Pacific Northwest
NATIONAL LABORATORY

DOE FOA 1861 Research Outcomes

October 18, 2022

NASPI Work Group Meeting
and Vendor Show
Charlotte, NC

Shuchismita Biswas

Electrical Engineer, PNNL

U.S. DEPARTMENT OF ENERGY BATTELLE

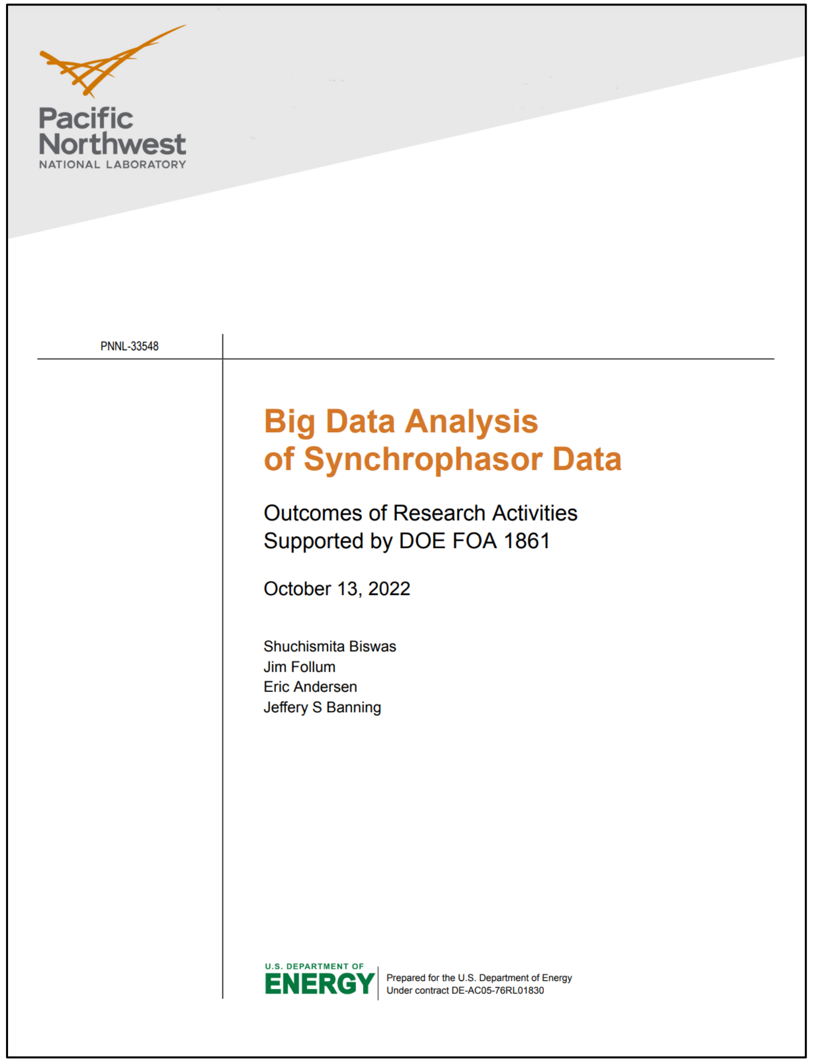
PNNL is operated by Battelle for the U.S. Department of Energy

PNNL-SA-178557

Background

- Aimed at advancing the state-of-the art in big data analytics applied to transmission-level PMU data
- First-of-its-kind large anonymized PMU dataset compiled- spanning multiple years, three US interconnections and with event logs (~20 TB)
- Eight research grants to teams formed by industry and academia
- Developed/evaluated methodologies at different TRLs
- Fast dissemination of major findings to the power systems community through a *meta-analysis report*

S. Biswas, J. Follum, E. Andersen and J. Banning, “Big Data Analysis of Synchrophasor Data: Outcomes of Research Activities Supported by DOE FOA 1861”, PNNL-33548, Oct 2022.


Pacific Northwest
NATIONAL LABORATORY

PNNL-33548

Big Data Analysis of Synchrophasor Data

Outcomes of Research Activities Supported by DOE FOA 1861

October 13, 2022

Shuchismita Biswas
Jim Follum
Eric Andersen
Jeffery S Banning

U.S. DEPARTMENT OF ENERGY
Prepared for the U.S. Department of Energy
Under contract DE-AC05-76RL01830

Please email shuchismita.biswas@pnnl.gov to request a copy of the report.

