



An Exelon Company



2021

Grid of the Future

OCTOBER 17-20, 2021

# Anomalous Voltage Data Detection: Utility Experience from Large Scale CVR Deployment

Md Shakawat Hossan, Wen Fan, Anthony Cook,  
Sepideh Fard, Aleksi Paaso

ComEd

Amin Khodaei

PLUG LLC

- ✓ Introduction
- ✓ Overview of the measurement terminologies
- ✓ Detection of data associated with operational concerns
- ✓ Case studies
- ✓ Discussions and conclusions

- Benefits from CVR deployment are captured by an index named CVR factor (CVR<sub>f</sub>)

$$CVR_f = \frac{\% \Delta E}{\% \Delta V}$$

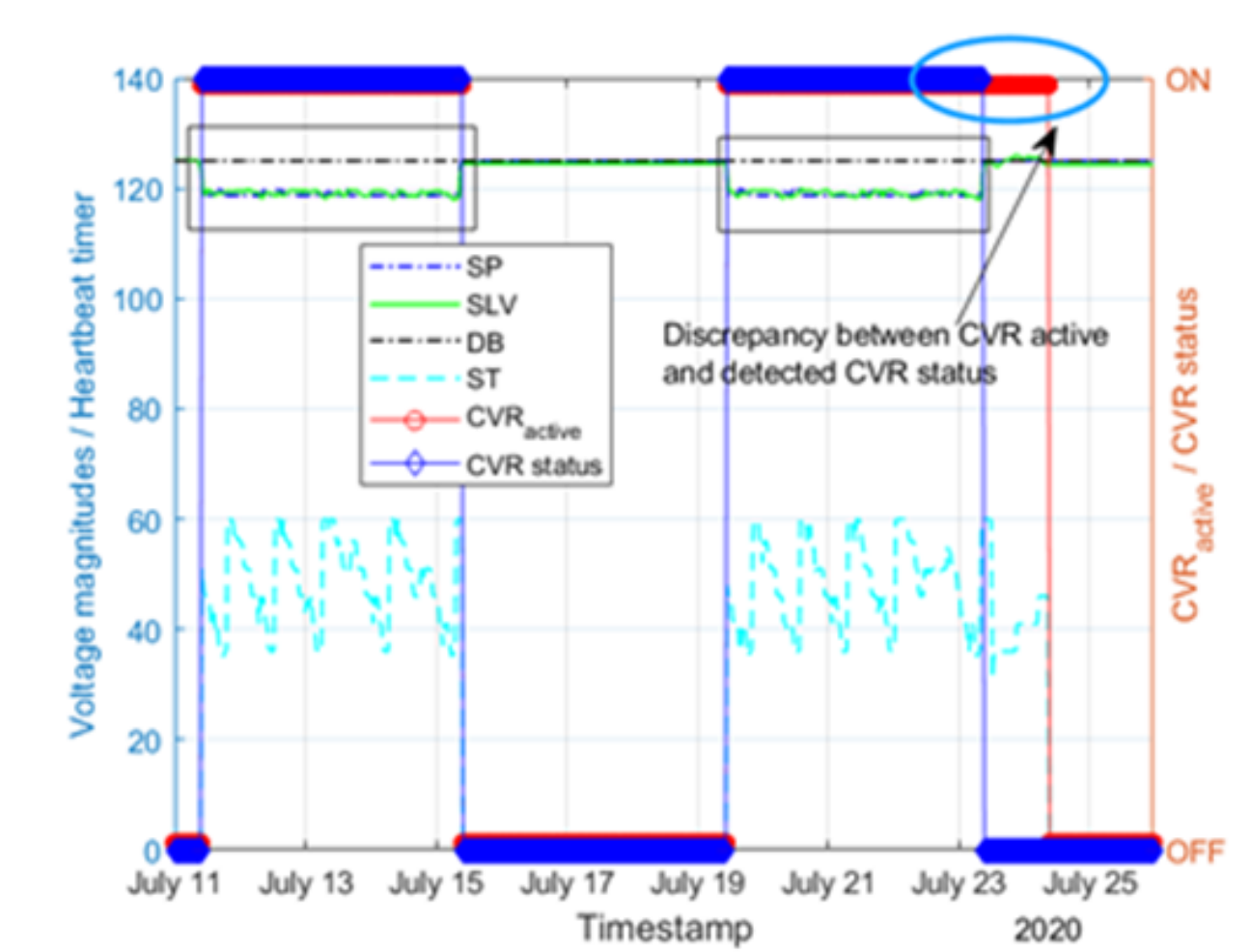
- Energy savings for an individual feeder is calculated by

$$E_{savings} = E_{baseline} \times \Delta V \times CVR_f$$

- Hence, the voltage reduction  $\Delta V$  plays a significant role in determining the CVR factor and the energy savings
- This study discusses several data issues that need to be identified and excluded for estimating  $\Delta V$  using time-series voltage data
  - Phase imbalance
  - Invalid CVR data
  - Primary and secondary measurement lagging
  - Voltage Spike(s)

