

PQLX

A Data Quality Control Solution

Orfeus Workshop

Lisbon, Portugal

27 May, 2011

- What is PQLX and What Can it Do?
- What is a PDF?
- Server-Side Functionality
- Client-Side Functionality
- Near- and Far- Future Functionality
- Conclusion
- Availability

What Can PQLX Do?

- Identify the **Existence** of Data Quality Problems
- Identify the **Source** of These Problems
- View Data in the **Time Domain**
- View Data in the **Frequency Domain**
- Switch Between the Two
- **Transform** Data in the Time Domain
- Create PNG-Format Plots For Web Display

What Can PQLX Not Do?

- Fix Your Data Problems!

But First....

- What is a PDF in the PQLX World?
 - Algorithm Conceived by D.E. McNamara, USGS - NEIC
 - PSD Calculation Parameters Following Peterson Method to Match High- and Low-Noise Models
 - A Bunch of PSD's Represented in a Single Plot

Example Problems Identified:

- Meta-Data Issues:
 - Response File Definition - via PDF
 - Advertised Orientation - via Particle Motion Plot

Problems Identified:

- Instrumentation Issues:
 - Dead Channel - via PDF
 - Channel-Specific Sensor Problems - via UVW Transformation (STS2 & Trillium Instruments)
 - Too Frequent Mass Re-Centers / Calibration Events / Re-Boots - via PDF
 - Gaps & Overlaps - via PDF and Time Domain Data Views

Problems Identified:

- Site Response Analysis:

- Noise Levels Across all Frequencies - via PDF

- Minimum Noise Base-Line, per Channel - via PDF

- Identify Existence of Human Noise Sources, for example:

- Nearby Mechanicals, e.g., Water Pump

- Nearby Transportation - Trains, Cars, etc.

- Ships - Causing Seiches/Sloshing Water (see Panama Canal)

- Identify Existence of Natural Noise Sources, for example:

- Weather

- Ocean Waves

- Earthquakes

- Calving Glaciers

Problems Identified:

- Telemetry Issues:
 - Data Drop-Outs - via PDF and Time Domain Data Views

Server-Side Functionality

- Auto-Identify ALL Relevant Files - Traces and Responses
- Read & Analyze Files
 - Database Auto-Configured, No User Set-up Required
 - Save Meta-Data and Analysis Results to Database
- Compute Statistics (per Channel):
 - Sample Rate Epochs
 - System PDF's
- PNG-Format Output of (per Channel):
 - System PDF's
 - Spectrograms

Client-Side Programs

- Two Graphical Programs:
 - Database Administration Program
 - pqlx-admin
 - Data Viewer Program
 - pqlx

Administration Program - pqlx-admin

- Database Management:
 - Creation
 - Modification / Update
 - Deletion
- Database Meta-Data Inspection:
 - Data Directory Statistics
 - Channel Information
- Server Execution
 - Set Execution Configuration Parameters
 - Execute From Within GUI
 - View Log / Error Message Output

Data Viewer Program - pqlx

- Time Domain Data Views:
 - By Trace File
 - Unlimited Number of Files
 - Sortable on Multiple Header Values
 - View Entire Trace, Magnified Portion and Spectra Simultaneously
 - By Station/Channel
 - Display up to 60 Days Per Channel
 - Select Data for Further Analysis / Transformation
 - Plot Predicted Arrivals for Events

Data Viewer Program - pqlx

- Frequency Domain Data Views (i.e., PSD's)
 - Probability Density Function - PDF
 - Server-Generated - Pre-Defined Time Periods
 - User Request
 - Interactive - Select a Sub-Section
 - Movie
 - Spectrograms
 - Pre-Defined Time Periods - Matching System PDF's
 - User Request
 - Interactive - Zoomable, Dynamic Color Bar

Data Viewer Program - pqlx

- Available Data Transformations:
 - High- and Low-Pass Filtering
 - De-Mean
 - De-Trend
 - Polarity Reversal
 - Differentiate
 - Integrate
 - ENZ to UVW - Trillium & STS2
 - Spectra
 - De-Convolution
 - Particle Motion

Near-Future Functionality

- Base-Line Comparisons

- Based on Single- and Double-Line Curves (period, power value pairs)

- Types:

- Pre-Defined, e.g., HNM & LNM

- User-Defined, Imported via Admin GUI

- Relative to Channel, e.g., 90% Percentile

- Identifying Out-Liers:

- Single Curve Line - Points Falling Above or Below

- Double Curve Lines - Points Falling Outside Both

- Identifying In-Liers:

- Single Curve - Points Matching dB Values Exactly

- Double Curve - Points Falling Inside Both Curves

- Displayable as PDE - Including All PDE Functionality

Not-So-Near-Future Functionality

- Network-Wide PDF's - Grouped by Like Channel
- Automatic Alerts Based on Base-Line Comparison Results
- Analyses Performed on Real-Time Data Transmissions
- Define Base-Lines vs. Known Signatures, for example:
 - Portion of PDF Defining Microseism Events
 - Portion of PDF Identifying Calibration Pulses
 - etc...
- Allowing For Automatic Identification of All Known Source Signatures!

Conclusion

- PQLX Performs a Wide Array of Functions:
 - Data Quality Assessment – Trace and Meta
 - Station Quality Assessment – Site and Instrument
 - Research – Only Some Imagination Required
 - Database of PSD's - Available for Query / Extract
 - Data Mining
 - Pattern Recognition – Known and Yet Unidentified

Sponsorship

The Following Organisations Have Significantly Contributed to PQLX Development:

- IRIS - DMC, PASSCAL, NSF
- USGS - NEIC, ASL
- NanoMetrics Inc., Canada

Availability

PQLX Can Be Obtained From:

Nanometrics Inc.

Peter Devanney
Sales Manager, Europe and Africa

250 Herzberg Rd.

Kanata, ON

Canada, K2K2A1

T: +1 613 592 6776 ext. 301

F: +1 613 592 5929

peterdevanney@nanometrics.ca