Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

Event Detection and the Importance of Feature Selection

NASPI – Session 5
April 2018

BRETT AMIDAN
JIM FOLLUM
NICK BETZSOLD
Pacific Northwest National Laboratory

TIANZHIXI (TIM) YIN
University of Wyoming

May 8, 2018

Why Feature Extraction?

- **Feature Extraction** – Any algorithm that transforms raw data into features that are used as input into analytical algorithms

Raw PMU Data
1 month, 26 PMUs, 60 Hz = **200 GB**

- Too much data for traditional algorithms when looking across many months
- Analyses will take a long time to complete and/or require extensive computing resources
- **Data, in its raw form, may not even help you identify the type of data behavior that you are interested in identifying and understanding**

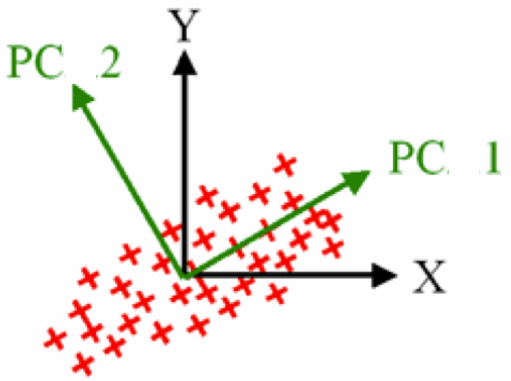
Possible Features from the Raw Data

- Derived Variables
  - Phase angle pair differences
  - Active and reactive power
  - Fast Fourier Transformation results
  - Correlations between variables
- Summaries
  - Mean over a specific time
  - Standard deviation over a specific time
- Trend
  - Rate of change over a specific time
  - Rate of rate of change (acceleration) over a specific time

Possible "Features" from the Features

- Dimensionality Reduction – Principal Component Analyses (PCA)

An orthogonal transformation to convert a set of correlated variables into a set of values that are linearly uncorrelated.



- Cluster the data, use results to calculate other measures, for example:

Mean silhouette value

A measure that determines how similar a set of features is to its cluster and to neighboring clusters. Also, informs when the observed system transitions to a different state.

To Impute or not To Impute – that is the Question

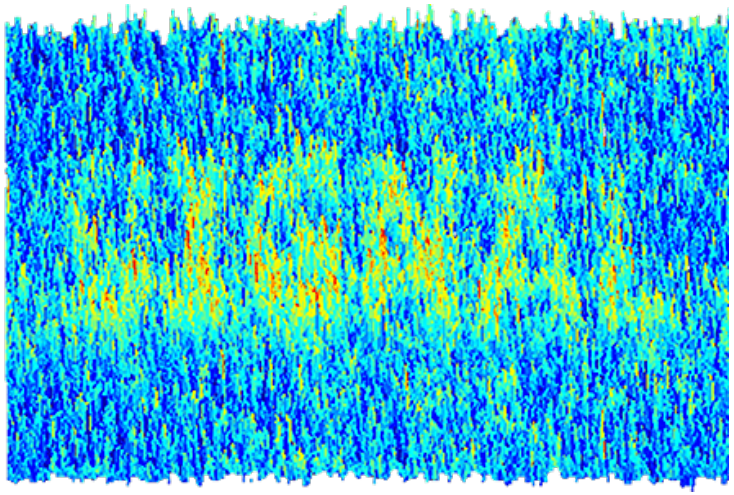
When data is bad or missing, is imputing the best solution?

- It Depends …
  - Imputing will introduce bias
  - Imputing for a randomly missing data point every once in awhile probably won't introduce too much bias
  - **BE AWARE** – Imputing usually introduces a different error structure than the original data and can often be detected with some investigation

- A better way …
  - Select and calculate features that are robust to missing data
  - Select analytical methods that are robust to missing data

Why Feature Selection?

- Event detection and identification is a signal to noise ratio problem.
- **If a feature is adding to the noise and not the signal, it will make finding the signal more difficult.**
- Events that are found may not be very interesting or insightful, if the feature is not very meaningful.



- If we can identify features that are helpful in detecting events of interest, then the algorithms that detect events and learn their identity will be more successful.

Feature Selection Methods

- Variable Importance

  A measure calculated when using CART (Classification and Regression Trees) or any other machine learning technique for supervised learning. Each variable is scored from 0 (no importance) to 100 (most important).

- LASSO or Ridge Regression

  Regression methods that model the data using a known response or outcome and includes performing variable selection and regularization, to enhance the prediction accuracy.

These methods require knowledge about what you are looking for

Unsupervised Event Detection

- This is a tough problem because the identity of the events is unknown and, because of that, the features that will identify them are unknown.
- Domain expertise and trial and error can be employed to help determine which variables and features help most in detecting interesting but unenvisioned events.
- As events are better understood, this process can help feed the information necessary to perform supervised learning and event identification.

