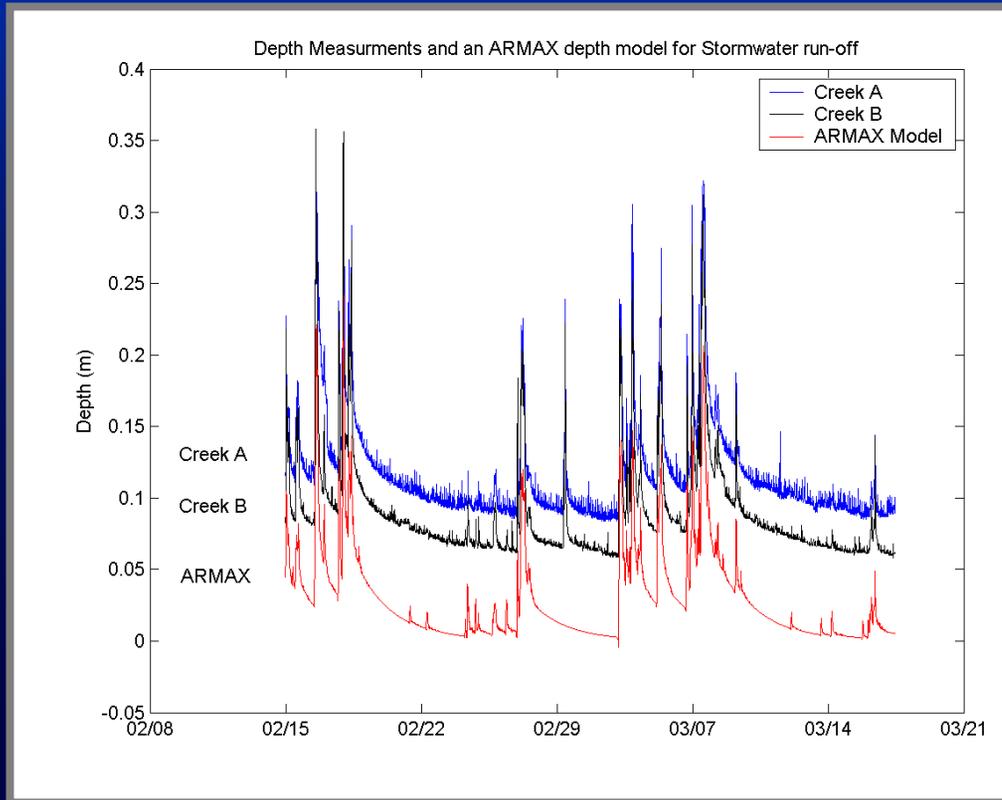
...Real world Applications

➤ Lang Creek watershed – Data grading results



