Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

Hunting for Anomalies in PMU Data

BRETT AMIDAN
JAMES FOLLUM
JEFFERY DAGLE

Pacific Northwest National Laboratory

NASPI Presentation (October 23, 2014)

“Big Picture” Objective

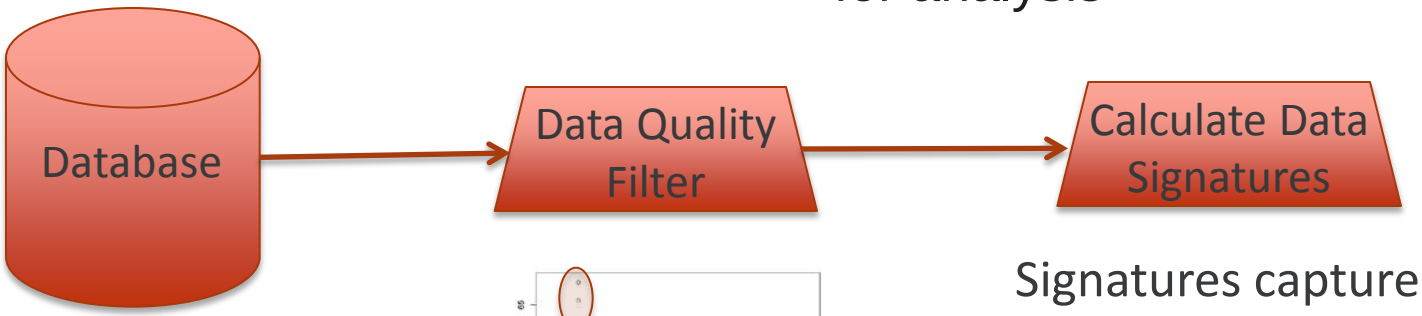
Power grid related data (PMUs, State Estimators, Load, etc)

Analytical Tool that provides:

- **Real time analytics, monitoring the state of the grid**
- **Capability to look at historical trends and events**
- **Reliable predictions about the forthcoming state of the grid**

Pre-Processing Steps

- Read and store raw **PMU data**
- Apply a data quality filter and calculate signatures (features) for analysis



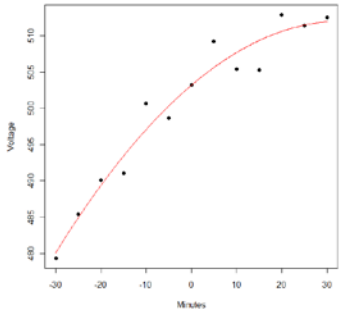
State Estimator Data

1 Yr = 5 GB

PMU Data

1 Yr (45 PMUs) = 3 TB

Signatures capture mean and variability as well as temporal aspects like slope and acceleration

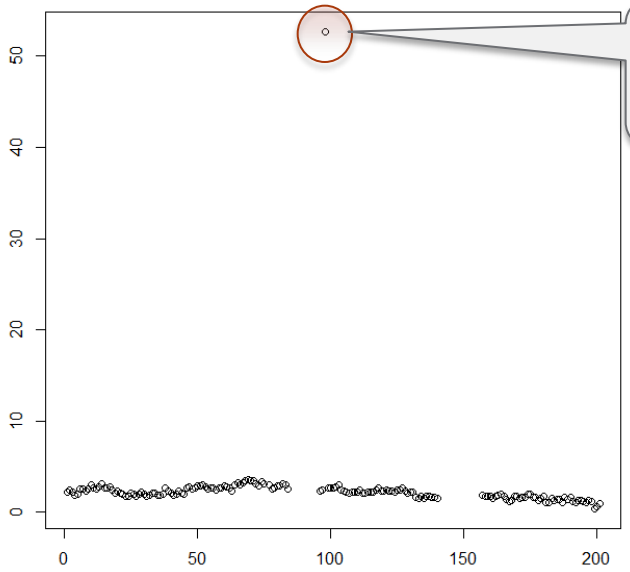


Data Storage

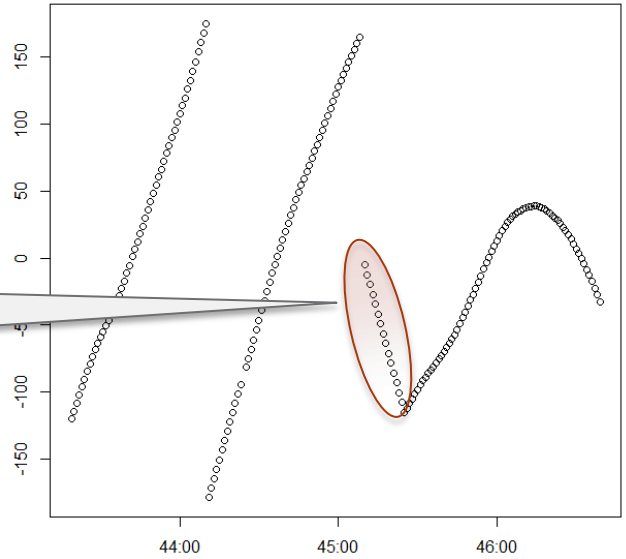
- 2 Months of Eastern Interconnect PMU data, down-sampled to 1 per second. Focus was studying phase angle pairs.
- Nearly 2 years of BPA PMU data, measured 60 times per second. Focus on frequency, voltage, current, & phase angle pairs.

