Power Quality Data Analytics:
A new world of applications

Walmir Freitas
University of Campinas – UNICAMP
http://www.dsee.fee.unicamp.br/~walmir

ANEEL - November 2022

Acknowledgement
Part of these slides are based on a series of presentations done by Prof. Wilsun Xu, University of Alberta, the visionary and former chair of the IEEE/PES Working Group on Power Quality Data Analytics
Agenda

✓ Power Quality (Disturbance) Data Analytics: definition

✓ Potential data sources

✓ Data analytics

✓ Potential applications: smart meters – low resolution data

✓ Potential applications: power quality meter – high resolution data

✓ Final comments
Power quality data analytics (or simply Power Disturbance Analytics) is the discipline specialized in collecting measurement-based power system data, extracting information from it, and applying the findings to solve several power system problems such as:

✓ Power quality
✓ Power system protection
✓ Equipment condition monitoring
✓ System condition monitoring
✓ Active risk-based asset management

IEEE/PES has recognized the relevance of this emerging area by establishing the Working Group on Power Quality Data Analytics, which reports to the IEEE/PES Power Quality Subcommittee (active since 2013)

Walmir Freitas:
✓ One of the Working Group founders
✓ Chair: 2018-2020
✓ Vice-chair: 2016-2018
✓ Secretary: 2013-2016
✓ Received, as chair, the IEEE PES T&D Committee Award for Outstanding Technical Report – 2020

Adapted from: IEEE Working Group on Power Quality Data Analytics (http://grouper.ieee.org/groups/td/pq/data/)
Power Disturbance Data Analytics: components

- Power Disturbance Analytics
  - Data collection
    - Individual monitors
    - Networked monitors
  - Data analytics
    - General purpose algorithms
    - Application specific algorithms
Data science is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge (information) from noisy, structured and unstructured data, and apply it to a broad range of domains. Data science is related to data mining, machine learning, big data, computational statistics and analytics (Adapted from Wikipedia)
Potential data sources

- AMI: E, P, Q, V and I data (billing and demand monitoring)
- SCADA: 60 Hz magnitude (P, Q, V and I) data
- PMU: voltage and current phasors
- PQ monitors: rms and event-triggered waveform data
- Waveform measurement units (WMU): gapless voltage and current waveforms (synchronized or unsynchronized)
- Other potential data sources:
  - Modern relays – mission critical, hard to access data
  - Digital fault recorders – specialized for fault recording
  - Condition monitors – specialized/customized devices
Potential applications

Many others have not been envisioned yet!

- Power Disturbance Analytics
  - System-based
    - Fault
      - Anticipation
      - Location
      - Adaptive reclosing
      - Impedance estimation
      - Secondary fault detection
    - Resonance
      - Harmonic resonance estimation
  - Component-based
    - NTL detection
    - Network equipment
      - Parameter estimation
        - Capacitor banks
        - Switches
        - Transformers
      - Lines
        - Transformers
      - Load monitoring
        - Transformer load modeling
  - Customer load
    - Load modeling

From individual devices to wide-area systems

From generators to refrigerators
Sample of potential applications

✓ Potential Applications (Smart Meter – low resolution data):
  ▪ Automated management of GIS/asset and other related databases (BDGD)
  ▪ Non-technical loss detection and location
  ▪ Technical loss management and evaluation
  ▪ Fault location
  ▪ Load modeling
  ▪ Customer load disaggregation
  ▪ DER hosting capacity

✓ Potential Applications (Power Quality Meter – high resolution data):
  ▪ Resonance detection and mitigation (wind and solar parks)
  ▪ Fault anticipation
  ▪ Detection and location of high impedance faults
Automated management of GIS (BDGD) and others related databases

**Issue:** Utilities GIS/Assets database presents errors and inconsistencies due to:
- ✓ wrong data registration
- ✓ absence of data or update
- ✓ Line/transformer parameter variations due to weather conditions and equipment aging
- ✓ manual procedures for database update from field crew

**Relevance:** these databases are the core for:
- ✓ technical decisions
- ✓ economic decisions
- ✓ regulatory decisions

**Idea:** Combine customer smart meter data and data analytics to automatically correct:
- ✓ MV and LV system topology
- ✓ line and transformer parameters
- ✓ customers phase connection
- ✓ status of switches
- ✓ regulators/compensators settings and parameters
GIS automated correction (BDGD): LV systems

Issue: how to correct
✓ system topology
✓ customers phase connection
✓ line parameters