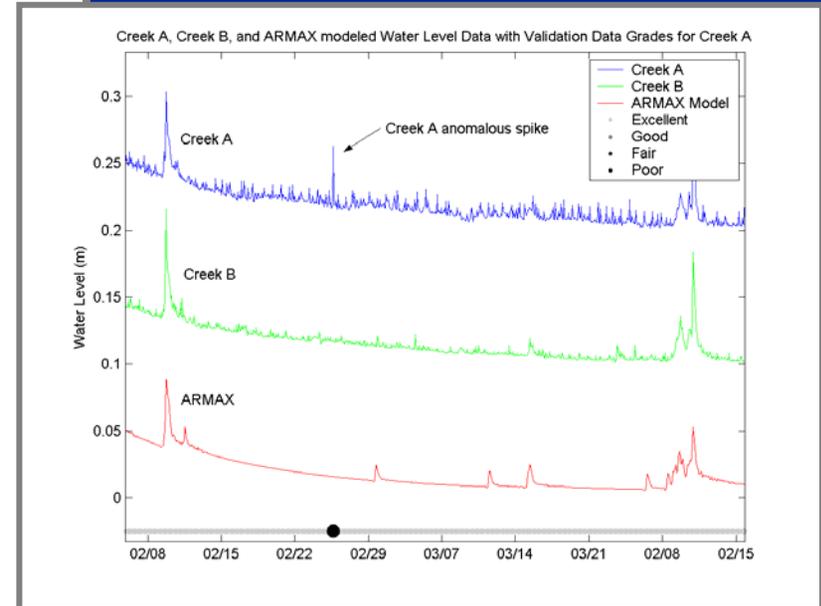
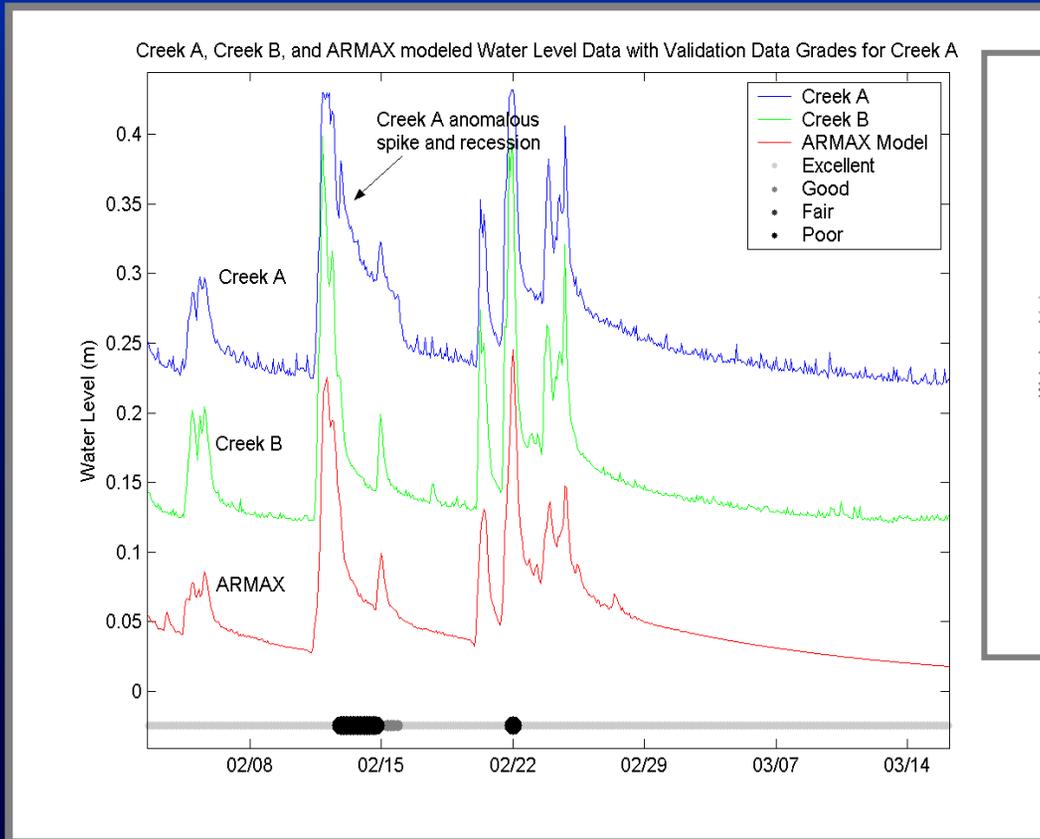
...Real world Applications