- Data were not stored in a Hadoop cluster or any databases. They are difficult to set up and can be difficult to input data or export data.
- Data were stored in small, 1 minute files in Rdata format. This enabled the data to be easily read into R or Matlab for any additional plotting or analyses.
- This shifted the burden of data management on us, instead of relying on the software and the Hadoop or database experts.

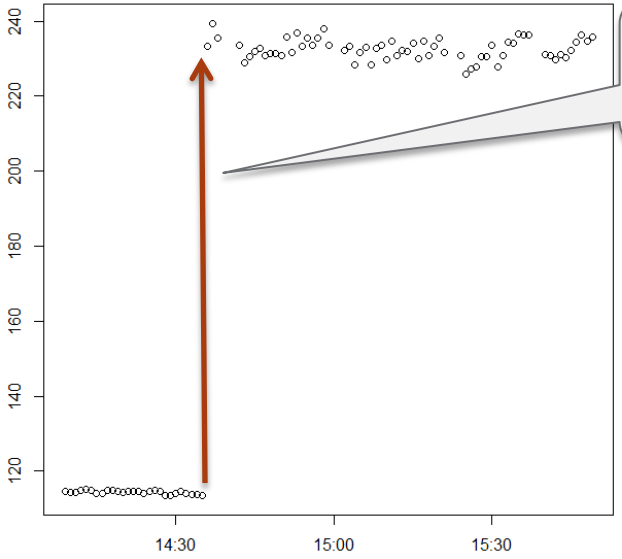
Data Quality Filters



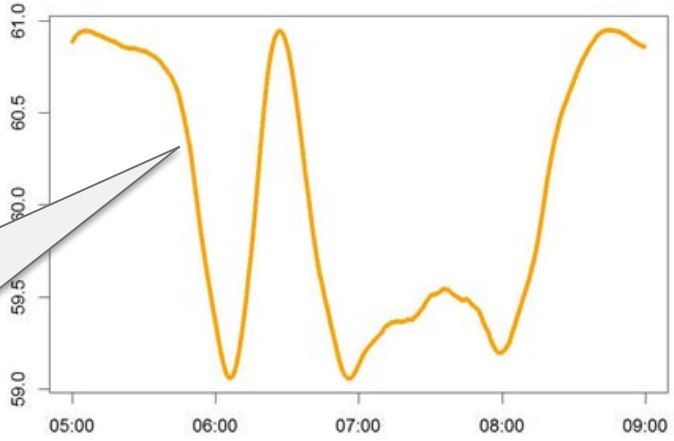
Usually easy to deal with



Phase angle wrap around issue

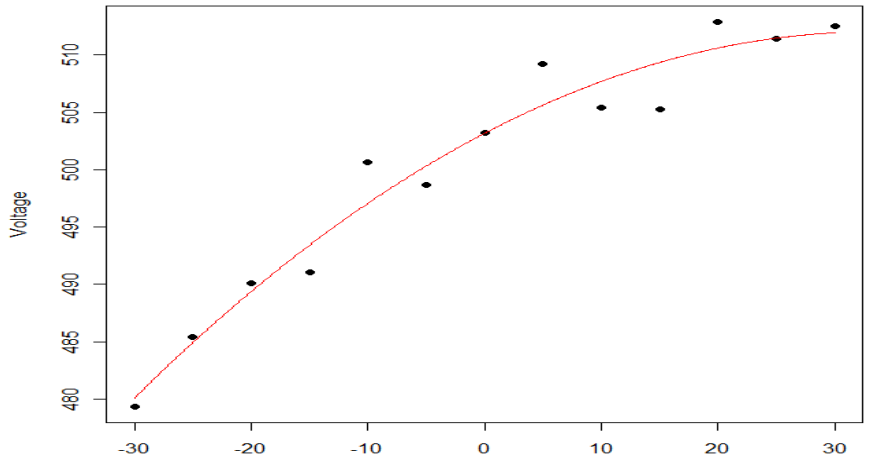


Sudden changes in phase angle



Impossible variable behavior (frequency ranging from 59 to 61)

Feature Extraction (Data Signatures)



- Regression fits through the data calculate estimates of value, slope, curvature (acceleration), and noise.
- Can be calculated in the presence of missing or data quality flagged values.
- Summaries of these features are used in the analyses.