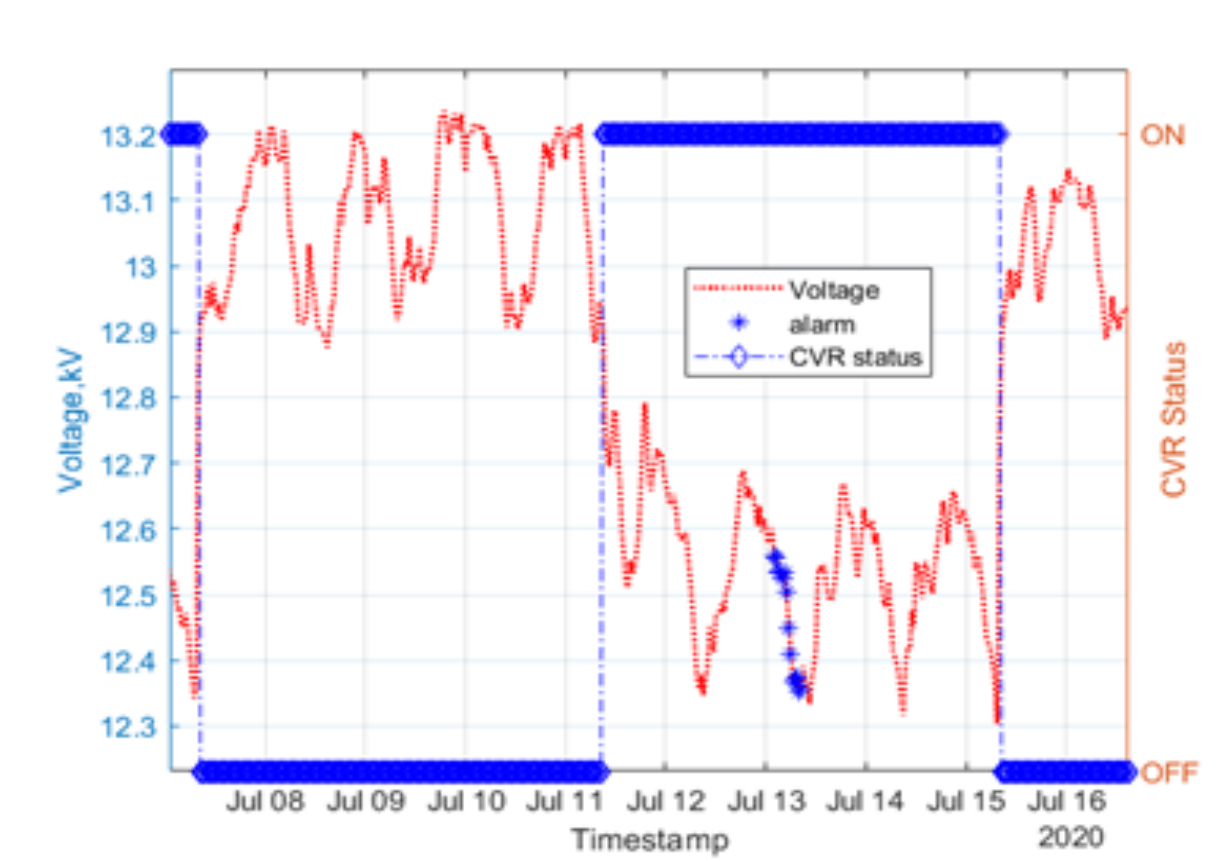
# Overview of the Measurement Terminologies

- CVR status



# Overview of the Measurement Terminologies

- LOLM Alarm



- In some cases, it is required for CVR to be activated on each individual phase
- This could cause a scenario where one or more phases won't operate at the correct SP due to backend calculation or communication issue
- Voltage difference between phases are calculated and checked if any phase has any pre-determined deviation compared to the other two phases based on the CVR status

---

## Algorithm 1: Detection of Phase Imbalance

---

- 1 Define the time range of data cleaning as  $T_{Range}$ ;
  - 2 Extract the three phase voltage data within  $T_{Range}$  and store as  $V_{T_{Range},A}$ ,  $V_{T_{Range},B}$ , and  $V_{T_{Range},C}$  for phases A,B, and C, respectively;
  - 3 Estimate the phase to phase difference as
 