➤ Storm Water Runoff in an Urban Setting in Southwestern British Columbia



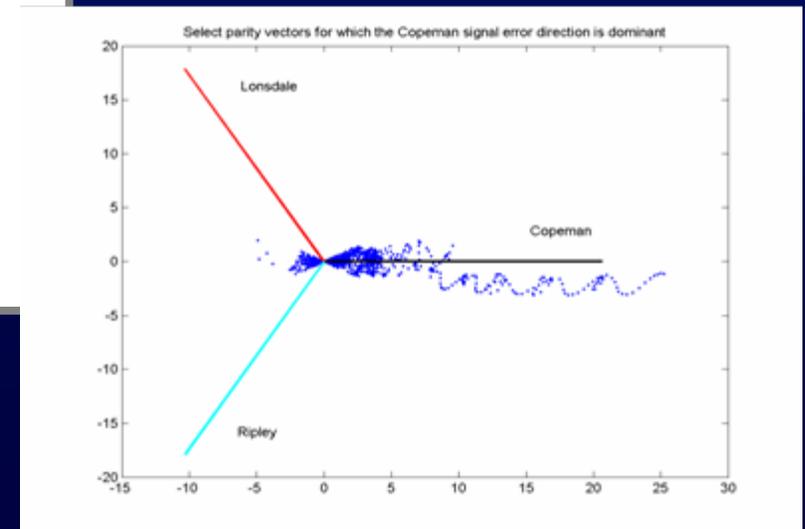
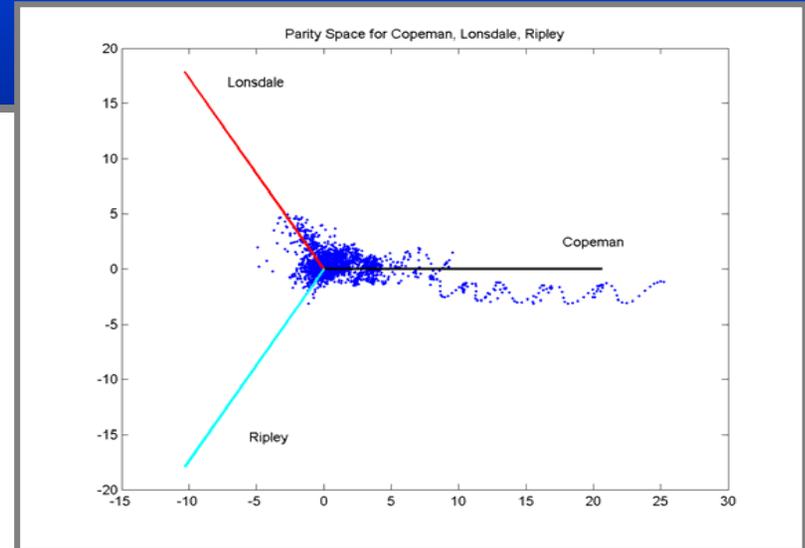
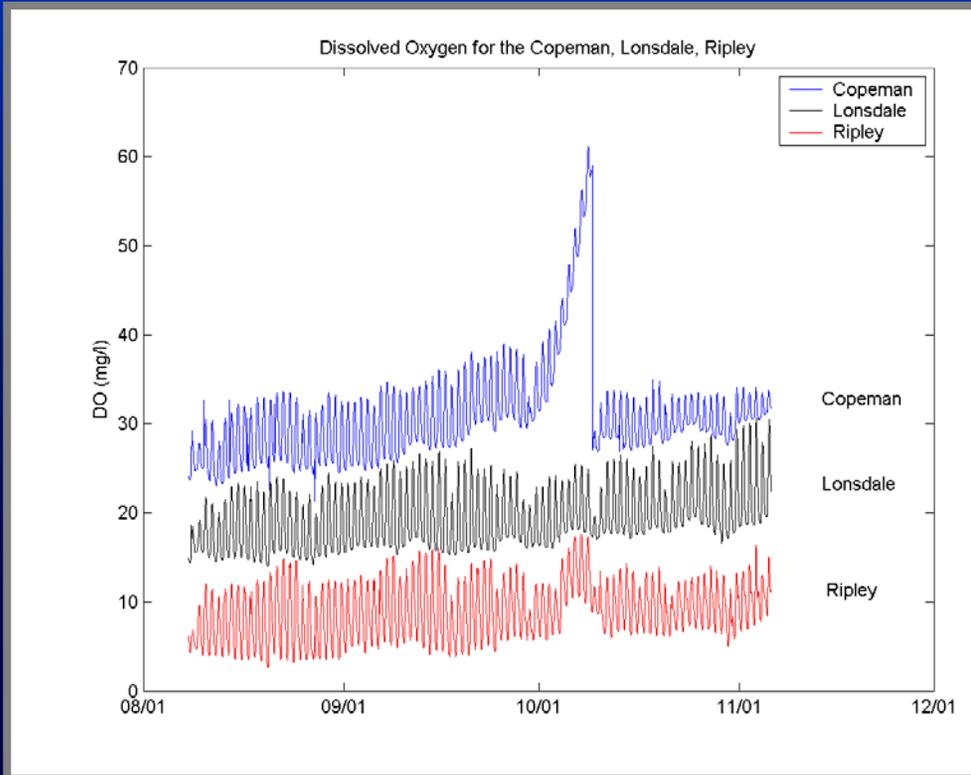
...Real world Applications