Data Analyses

- Find **patterns** in the data

  Apply clustering algorithms to determine patterns (similarities) in the data.

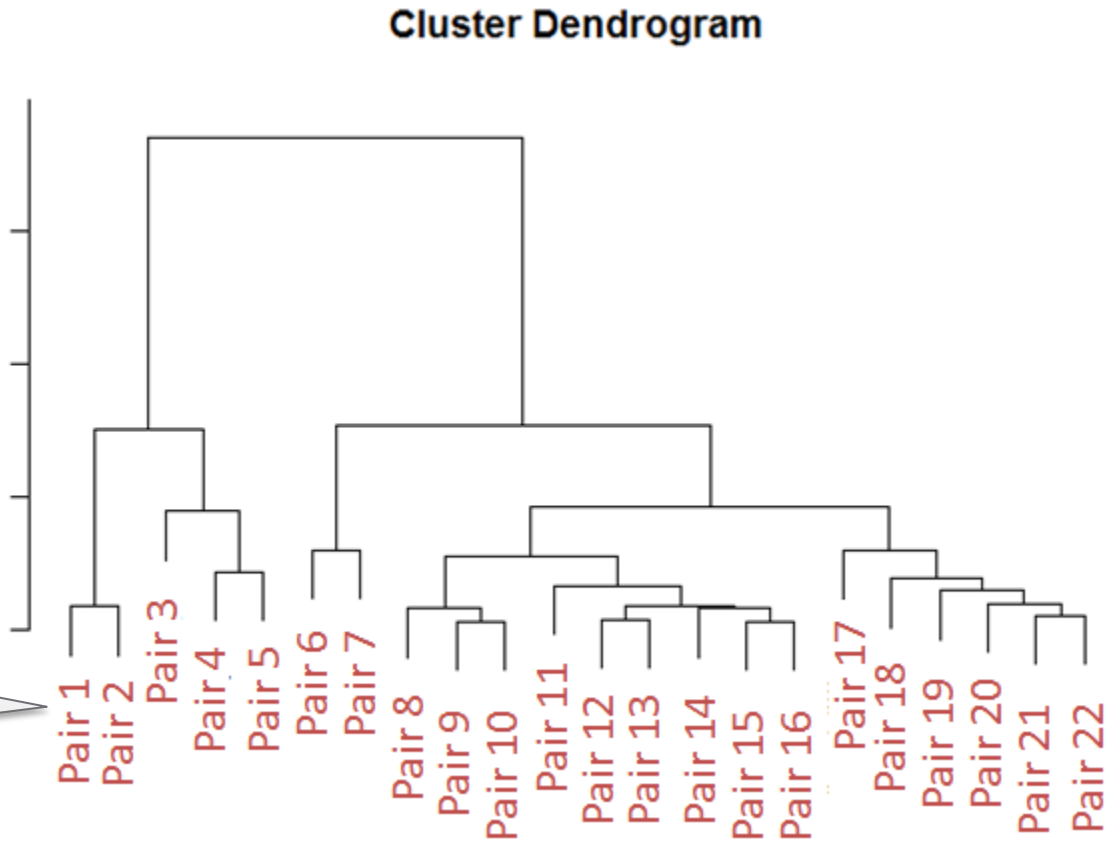
  These similarities could be determined by time (i.e. periods of time in which the grid is in a similar state) or by variables (i.e. variables that are closely related).

- Find **anomalies** in the data (referred to as atypicalities)

  Apply multivariate analytical techniques to find unusual moments in time with respect to a baseline that was determined from historical data.