Awardees

Lead	Partners	Project
PingThings		Combinatorial Evaluation of Physical Feature Engineering and Deep Temporal Modeling
GE Research	GE Grid Solutions	PMU-Based Data Analytics using Digital Twin and PhasorAnalytics Software
Schweitzer Engineering Laboratories	Oregon State University	Machine Learning Guided Operational Intelligence from Synchrophasors
Siemens Corporation	Southern Methodist University, Temple University	MindSynchro
University of California, Riverside	Electric Power Group (EPG), Michigan Technological University	Discovery of Signatures, Anomalies, and Precursors in Synchrophasor Data with Matrix Profile and Deep Recurrent Neural Networks
University of Nevada, Reno	Arizona State University, IBM, Virginia Tech	A Robust Event Diagnostics Platform: Integrating Tensor Analytics and Machine Learning Into Real-time Grid Monitoring
Iowa State University of Science and Technology	Electric Power Group (EPG), Google Brain, IBM	Robust Learning of Dynamic Interactions for Enhancing Power System Resilience
Texas A&M Engineering Experiment Station	Temple University, Quanta Technology	Big Data Synchrophasor Monitoring and Analytics for Resiliency Tracking (BDSMART)

Pacific Northwest
NATIONAL LABORATORY

Key Outcomes

- Robust data management pipelines
- Event detection and classification algorithms
- Feature engineering approaches
- Large repository of events, synthetic data
- Transfer learning techniques

The performance of proposed approaches may be enhanced by including additional information like topography, SCADA data, outage reports etc.

Contents

Pacific Northwest
NATIONAL LABORATORY

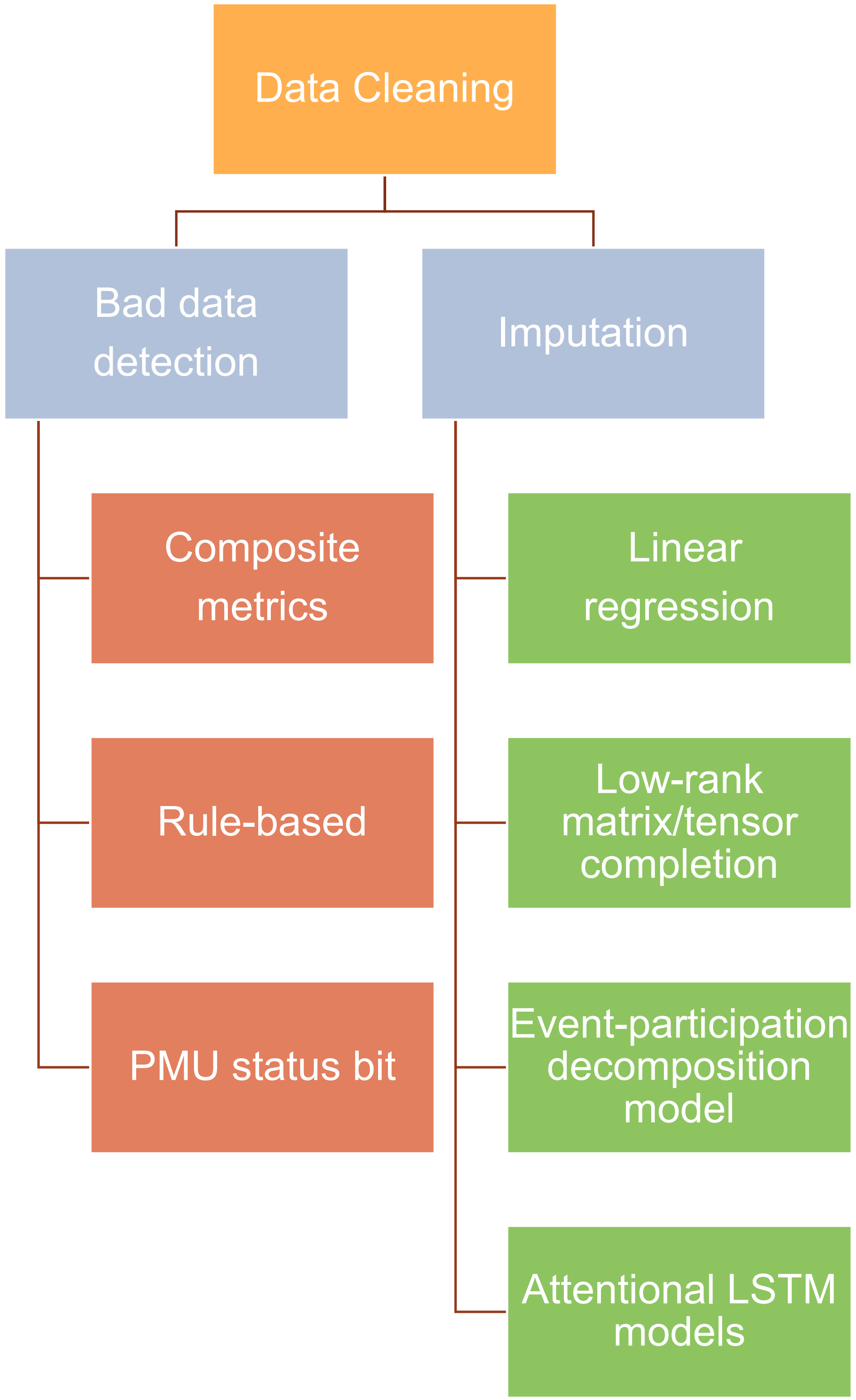
Algorithms Developed

- Data Cleaning
- Feature Engineering
- Event Detection
- Event Classification
- Supporting Tools

- Single PMU or multi-PMU implementation?
- Single-channel or multi-channel implementation?
- Can off-the-shelf AI/ML applications be applied directly?
- How to leverage SME knowledge?