➤ Storm Water Runoff in an Urban Setting in Southwestern British Columbia – Data grading results



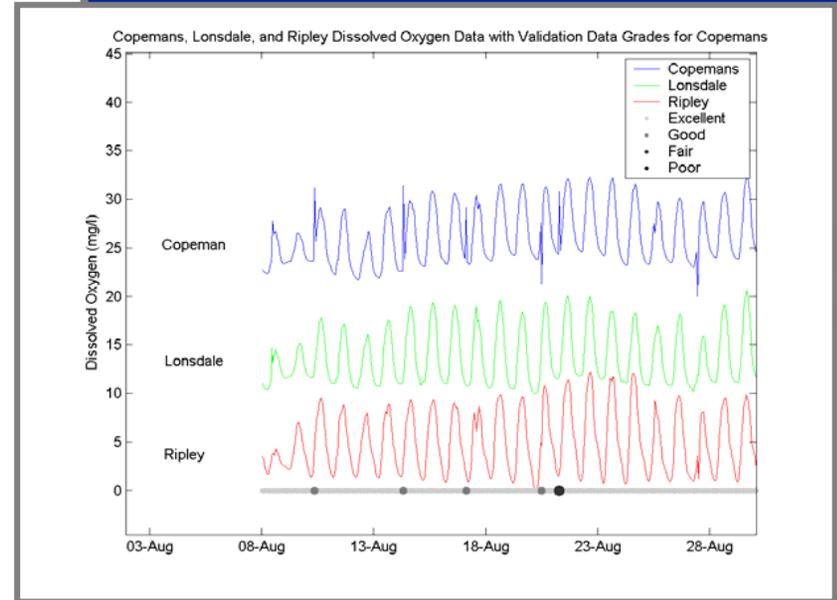
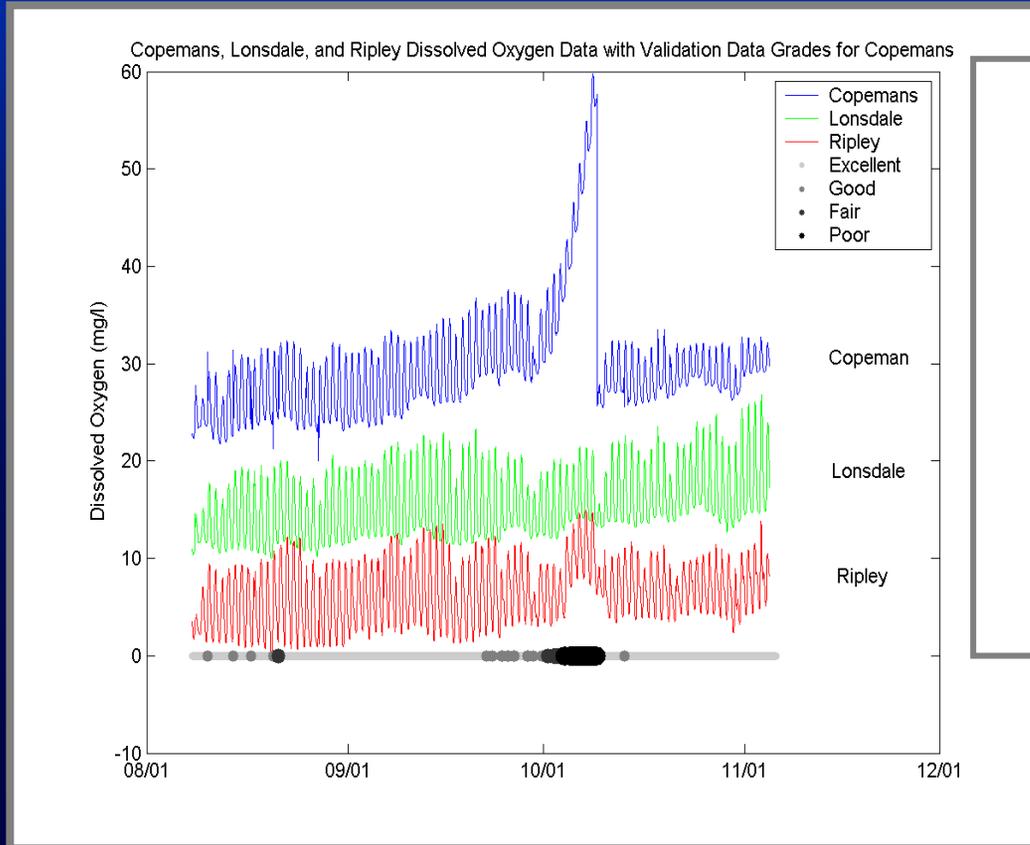
...Real world Applications