Idea: use multiple linear regression and data from customers smart meters:

\[
V_{an1} - V_{an2} = \begin{bmatrix}
1 - I_{R1}^T X_{1} \eta_1 \\
X_{2} \eta_1 
\end{bmatrix} + \begin{bmatrix}
\beta_0 \\
R_1^T X_{1} + R_2^T X_{2} 
\end{bmatrix}
\]

\[
R_1^+ = \beta_3 - (\beta_2 \text{ or } \beta_3) \approx R_{an1} - (R_{ab1} \text{ or } R_{ac1})
\]

\[
X_1^+ = \beta_2 - (\beta_2 \text{ or } \beta_3) \approx X_{an1} - (X_{ab1} \text{ or } X_{ac1})
\]

\[
R_{n1} = (\beta_2 \text{ or } \beta_3) \approx R_{nn1} - R_{an1}
\]

Topology is built from the bottom (customers) up (transformer) approach to the transformer, by pairing meters (real meters and virtual meters)

GIS automated correction (BDGD): LV systems

Real case: MV/LV systems: 2,175 buses; 2,000+ customers (87% residential); 76 MV/LV transformers

Low resolution:
- Metering error: 1.0%
- Measurement desynchronization: 10 sec
- 30 days of sample size

High resolution
- Metering error: 0.2%
- Measurement desynchronization: 0 sec
- 30 days of sample size

<table>
<thead>
<tr>
<th>Metric</th>
<th>High-Precision Scenario (%)</th>
<th>Low-Precision Scenario (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution (min)</td>
<td>15 30 60</td>
<td>15 30 60</td>
</tr>
<tr>
<td>Branch</td>
<td>92 92 91</td>
<td>33 48 58</td>
</tr>
<tr>
<td>Line length</td>
<td>87 88 87</td>
<td>31 41 49</td>
</tr>
<tr>
<td>Phasing</td>
<td>100 100 100</td>
<td>100 100 100</td>
</tr>
</tbody>
</table>

High success rate
Phasing estimation still has high success rate

Line parameters successfully estimated (more than 90% of the parameters estimated accurately)

**Issue:** Utilities GIS Database presents errors, missing data, and inconsistencies on MV systems regarding to:

- Tap position of service transformers (MV/LV transformers)
- Status of switches

**Idea:** Combine customer smart meter data and a generalized state estimation formulation to correct this equipment

---

GIS correction: tap position of service (MV/LV) transformers and status of switches

Real case: MV/LV system with 5,000+ customers (87% residential) and 190 MV/LV transformers

Status of switches: 100% accurate – 24 hours of operation

<table>
<thead>
<tr>
<th>Bus/Switch</th>
<th>Phase</th>
<th>$r_N^P$</th>
<th>$r_N^Q$</th>
<th>$r_N^V$</th>
<th>$r_N^\theta$</th>
<th>Iteration</th>
</tr>
</thead>
<tbody>
<tr>
<td>2794</td>
<td>c</td>
<td>12.52</td>
<td>12.56</td>
<td>-12.65</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>S2</td>
<td>a</td>
<td>0.22</td>
<td>-5.07</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>S3</td>
<td>b</td>
<td>-</td>
<td>-</td>
<td>8.59</td>
<td>-0.19</td>
<td>3</td>
</tr>
<tr>
<td>S1</td>
<td>c</td>
<td>-</td>
<td>-</td>
<td>-0.01</td>
<td>-1.53</td>
<td>4</td>
</tr>
</tbody>
</table>

MV/LV transformer tap: 100% accurate

The method is robust against:
- ✓ Meter Errors (precision class and clock desynchronization of meters)
- ✓ Gross Errors (e.g., incorrect power measurements)
- ✓ Switch Errors (incorrect status of switches)

**Issue:** Control settings of voltage regulators and capacitor banks are constantly updated on field, but this information is often not updated on the database.

**Idea:** Combine customer smart meter data and a generalized state estimation formulation to estimate the physical status of capacitor banks and voltage regulators, and the control settings of this equipment.

**Voltage regulator: what is to be estimated**

- Control settings:
  - $V_{\text{sup}} (V)$
  - $V_{\text{inf}} (V)$
  - $V_{\text{ref}} (V)$
  - $B (V)$

**Capacitor bank: what is to be estimated**

- Control mode:
  - reactive power
  - power factor
  - time
  - current
- Control settings:
  - $Q_{\text{ON}}$
  - $Q_{\text{off}}$
- Operation period (status):
  - $ON$
  - $OFF$

Real case 1: Why estimate the control settings of voltage regulators?