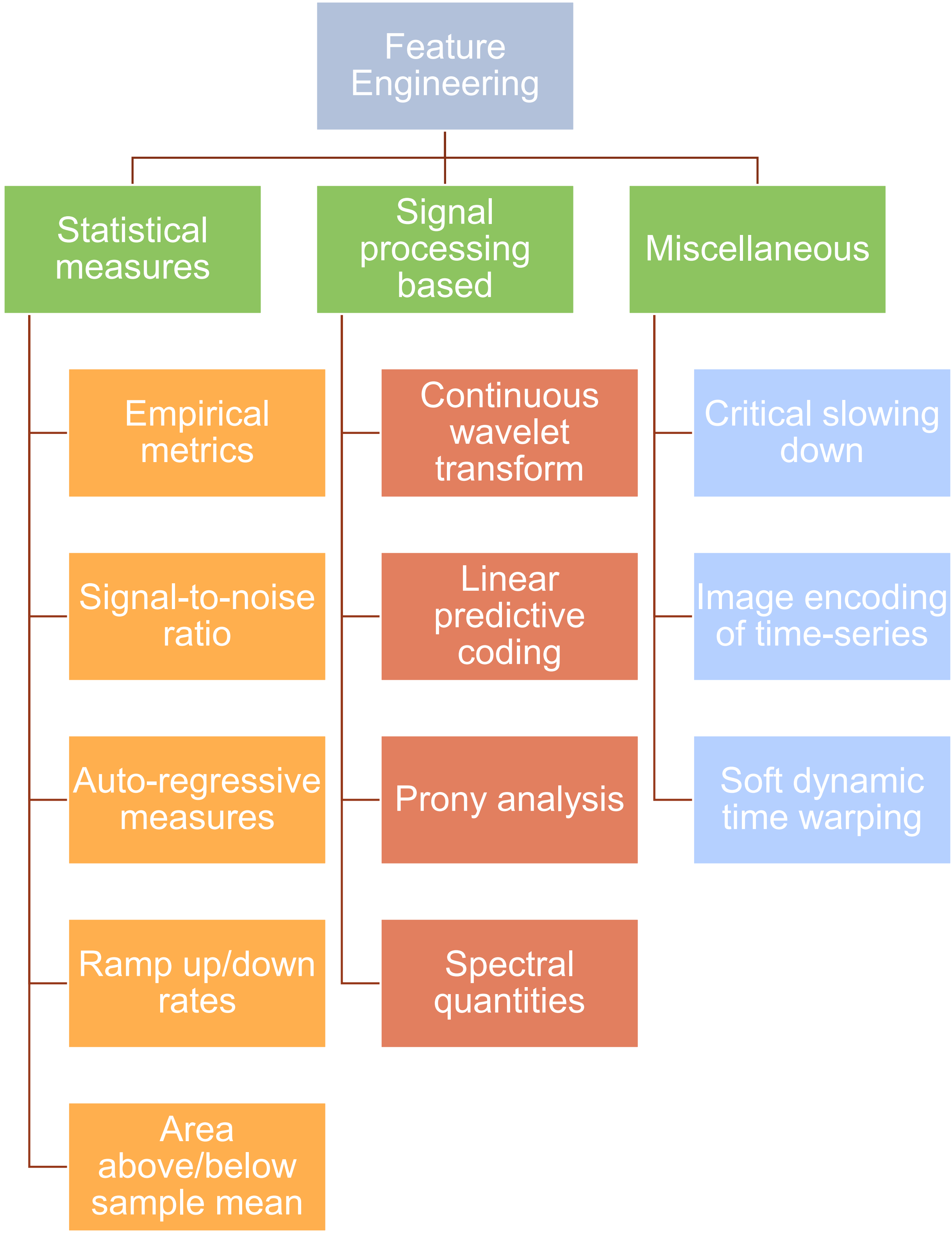
Data Cleaning

- Bad data detection:
  - Rule-based methods for statistical outliers, stale values, physically impossible values
  - PMU status bits flag erroneous values
- Imputation:
  - Leverage spatiotemporal correlation
  - Reconstructing archived data
  - Forecasting incoming measurements