Eastern Interconnect
Clustering of Phase Angle Pairs

Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle *Since 1965*

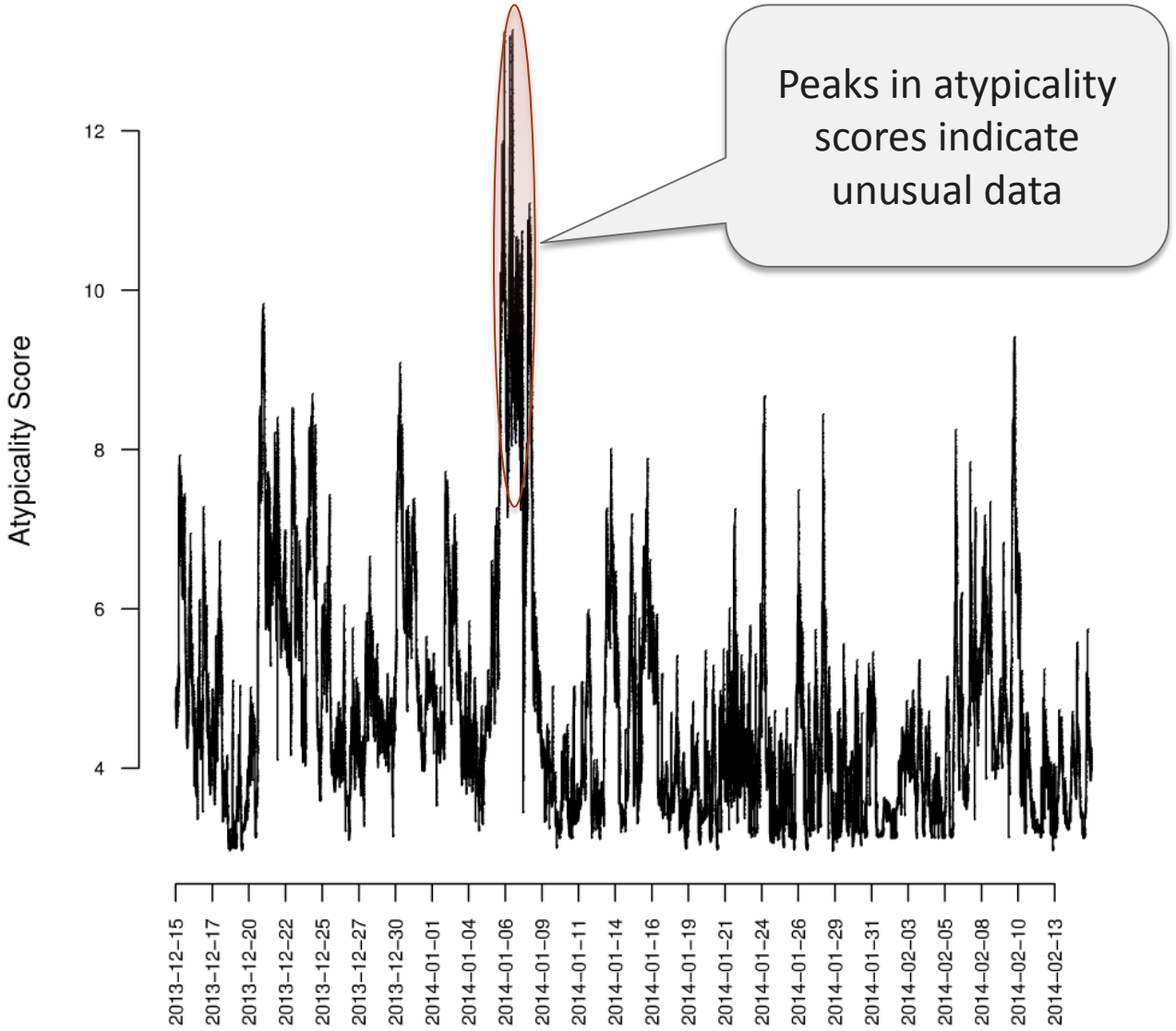


Proximity on tree indicates similarity

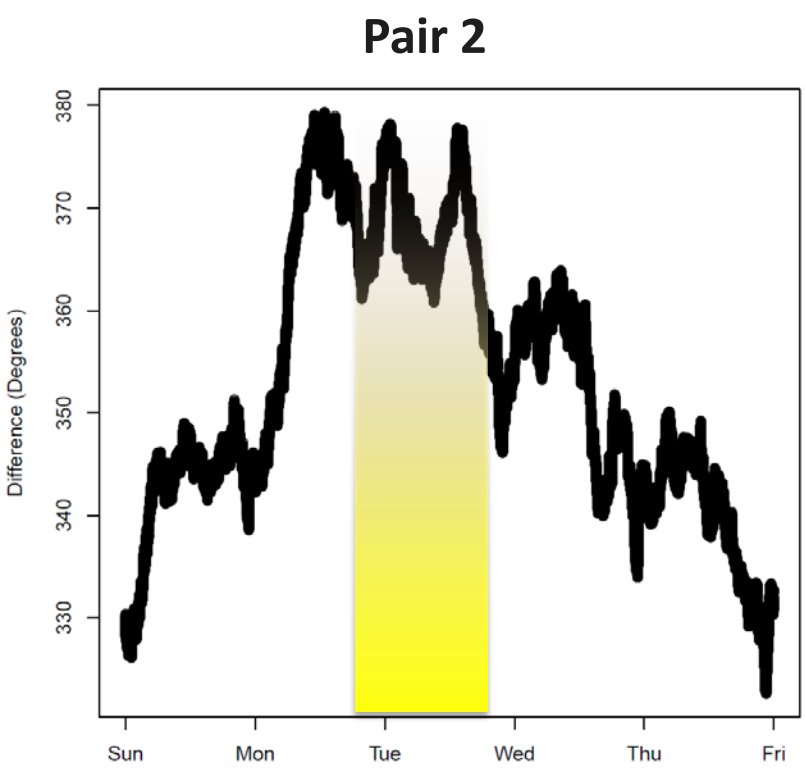
Eastern Interconnect PMU Data Atypicality Detection