Real case 2: Existence of a 1.2 MVA capacitor bank from a disactivated industry without the knowledge by the utility.
GIS correction: estimation of physical status and control settings of voltage regulators and capacitor banks

Real case: MV/LV system with 5,000+ customers (87% residential) and 190 MV/LV transformers

<table>
<thead>
<tr>
<th>Setting</th>
<th>Original</th>
<th>Resolution (15 min)</th>
<th>Resolution (60 min)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Days 1</td>
<td>Days 3</td>
</tr>
<tr>
<td>$V_{sup} (V)$</td>
<td>11,775</td>
<td>11,787</td>
<td>11,791</td>
</tr>
<tr>
<td>$V_{inf} (V)$</td>
<td>11,625</td>
<td>11,609</td>
<td>11,617</td>
</tr>
<tr>
<td>$V_{ref} (V)$</td>
<td>11,700</td>
<td>11,698</td>
<td>11,704</td>
</tr>
<tr>
<td>$B (V)$</td>
<td>1.5</td>
<td>1.77</td>
<td>1.73</td>
</tr>
<tr>
<td>$Q_{ON} (kvar)$</td>
<td>1,200</td>
<td>1.196&lt;Q&lt;1.208</td>
<td>1.202&lt;Q&lt;1.189</td>
</tr>
<tr>
<td>$Q_{OFF} (kvar)$</td>
<td>800</td>
<td>801</td>
<td>805</td>
</tr>
</tbody>
</table>

The control settings and operation period are properly estimated

Non-Technical Losses: detection and location (Idea 1)

✓ Issue: Illegal load connection (NTL) tampers active and reactive power measurements, but no voltage measurement. **Power flows, voltage does not**

✓ Idea 1: By using P, Q and V measured by the **smart meter** in each customer, one can estimate the $V_{PCC}$. The lowest estimated value indicates potential NTL

---

**Illegal load connection**  
P & Q: tampered data  
V: correct data

$\hat{V}_{PCCI} = V_{mi} + Z_{Si} \cdot I_i$

---

**Volt drop method:**  
$V_{PCC}$ is estimated by using data from each customer smart meter connected to the same MV/LV transformer. Estimated values different (lower) indicated NTL

Non-Technical Losses: detection and location (Idea 2)

**Issue:** Energy theft by the connection of irregular loads

**Idea 2:** Use data from **customer smart meters** to run a **state estimation process**, as active and reactive power measurements are tampered, but voltage is not

Method can typically detect and locate NTL as small as 2 kW for LV illegal loads and 23 kW for MV illegal loads

Non-Technical Losses: detection and location (Idea 2)

Case study:
✓ 13.8-kV feeder (real): 55 LV systems + 64 MV customers – 1,682 buses.
✓ LV–NTL: magnitude from 1 kW to 10 kW (~1,000 occurrences)
✓ MV–NTL: magnitude from 20 kW to 200 kW (56 occurrences)

### LV–NTL

<table>
<thead>
<tr>
<th>NTL (kW)</th>
<th>Successful Cases (%)</th>
<th>e) $N_{\text{NTL}}$ (percentile)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a)</td>
<td>b)</td>
</tr>
<tr>
<td>1</td>
<td>24.8</td>
<td>19.6</td>
</tr>
<tr>
<td>2</td>
<td>86.9</td>
<td>71.8</td>
</tr>
<tr>
<td>3</td>
<td>97.3</td>
<td>88.3</td>
</tr>
<tr>
<td>4</td>
<td>99.3</td>
<td>95.3</td>
</tr>
<tr>
<td>5</td>
<td>99.3</td>
<td>97.7</td>
</tr>
<tr>
<td>6</td>
<td>99.5</td>
<td>98.6</td>
</tr>
<tr>
<td>7</td>
<td>99.5</td>
<td>98.9</td>
</tr>
<tr>
<td>8</td>
<td>99.5</td>
<td>99.1</td>
</tr>
<tr>
<td>9</td>
<td>99.5</td>
<td>99.1</td>
</tr>
<tr>
<td>10</td>
<td>99.5</td>
<td>99.1</td>
</tr>
</tbody>
</table>

### MV–NTL

<table>
<thead>
<tr>
<th>NTL (kW)</th>
<th>Successful Cases (%)</th>
<th>e) $N_{\text{NTL}}$ (percentile)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a)</td>
<td>b)</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>21</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>22</td>
<td>39.6</td>
<td>39.6</td>
</tr>
<tr>
<td>23</td>
<td>96.9</td>
<td>96.9</td>
</tr>
<tr>
<td>24/25/50/100/150</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

a) NTL is detected; b) NTL bus is among the suspect buses; c) NTL bus is indicated with the maximum $\text{Err}_{\text{NTL}}$(%) value; d) NTL bus or a first neighbor bus is indicated with the maximum $\text{Err}_{\text{NTL}}$(%) value; e) Number of buses indicated as suspects of NTL.