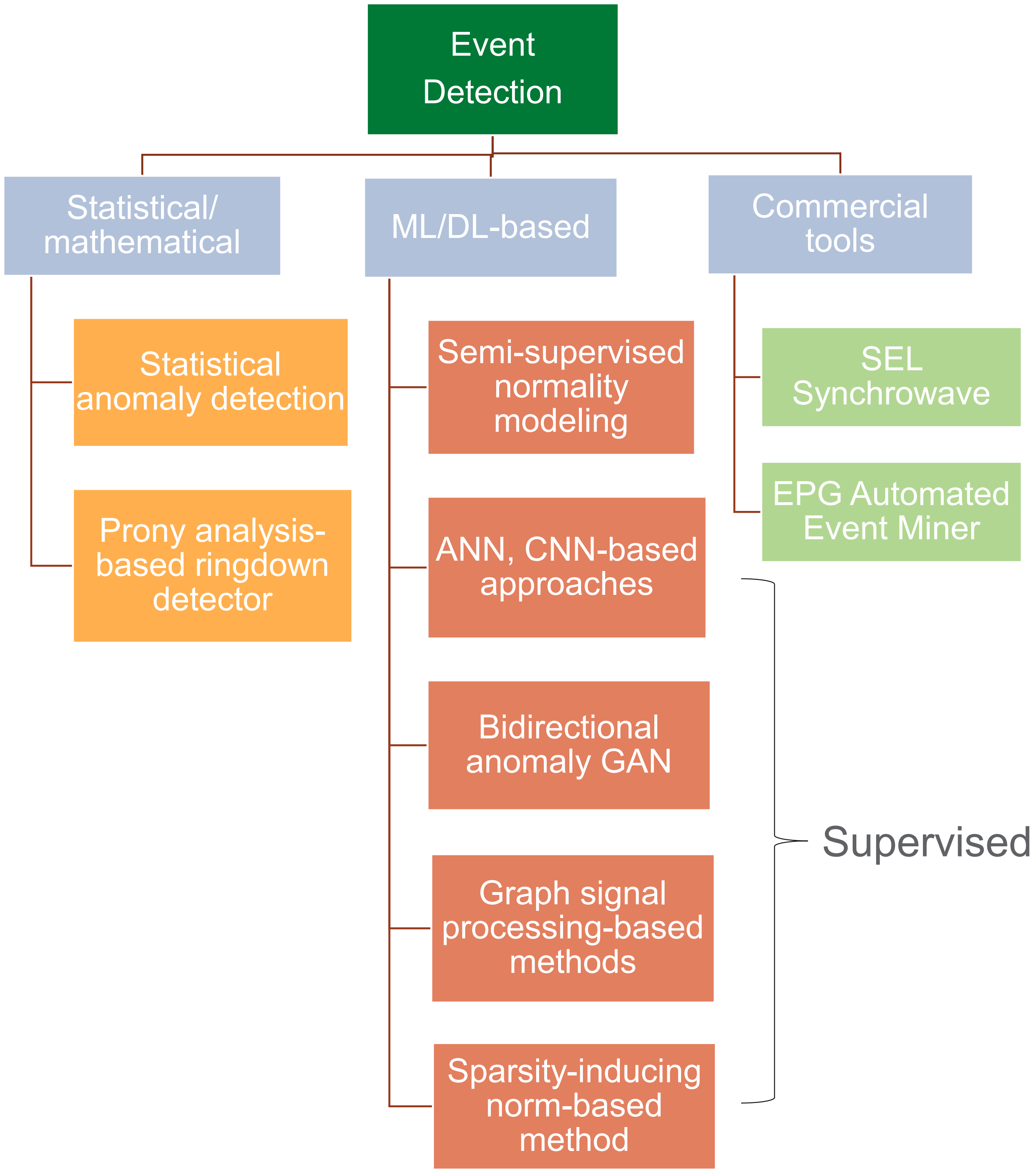
Feature Engineering

- Various feature-engineering methods developed that may also find use in other applications
- Statistical and spectral parameters are easier to interpret
- Dimensionality reduction techniques like PCA used to reduce the number of features to be fed to ML models