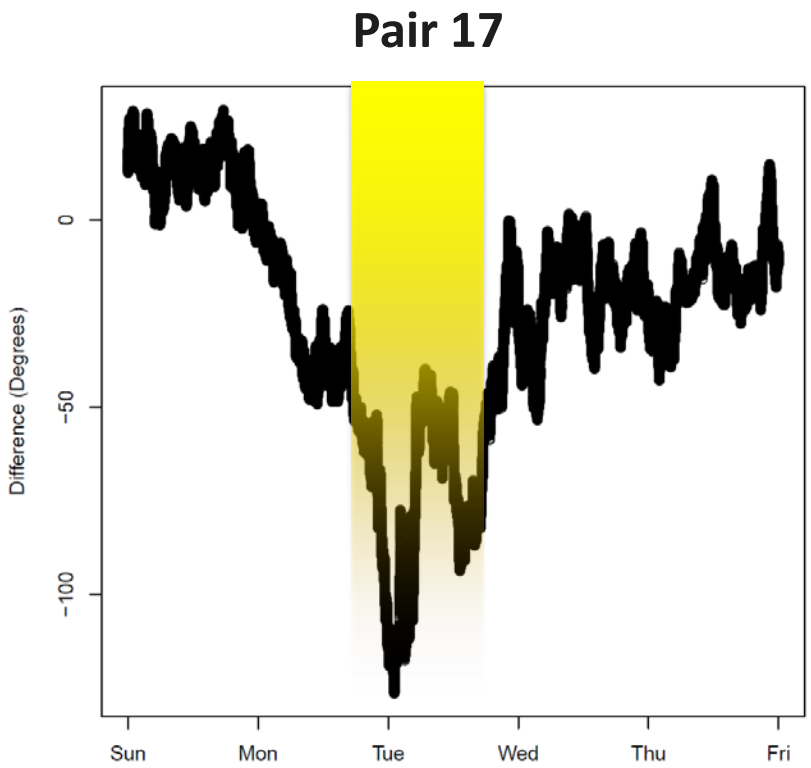
Analyses performed on selected phase angle pairs across the Eastern Interconnect

This anomaly corresponded to a heavy load time period



Eastern Interconnect Drilldown Plots





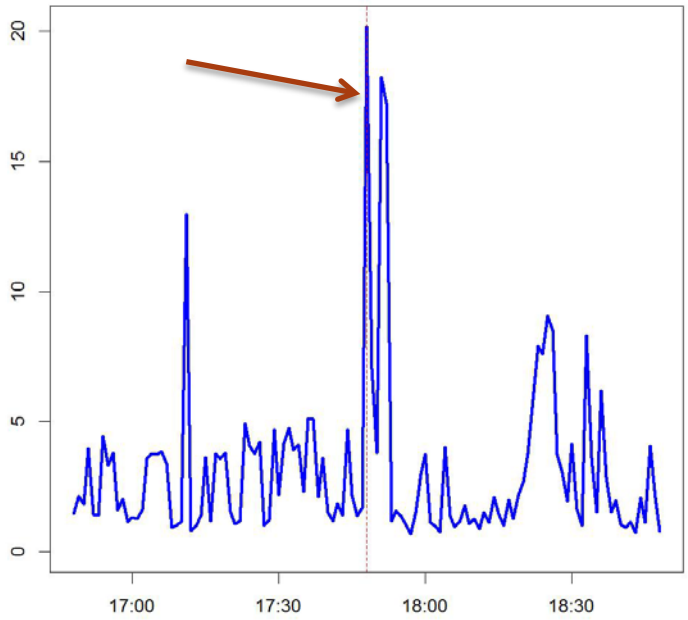
Two angle pairs that showed atypical behavior when the atypicality score was high. Pair 2 shows atypical behavior before the atypicality score triggered.

Studying this type of behavior may help discover possible precursors to atypical events.