Fault location: distribution systems (idea 1)

Idea 1:
✓ Collect V&I at feeder terminal
✓ Calculate downstream Z using V&I
✓ Estimate fault distance using Z
✓ PQ monitor is used to collect data

Basic idea of impedance-based fault location technique

Examples of Algorithms for Single-Phase Fault Location

• Positive-Sequence and Zero-Sequence
  – Loop Impedance ($Z_L$)
  – Loop Resistance ($R_L$)
  – Loop Reactance ($X_L$)

• Positive-Sequence Algorithms
  – Resistance-to-Fault (RTF)
  – Impedance-to-Fault (ZTF)
  – Reactance-to-Fault (XTF)

• RMS Voltage and RMS Current Only
  – Absolute Impedance (Z)
Fault location: distribution systems (idea 1)

Case:

DTE Energy - Detroit Edison (DECo)

PQ monitor installation at HV/MV substation for fault location - © 2012 IEEE

Single-line-to-ground fault measured by the PQ monitor - © 2012 IEEE

Example Street View Map of estimated (green lightning) and actual (red lightning) fault location - © 2012 IEEE

Fault location: distribution systems (idea 2)

Issue:
✓ How to avoid identification of multiple locations?

Idea 2:
Use information from **customer smart meters**:
✓ Outage mapping (might not be sufficient)
✓ Voltage magnitude (concept of low voltage zone)

Functionalities:
✓ Outage mapping: simpler and lower accuracy
✓ Voltage measurement: more complex and higher accuracy

Fault location: distribution systems (idea 2)

Results of fault location for a **single-phase** fault with $R_f = 0.5 \, \Omega$.

<table>
<thead>
<tr>
<th>Estimated fault location</th>
<th>Voltage magnitude (pu)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.171</td>
</tr>
<tr>
<td>2</td>
<td>0.236</td>
</tr>
<tr>
<td>3</td>
<td>0.365</td>
</tr>
<tr>
<td>4</td>
<td>0.365</td>
</tr>
<tr>
<td>5</td>
<td>0.363</td>
</tr>
<tr>
<td>6</td>
<td>0.365</td>
</tr>
<tr>
<td>7</td>
<td>0.365</td>
</tr>
</tbody>
</table>

Load modeling: what are to be modeled?

Load parameters for loads with the above responses:

- Transient load responses: $\Delta P_t$ and $\Delta Q_t$
- **Steady-state load responses:** $\Delta P_s$ and $\Delta Q_s$
- Time to recover: $\tau_P$ and $\tau_Q$

Why is it important to correctly model the steady-state load responses?

✓ Determination of technical and non-technical losses
✓ Allocation of capacitor banks and voltage regulators
✓ Decision-making of strategies for voltage regulation and var compensation
✓ Ampacity calculations
✓ Expansion studies

Load modeling: parameter estimation

Which model should be used?
Which signals should be monitored?
How to automatically detect (select) a voltage disturbance useful for load monitoring? (Upstream versus downstream disturbance)
Which level of voltage variation should be detected?
How large should the measurement window be?
Is the number of events enough to be representative?

\[ P = P_0 \left( \frac{V}{V_0} \right)^{np} \quad \text{and} \quad Q = Q_0 \left( \frac{V}{V_0} \right)^{nq} \]

Exponential model

Case study: In a pilot project, measurements were carried out for 3 and 1/2 days (82 consecutive hours)

Issue: How to use V&I current from a single sensor to monitor individual home appliances consumption?