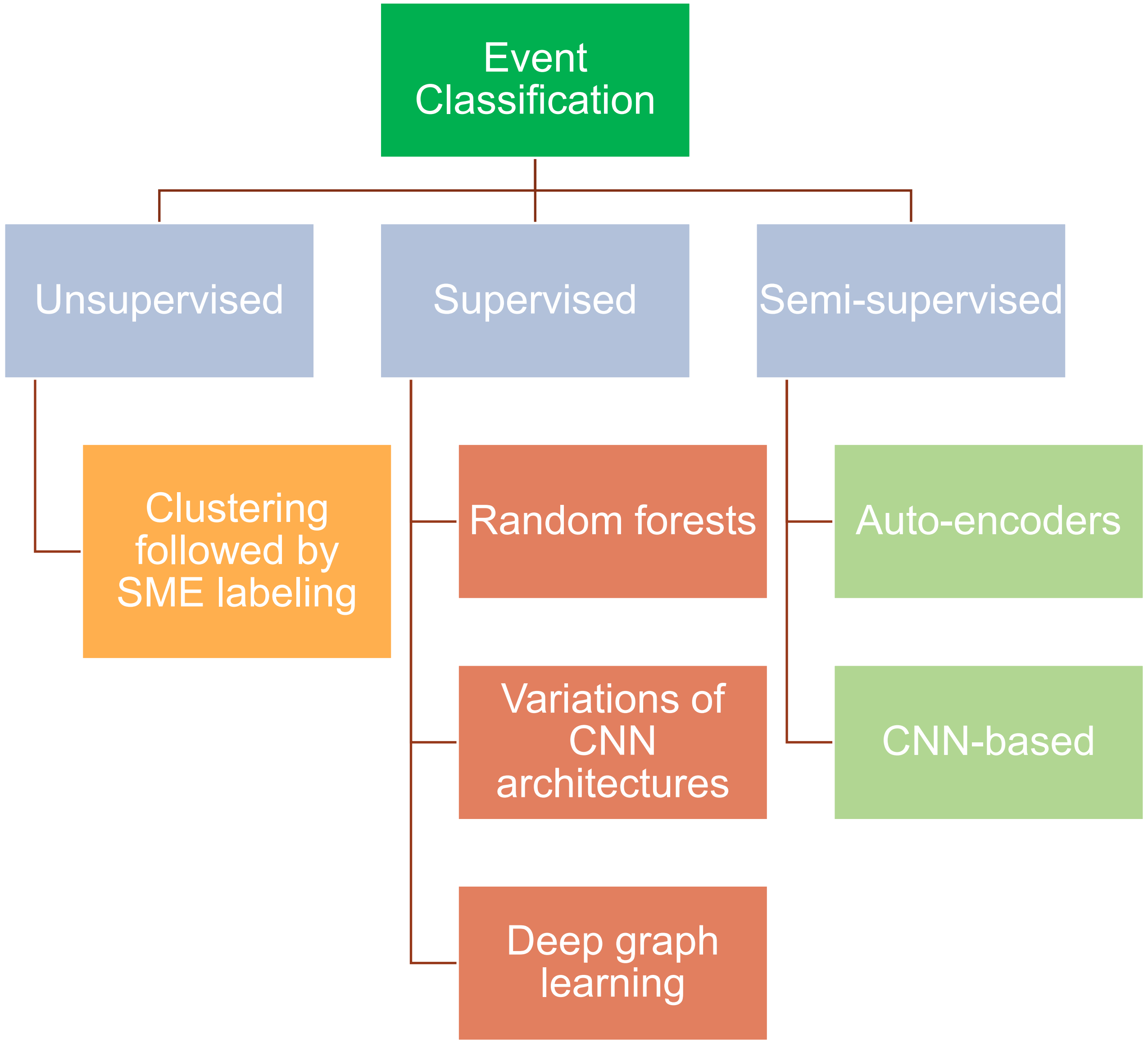
Event Detection

- Summary statistics are computationally efficient at detecting anomalous data periods, but may not be adequately selective
- Lack of refined labels may necessitate semi-supervised learning
- The proposed algorithms successfully detected thousands of events not documented in the event logs



Event Classification

- One *multi-class classifier* or an *ensemble of binary classifiers*?
- Hierarchical classification: how granular should event classifiers be?
- Popularity of CNN-based architectures



Supporting Tools

- Supporting tools
  - Automated labeling applications
  - Data Visualization
    - Multiresolution plotter (proprietary)
    - Largest triangle three buckets
  - Transfer learning
  - Synthetic data generation

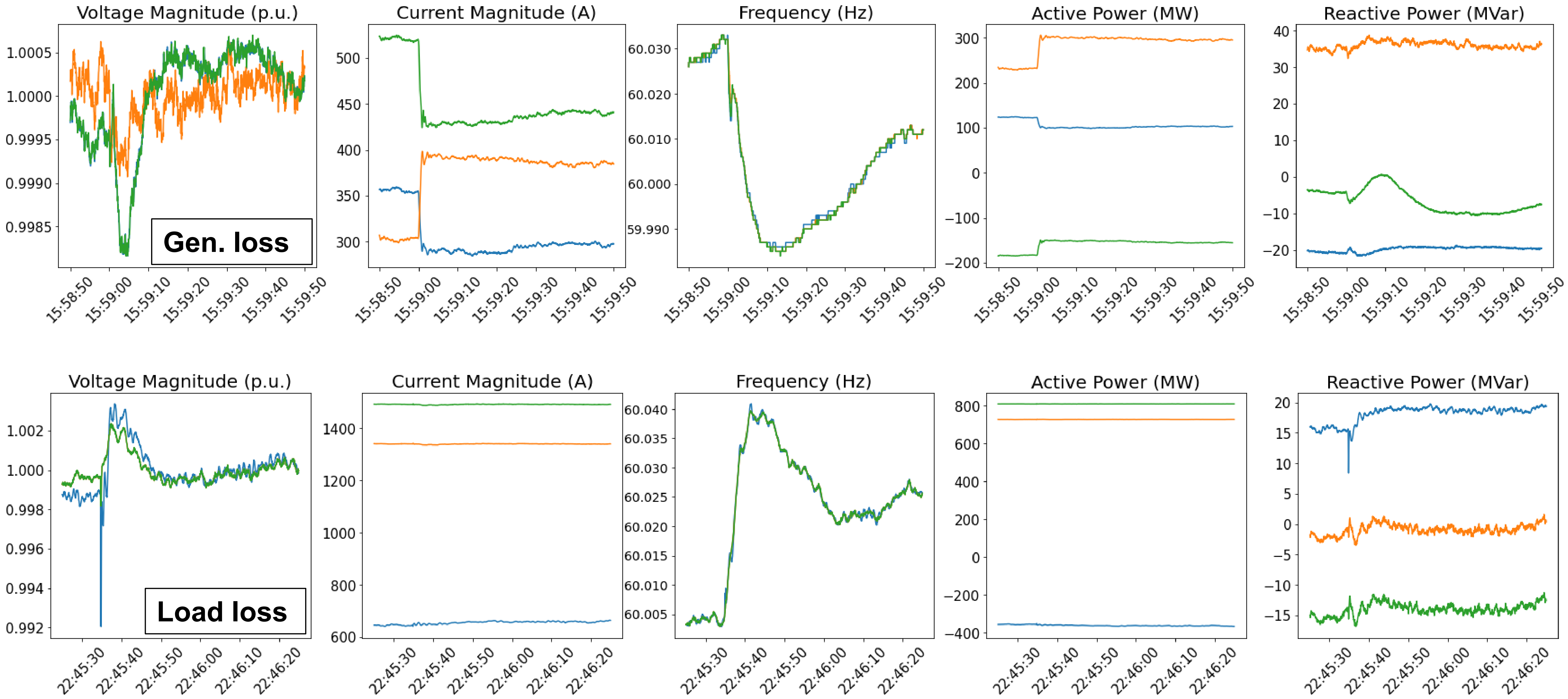
- Oscillation analysis
- Identifying precursors to transformer failures
- GPS spoofing mitigation using PCA