BPA PMU Data
Atypicality Detection

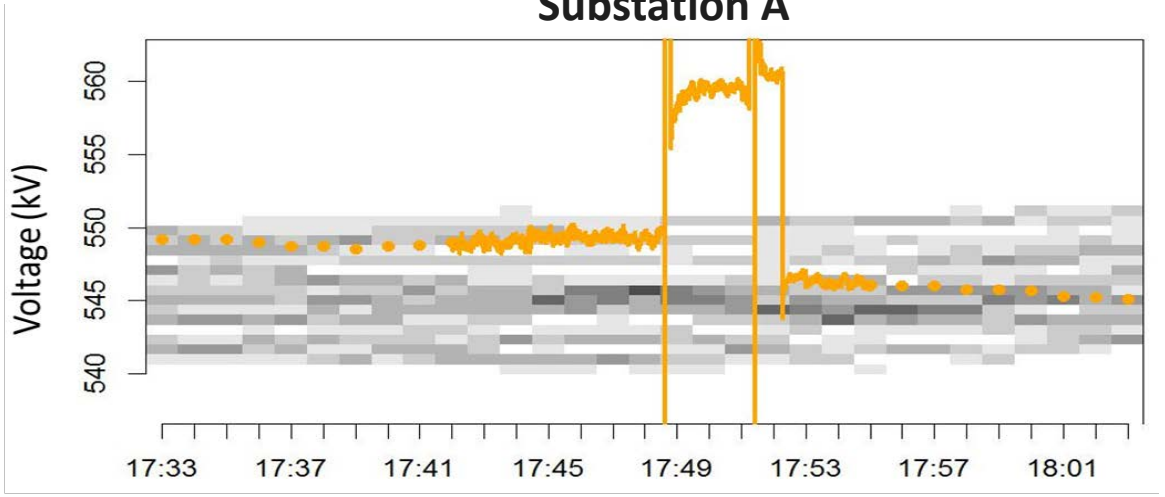
Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