➤ Dissolved Oxygen Sensor Validation from a River System in Southern Ontario

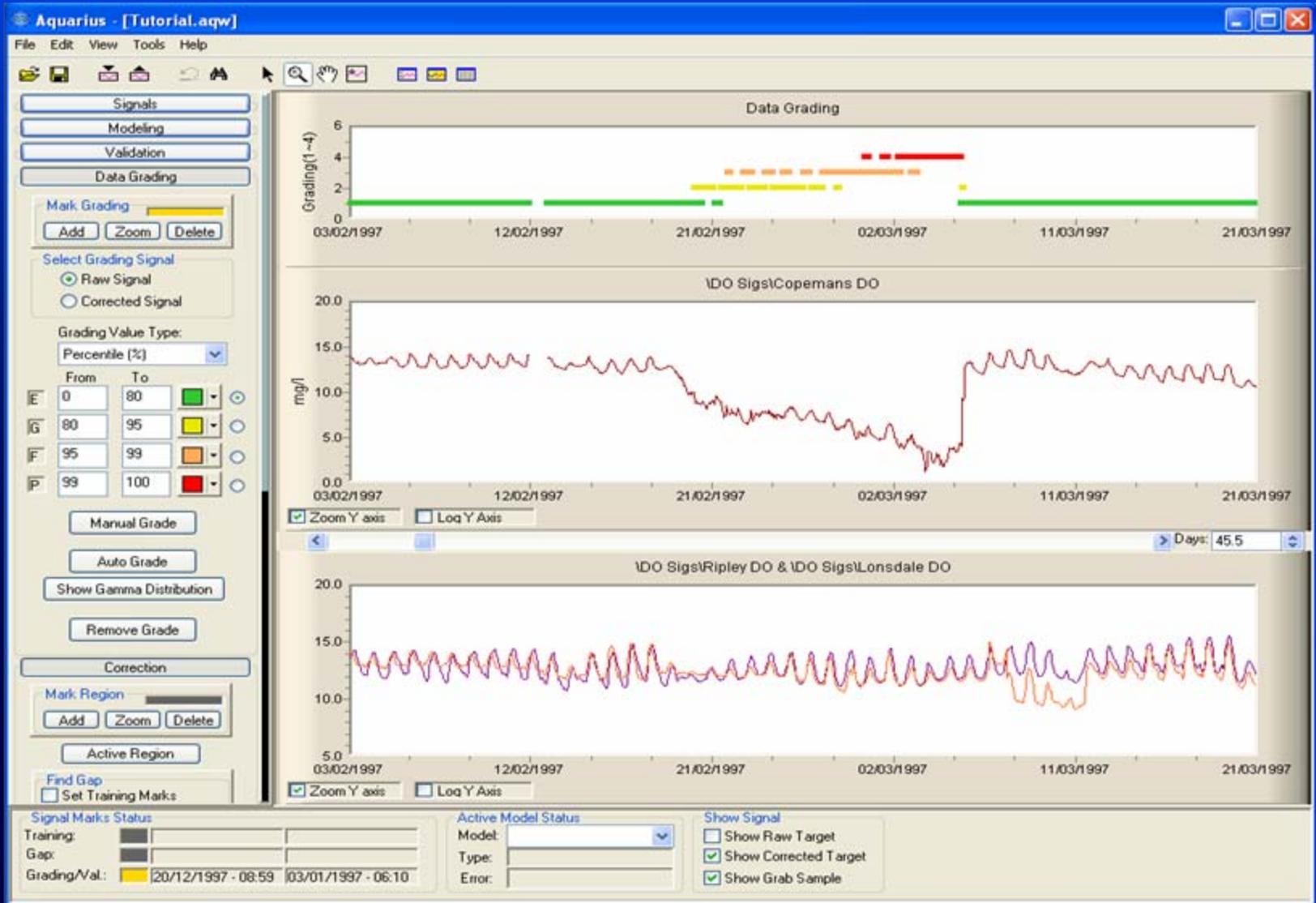


...Real world Applications