Baselining Leading to Anomaly Detection

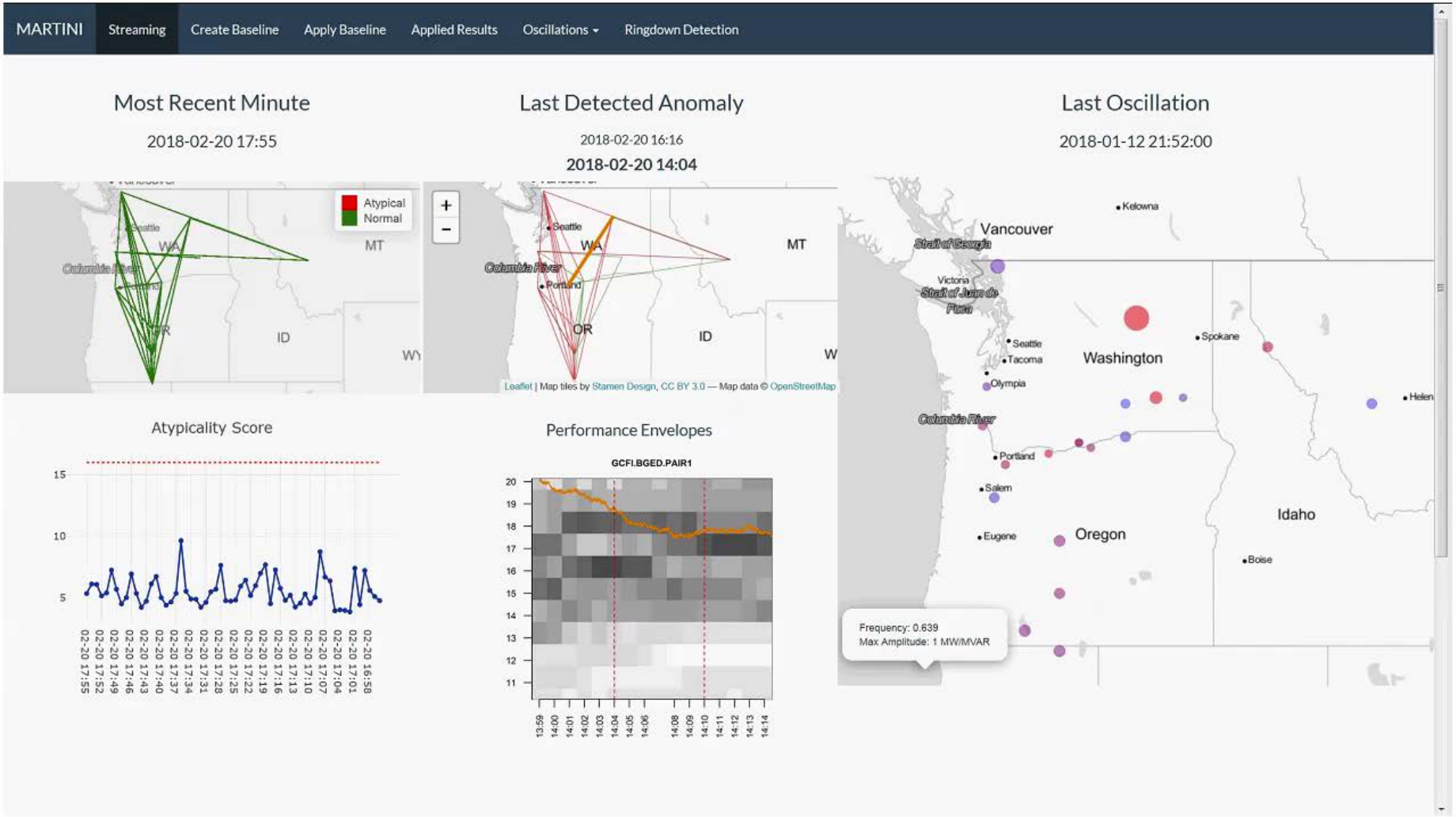
We have applied and tested multivariate statistical algorithms that define the baseline of commonly seen behavior and find departures from that baseline, using –

- 16 months of PMU data (Pacific NW data)
- Focus on phase angle pair difference (23 pairs)