Event Signatures

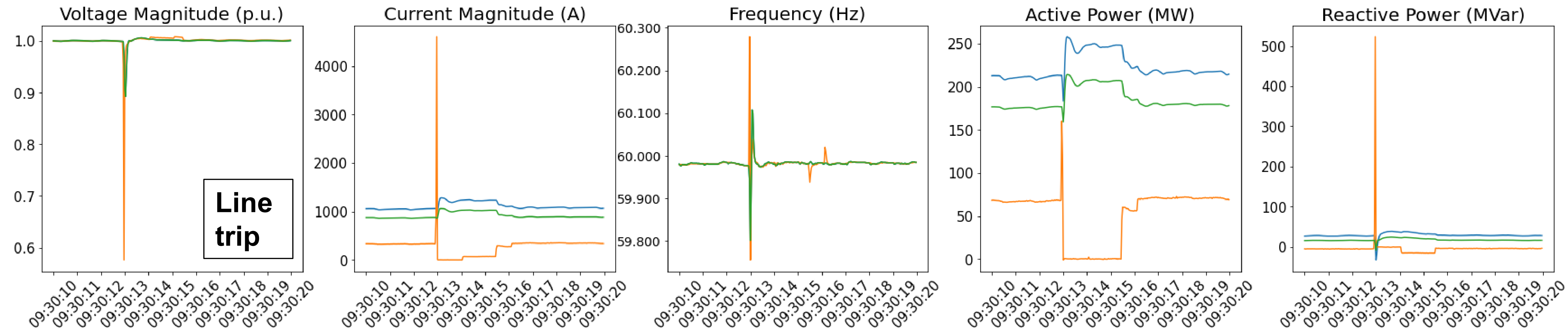
- With the large repository of detected events, signatures and commonalities for different event groups could be identified.
- These signatures helped awardees fine-tune their event detection and classification strategies.
- Confirms many power engineering intuitions, and useful in illustrating expected behavior to data science SMEs without power engineering backgrounds.
- Signatures across interconnections are consistent. Hence, algorithms developed are generalizable.

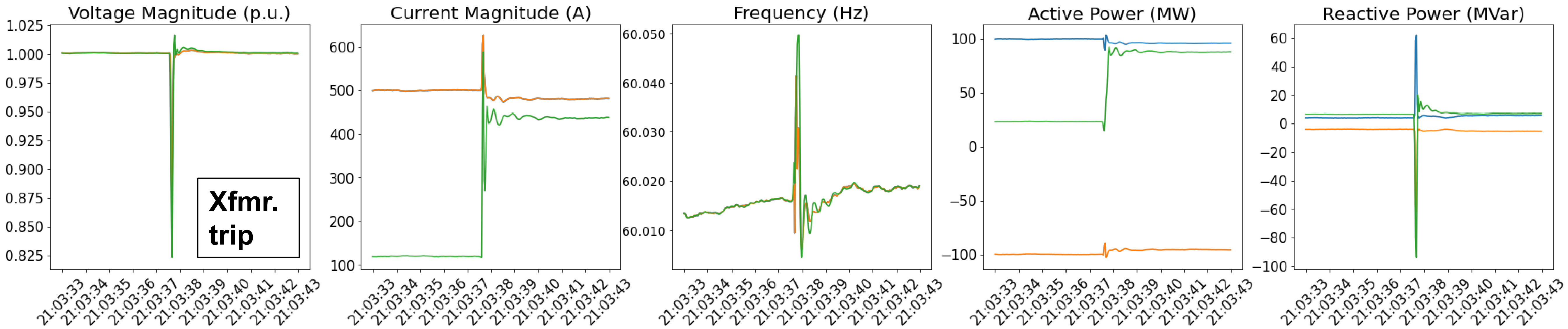
Event Group	Event Type
Frequency	Loss of generation
	Loss of load
Voltage	Line trip
	Transformer trip
Oscillation	Ringdown
	Forced Oscillation