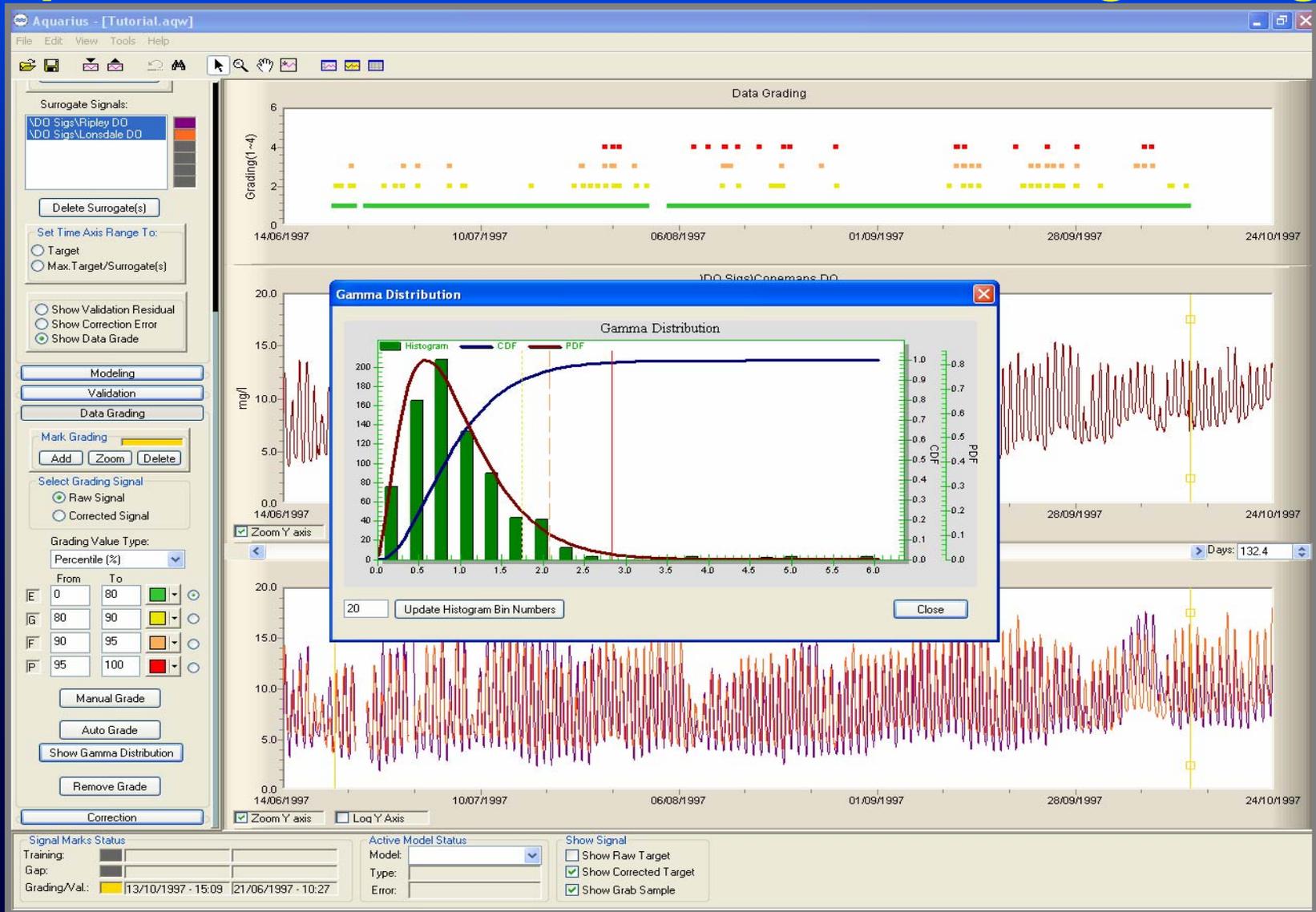
➤ Dissolved Oxygen Sensor Validation from a River System in Southern Ontario – Data grading results