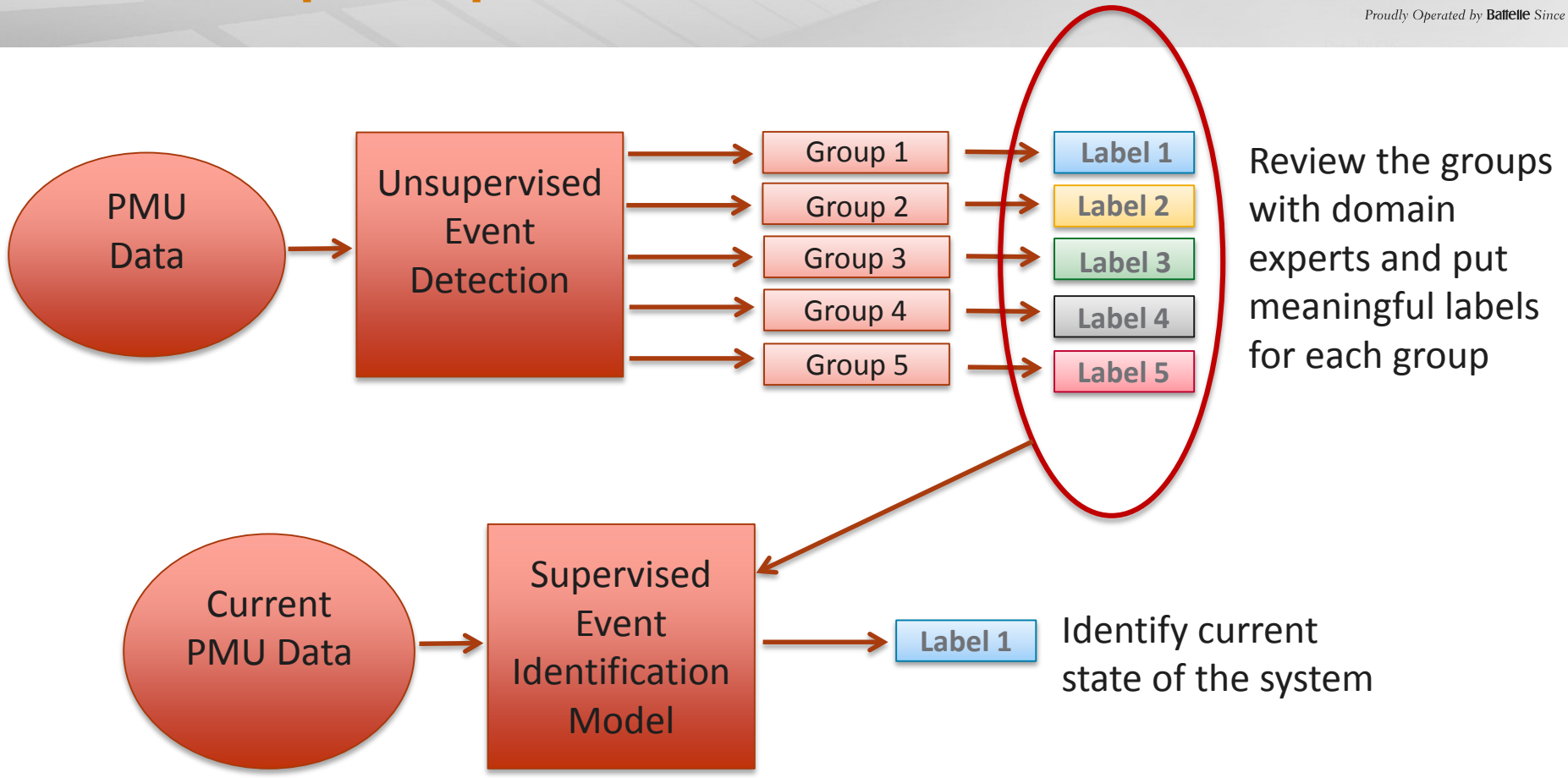
Results –

- MARTINI app – currently processing in near-real time and displaying results at the EIOC (Electricity Infrastructure Operations Center) at PNNL
- ESAMS (Eastern Interconnect Situational Awareness Monitoring System) – to be installed very soon to process Eastern Interconnect data

MARTINI App


MARTINI
Streaming
Create Baseline
Apply Baseline
Applied Results
Oscillations
Ringdown Detection
Most Recent Minute
2018-02-20 17:55
Last Detected Anomaly
2018-02-20 16:16
2018-02-20 14:04
Last Oscillation
2018-01-12 21:52:00
Atypical
Normal
Seattle
WA
MT
Columbia River
Portland
OR
ID
Leaflet | Map tiles by Stamen Design, CC BY 3.0 — Map data © OpenStreetMap
Kelowna
Vancouver
Strait of Georgia
Victoria
Strait of Juan de Fuca
Spokane
Washington
Tacoma
Olympia
Portland
Salem
Eugene
Oregon
Idaho
Boise
Helen
Frequency: 0.639
Max Amplitude: 1 MW/MVAR
Atypicality Score
Performance Envelopes
GCFI.BGED.PAIR1

Next Step – Supervised Event Identification



Conclusions

- The power grid community is in the infancy stages of applying statistical and machine learning algorithms.
- Care must be taken in determining which data should be used, how features can be extracted from the data, and selecting which features will provide insight.
- Data driven anomalies can be identified using multivariate analyses techniques. It's important to learn from these to help inform the next steps.
- As events are identified and better understood, supervised learning will automate the process and allow for real-time, decision-rich information.

… That's All Folks

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

Questions, comments, clairvoyant thoughts???

- Contact Info

Brett Amidan
b.Amidan@pnnl.gov