Idea: Develop an event-window-based approach using unique characteristics (signatures) of typical appliances such as:

- edge signatures
- sequence signatures
- trend signatures
- time/duration signatures
- phase signatures
- power signatures
- harmonic signatures

Appliance characteristics (signatures)

Characteristic 1: Power levels
- A microwave oven draws about 1000W when turned on
- A fridge draws about 100W when turned on

Characteristic 2: Current waveforms

Characteristic 3: Turn on transients and operating cycles

Characteristic 4: “Electrical location”

Characteristic 5: Duration and time of use

Issue: interactions among the wind farms and the system capacitances and inductances can produce sub-synchronous resonances (series compensation) and harmonic resonances (shunt compensation), which can be weakly damped or unstable (high frequency)
Wind generation: resonances – real cases

✓ Texas, USA: installed capacity: 200 MW (345 kV, Type III generators)
✓ Unstable resonance was caused by a short-circuit, followed by a transmission line being tripped close to the wind park
✓ Sub-synchronous currents reached 4.0 pu in 1 s
✓ Sub-synchronous voltages reached 2.0 pu in 3 s
✓ The event damaged the crowbar circuit of several wind park generators, and the series capacitor of a transmission line close to the wind park

✓ Heibei, China: installed capacity: 3.4 GW (220 kV, 82.5% - Type III, 15.4% - Type IV and 1.8% Type II generators)
✓ 58 events of unstable sub-synchronous resonance were detected from Dec. 2012 to Dec. 2013
✓ Event (Mar. 19th, 2013): power generation was 219.5 MW. 30 s after the start, the oscillation magnitude reached 25% of the average power generation. A total of 66% of the generation was lost during the event


Wind generation: protective methods for mitigation (real-time)

Issue:

SSR events in series compensated transmission lines connected close to DFIG-based wind farms have been reported.

SSR currents can become significant in less than 1 second, causing equipment damages.

Early detection of SSR characteristics is critical to avoid equipment damages and implement mitigation actions.

Challenge: obtain SSR current, with unknown frequency, with high speed and accuracy.

Wind generation: protective methods for mitigation (real-time)

Idea: voltage and current waveforms at the line terminals can be collected to extract the sub-synchronous current by using the line as a natural analog filter.

SSR characteristics can be obtained in ~1 SSR cycle.

Synchronized waveforms

\[ \Delta V_{\text{line}}(t) - \Delta V_{\text{aux}}(t) = A I_{1\omega} e^{-at} \cos(2\pi f_{1\omega} t + \gamma_{1\omega} + \theta) \]

Extracted sub-synchronous component

Depend on line parameters

Incipient fault detection (fault anticipation)

**Issue:** Several faults in distribution systems are preceded by incipient faults, especially if the faults are related to equipment failures:

- overgrown trees under power lines
- insulation failure
- failure of transformer tap

Detection of incipient faults allows the adoption of predictive actions, avoiding the occurrence of a permanent fault

**Characteristics of incipient faults:**

- Small magnitude not enough to trigger relays
- Short duration not enough to trigger relays
- Distorted waveforms

**Idea:** detect abnormal voltage and current waveforms (PQ monitors)

**Concept of incipient fault detection (or fault anticipation)**

Incipient fault detection (fault anticipation)

**Real case 1:** Fault-induced conductor slap (FICS): occurs when magnetic forces from an initial fault cause movement in upstream conductors sufficient to cause contacts, resulting in a second, higher magnitude fault closer to the substation.

RMS current waveforms for FICS sequence - © 2015 IEEE

Span with multiple FICS events - © 2015 IEEE

1. Failure signatures of equipment
   - Underground cables
   - Overhead lines
   - Transformers & tap changers
   - Switches
   - Capacitors
   - Lightning and surge arresters
   - Potential transformers

2. Review of waveform abnormality detection methods

3. Discussions on how to move forward

https://resourcecenter.ieee-pes.org/technical-publications/technical-reports/PES_TP_TR73_TD_122019.html

IEEE PES T&D Committee Award for Outstanding Technical Report: “For Advancing the Power Quality Data Analytics Domain by Demonstrating Techniques for Prediction and Analysis of Electric Power Equipment Failure” – 2020
Due to uncertainties, variabilities and unpredictability of demand and generation, economic and environmental concerns, the electrical energy systems of the future will be planned and operated based more and more on risk-based methods, stochastic approaches and active (predictive) philosophies.

Data analytics will be essential for the future of the electrical energy systems.

Smart meters (and other sensors) are one of the core technologies to promote this paradigm change (more killer applications can make this solution a business case).

Models and methods must be developed considering the availability and quality of data.

Worse than make a decision with no data, it is to make a decision with bad data.
Thank you

Walmir Freitas
http://www.dsee.fee.unicamp.br/~walmir