Frequency Events



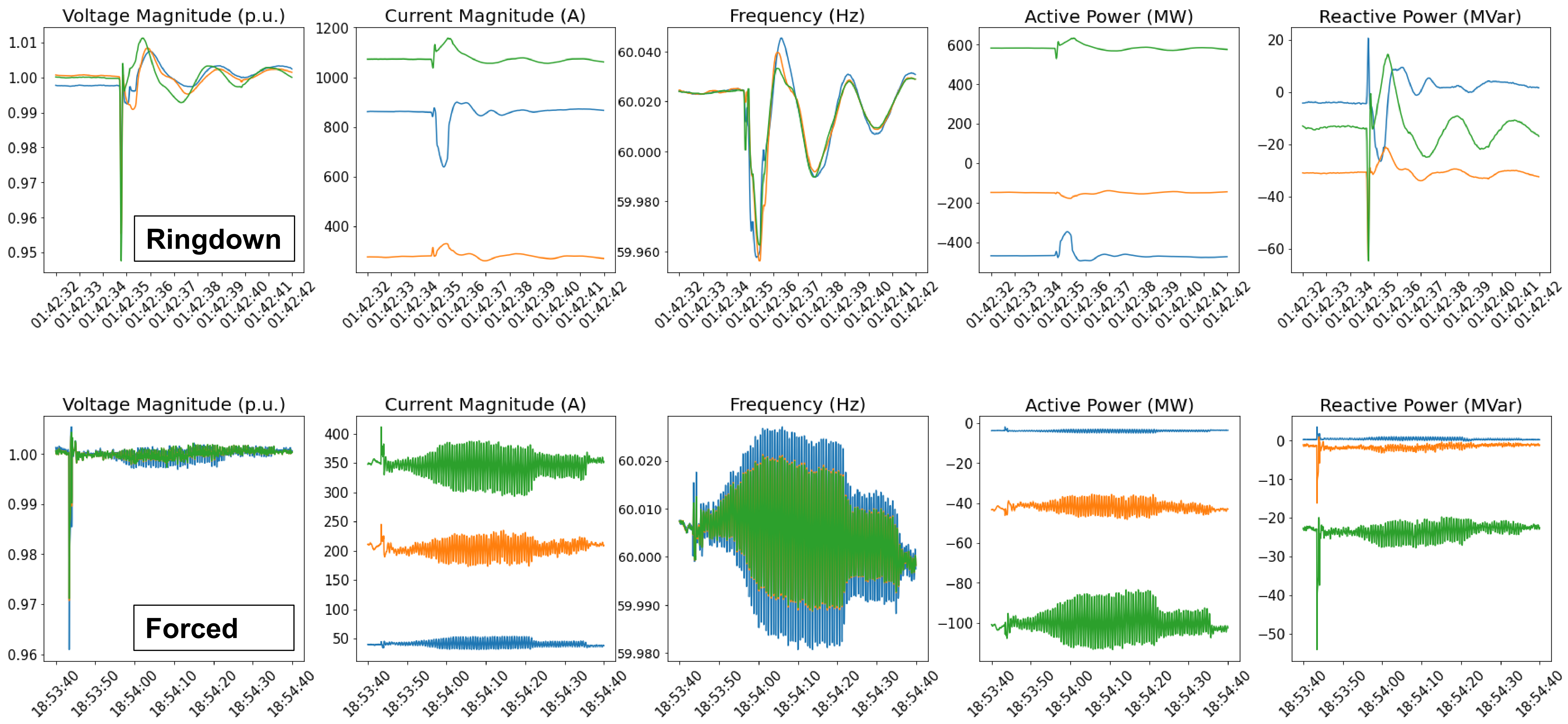
Voltage Events



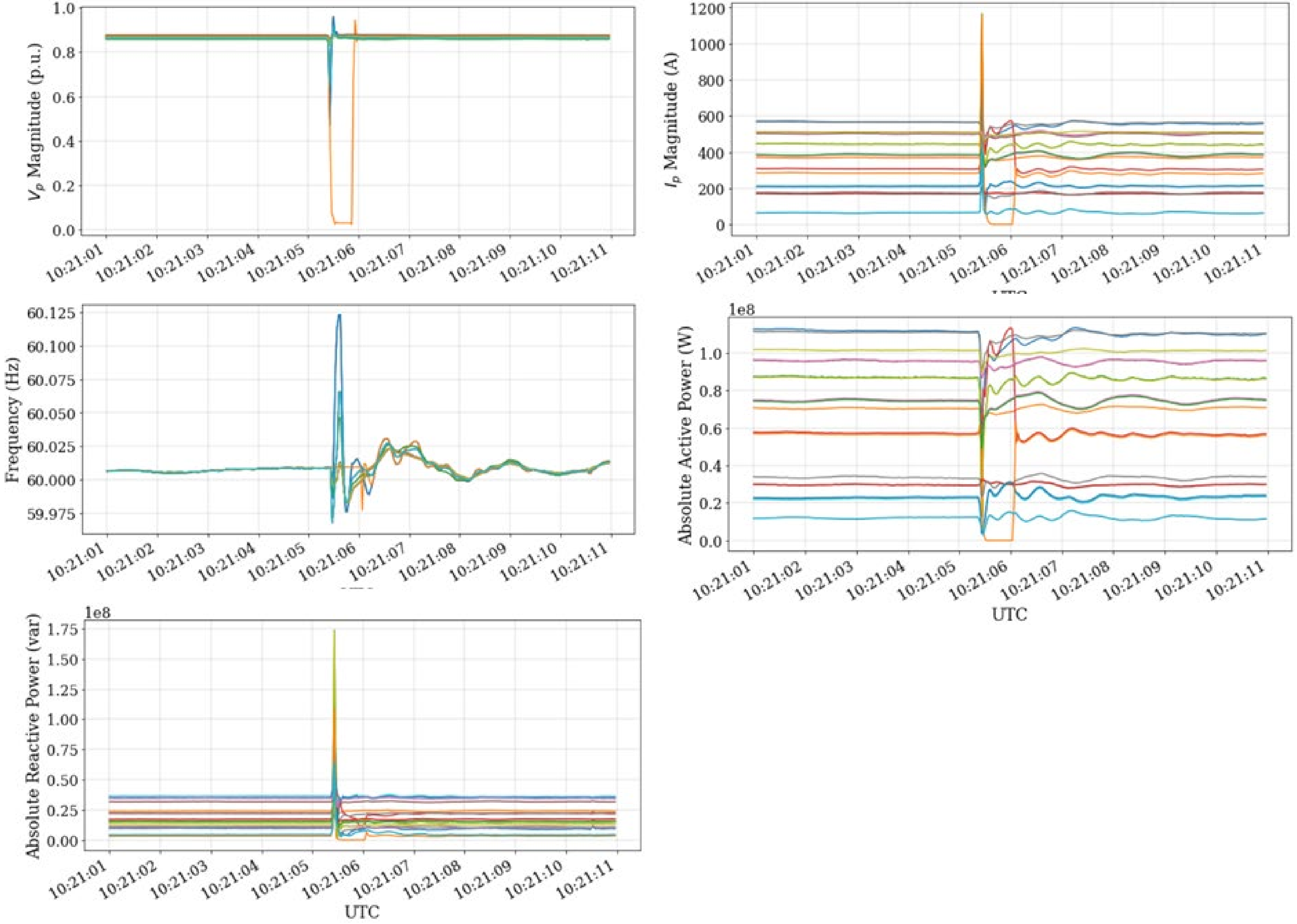


Pacific Northwest
NATIONAL LABORATORY

Oscillation Events



Autoreclosure



Some Closing Thoughts

- *What can AI/ML do with PMU data?*

Not yet ready to provide full diagnosis of events and their root causes, but can *automate aspects of operator and engineering workflows*

- Near-term deployment focus should be on algorithms that-
  - Better filter or highlight information
  - Augment operator memory and knowledge-retrieval, reduce cognitive load
  - Develop trust between humans and tools
- Statistical-feature based methods may be well-suited for such applications. These may also generate good labeled data that can train more sophisticated ML models.

Some Closing Thoughts

Awardee recommendations:

- Fostering discussions in working groups on standardized labeling practices
- Catalogue of data quality signatures
- Data storage formats for high computation efficiency and low memory requirement
- Incorporating algorithms within existing WAMS software platforms
- Utilizing the large repository of events to refine methodologies in existing WAMS platforms
- Other low-hanging fruits: Gen. trip classifiers/ringdown detectors for automatically exporting interesting events to event analysis/model tools, educational tools

What Next?

- Big data visualization tools
  - Effectively designed interfaces are critical for collaboration between humans and algorithms
  - Large amount of complex information must be conveyed
  - ***Must not*** increase cognitive burden on operators
- Creation of *golden datasets*
  - Labeled and validated real datasets will help benchmark the performance of proposed algorithms
  - An open-source signature library being compiled by PNNL-ORNL
- Feedback learning/ML as recommender systems
  - Human users can confirm/flag ML predictions, enabling learning in deployment

Other Useful Links

- Awardee reports and presentations will be listed here: https://www.energy.gov/oe/big-data-synchrophasor-analysis

Thank you

shuchismita.biswas@pnnl.gov

PNNL-SA-178557