Atypicality Score

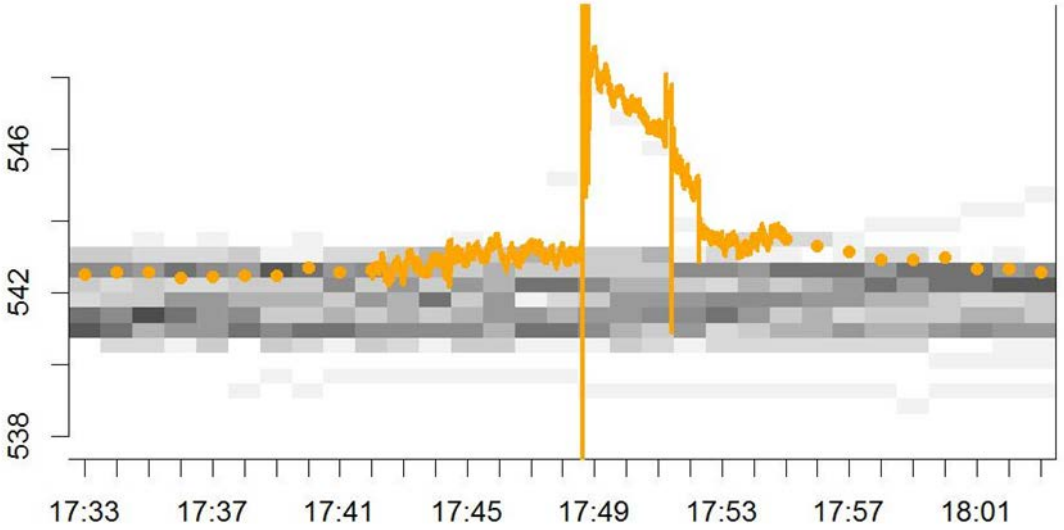


Lightning related anomaly

Substation A



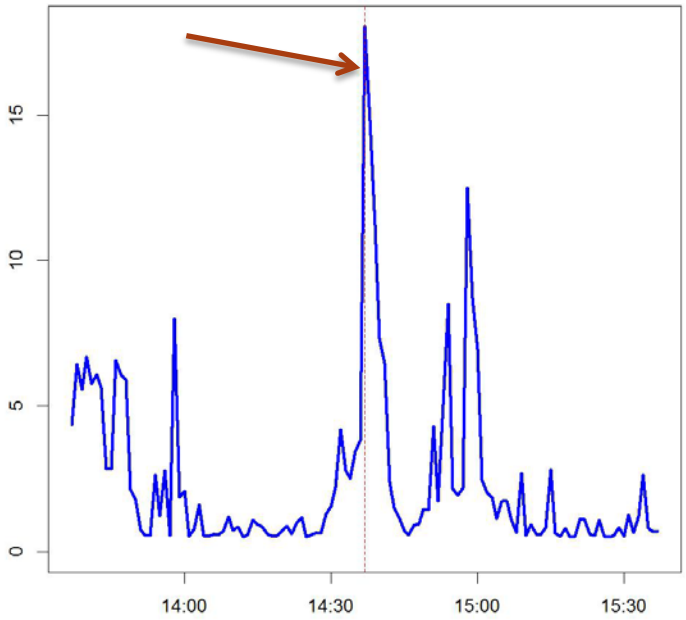
Substation B



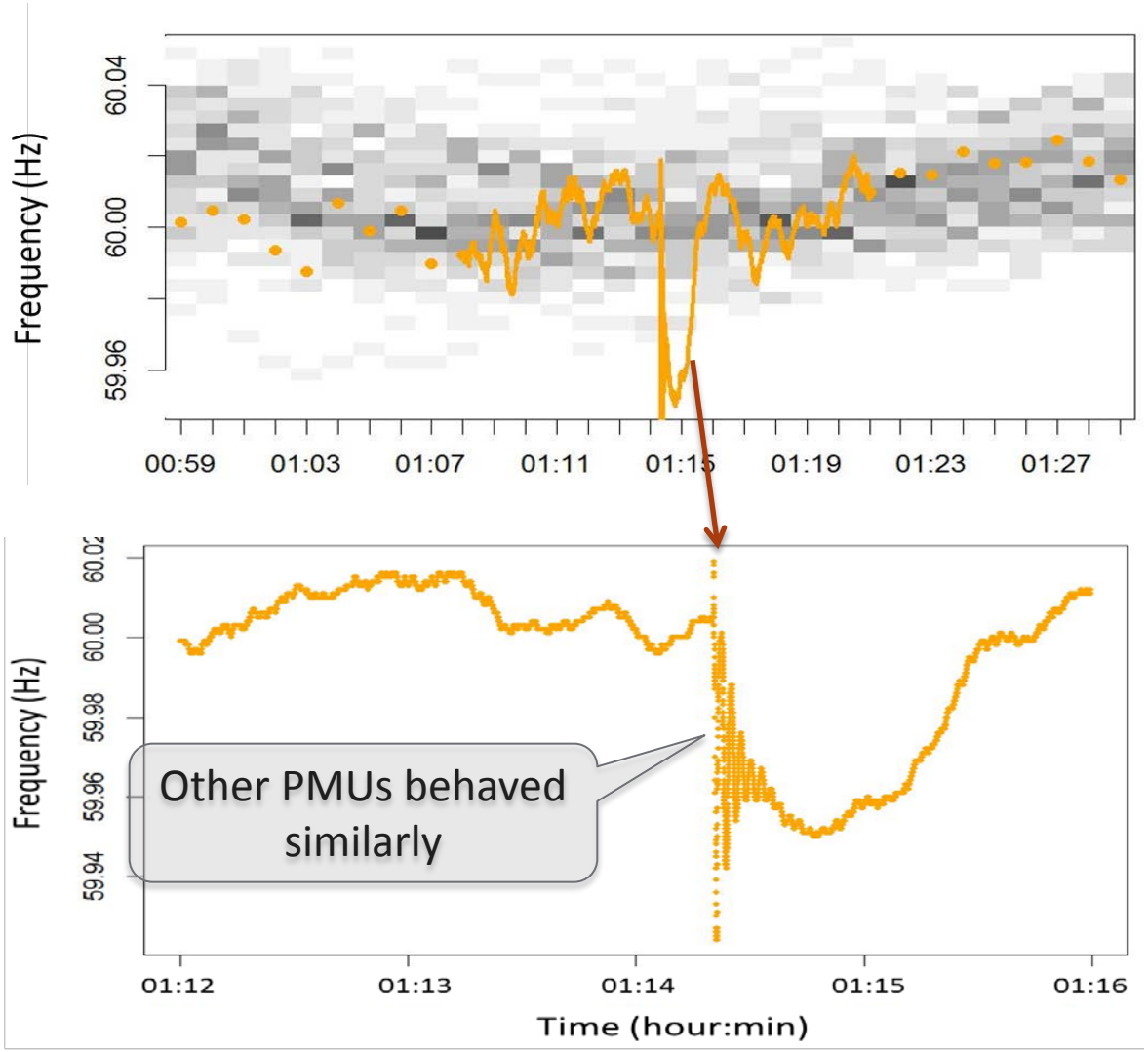
Atypicality Detection - Example

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

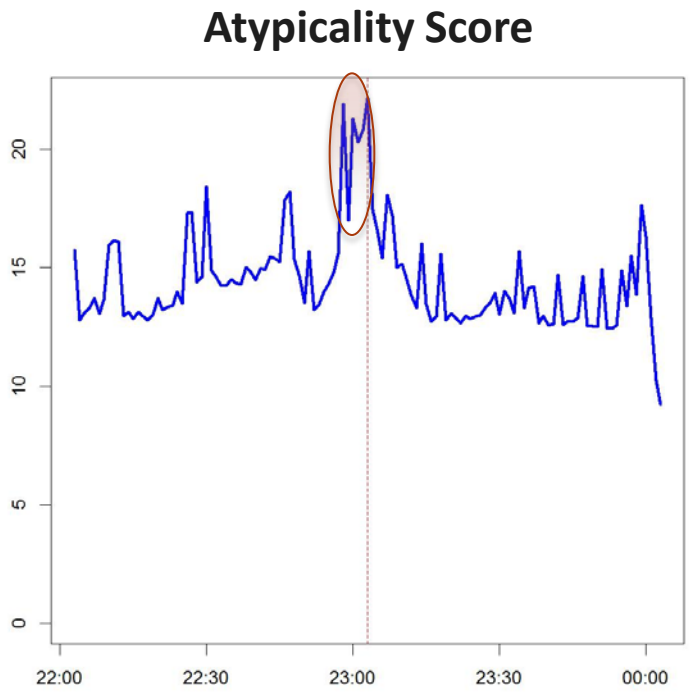
Atypicality Score

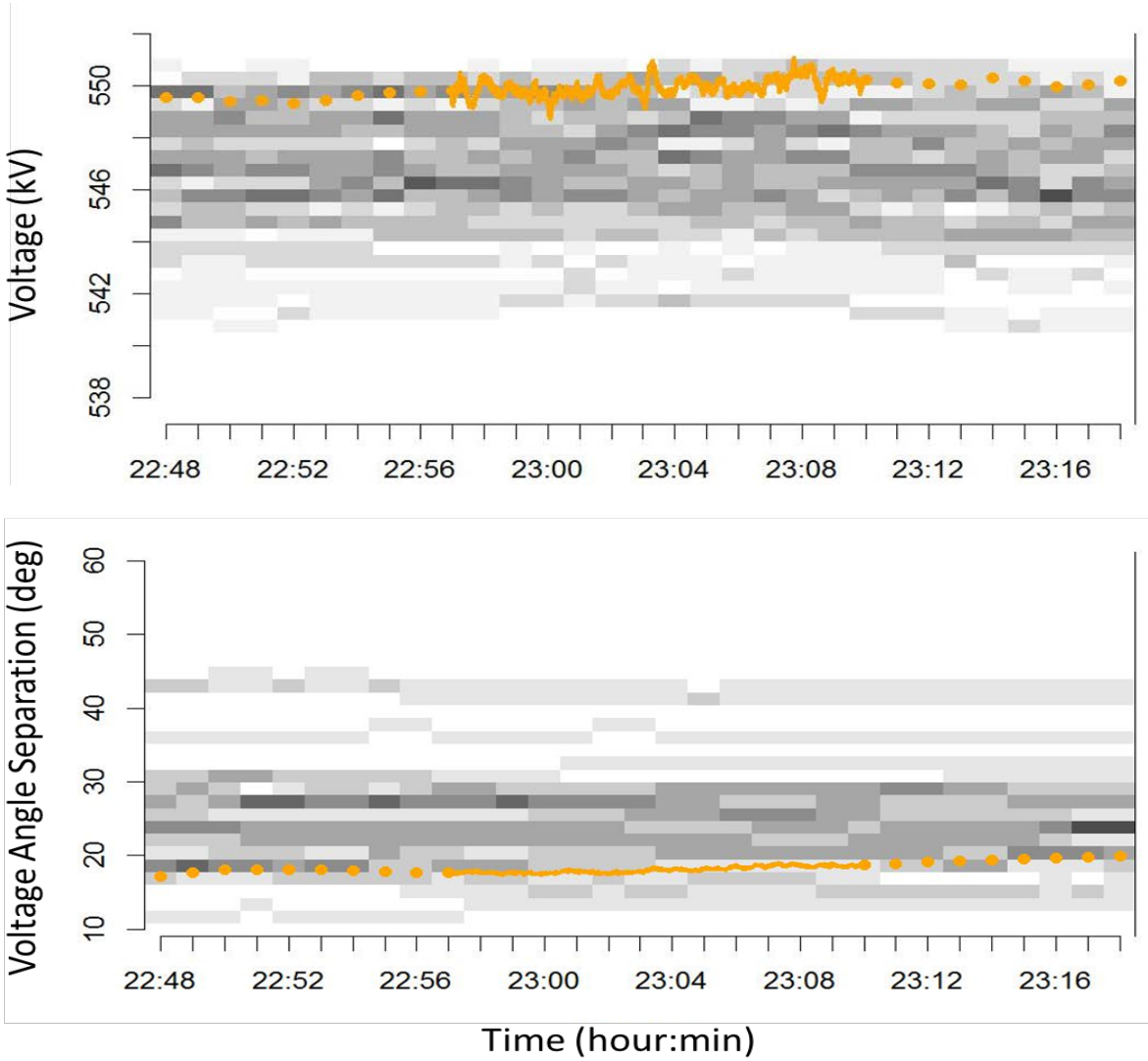


Equipment failure related anomaly



Atypicality Detection - Example





An example of a detected anomaly that was not related to a specific system event

Next Steps

- Attach these tools to a PMU data stream for near-real time results.
- Add other data streams for analyses.
- Process a longer time period.
- Determine what thresholds should be used to alert when data are atypical.
- Investigate data for predictive patterns.