5. Aquarius Time-Series Software: data grading



...Aquarius Time-Series Software: data grading



Aquaticinformatics



6. Conclusions and Discussions

The method of probabilistic parity space data validation offers water scientists and data managers another arrow in their quiver:

- ✓ ***A tool to quickly highlight particular regions of (often vast) data series that must be further examined for quality control***
- ✓ ***Point by point data grades can be assigned to a data series, in contrast with the traditional method of applying bulk grades to sections of data between site visits***
- ✓ ***Data grading can be based on an independent set of intuitive percentile grades, rather than complex parameter-specific grades***
- ✓ ***Method is also applicable to marine, atmospheric or any other environmental times series for which redundancy (physical or analytic) can be established***
- ✓ ***Fast and efficient for automatic real-time data grading***



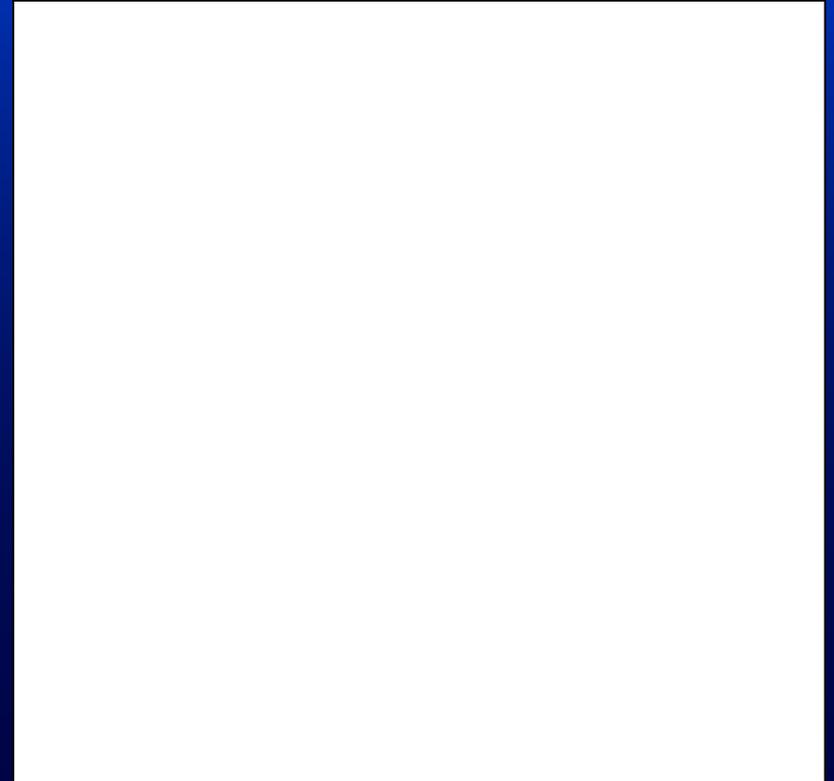
QUESTIONS?

Touraj Farahmand, B.Sc. M.Sc.

(604) 873-2782

tourajf@aquaticinformatics.com

www.aquaticinformatics.com



...THANKS!



Aquaticinformatics