$$Diff_{Ph_1,Ph_2} \rightarrow \frac{V_{T_{Range},Ph_1} - V_{T_{Range},Ph_2}}{V_{T_{Range},Ph_1}}, \quad Ph_1, Ph_2 = A, B, C;$$
  - 4 Accumulate the differences together as  $Diff_{Total} = [Diff_{AtoB}, Diff_{AtoC}, Diff_{BtoC}]$ ;
  - 5 Find the global indices of time instants where two of the differences are greater than a pre-determined value  $k$  in percentage and store in  $T_{diff}$  as  $T_{diff} = any[Diff_{Total} > k, 2]$ ;
  - 6 Find the flags per phase based on the global indices, CVR status, and boundary of the voltage as  $T_{flag,Ph}$  in  $T_{flag,Ph}(CVR_{status} == 0) = [V_{T_{Range},Ph} < \ddot{z} \ \& \ V_{T_{Range},Ph} > 0.9 * nominalV \ \& \ T_{diff} == 1]$ ; and  $T_{flag,Ph}(CVR_{status} == 1) = [V_{T_{Range},Ph} > \ddot{z} \ \& \ V_{T_{Range},Ph} > 1.1 * nominalV \ \& \ T_{diff} == 1]$ , Where  $Ph = A, B, C$ ;
  - 7 Accumulate the flags for both CVR OFF and ON instants as  $T_{anom_{imb},Ph} = [T_{flag,Ph}(CVR_{status} == 0), T_{flag,Ph}(CVR_{status} == 1)]$  Where  $Ph = A, B, C$ ;
  - 8 Declare phase imbalance related time instants,  $T_{anom_{imb},Ph}$  as  $NaN \ V_{T_{Range}}[T_{anom_{imb},Ph}] = NaN$ ;
-

## Detection of Invalid CVR Data

- Another scenario is that LOLM alarm is active while CVR is OFF, mostly being seen on feeders undergoing CVR ON/OFF testing
- The SP is not able to go back to the DB due to any software, communication, or equipment anomaly that disables the LTC to restore to default bandcenter
  - Resulting neither CVR ON nor CVR OFF voltages

---

### Algorithm 2: Detection of anomalous instants due to LOLM

---

- 1 Define the time range of data cleaning as  $T_{Range}$ ;
  - 2 Extract the voltage data within  $T_{Range}$  and store as  $V_{T_{Range}}$ ;
  - 3 Find the CVR OFF time instants within  $T_{Range}$  and store as  $T_{CVR_{OFF}}$ ;
  - 4 Find the time instants within  $T_{Range}$  where LOLM alarm is active and store as  $T_{LOLM}$ ;
  - 5 Find the time instants where CVR OFF operating voltage is lower than a pre-determined value  $\bar{z}$  and store as  $T_{Low}$ ;
  - 6 Find the common time instants and store as  $T_{bad}$  where  $T_{anom_{lolm}} = T_{range} \cap T_{CVR_{OFF}} \cap T_{LOLM} \cap T_{Low}$ ;
  - 7 Declare lolm related time instants,  $T_{anom_{lolm}}$  as NaN;  
 $V_{T_{Range}}[T_{anom_{lolm}}] = NaN$ ;
-

- Measurement collection may have certain lag due to issues related to settings of communication channels and their frequency
  - This may provide some discrepancies on how these data can be interpreted
- For example, transformer primary and secondary side measurements lagging between the timestamp and their data collection

---

**Algorithm 3:** Anomalous data detection due to time lag between Secondary and Primary Voltage

---

- 1 Define the time range of data cleaning as  $T_{Range}$ ;
  - 2 Extract the voltage data within  $T_{Range}$  and store as  $V_{T_{Range}}$ ;
  - 3 Find the moving mean and standard deviation based on the CVR status:
  - 4 CVR OFF and ON moving mean are calculated over  $K$  time horizon as, respectively  $V_{MM}(CVR_{status} == 0) = Movemean(V_{T_{Range}}(CVR_{status} == 0), k)$  and  $V_{MM}(CVR_{status} == 1) = Movemean(V_{T_{Range}}(CVR_{status} == 1), k)$ ;
  - 5 Find the indices of time instants where secondary and primary voltage are not aligned based on moving mean, CVR OFF status, and ANSI boundary  
 $T_{anom\_secdiff}(CVR_{status} == 0) = abs(V_{T_{Range}}(CVR_{status} == 0) - V_{MM}(CVR_{status} == 0)) > abs(V_{T_{Range}}(CVR_{status} == 0) - mean(V_{T_{Range}}(CVR_{status} == 1))) \& (V_{T_{Range}} < 1.1 * nominalV | V_{T_{Range}} > 0.9 * nominalV)$ ;
  - 6 Find the indices of time instants where secondary and primary voltage are not aligned based on moving mean, CVR ON status, and ANSI boundary  
 $T_{anom\_secdiff}(CVR_{status} == 1) = abs(V_{T_{Range}}(CVR_{status} == 1) - V_{MM}(CVR_{status} == 1)) > abs(V_{T_{Range}}(CVR_{status} == 1) - mean(V_{T_{Range}}(CVR_{status} == 0))) \& (V_{T_{Range}} < 1.1 * nominalV | V_{T_{Range}} > 0.9 * nominalV)$ ;
  - 7 Declare associated time instants where voltage difference between primary and secondary are observed,  $T_{anom\_secdiff}$  as NaN  $V_{T_{Range}}[T_{anom\_secdiff}] = NaN$ ;
-

## Detection of Voltage Spike(s)

- Voltage spikes in SCADA data is very common and are not very straightforward to be pinpointed whether they are true outliers
- Since voltage spikes are usually not that extreme similar to outliers, simply using the 3 standard deviation may not detect the spikes
- Instead, if the mean and standard deviation are utilized with a rolling horizon, these spikes are easily detectable

---

### Algorithm 4: Detection of Voltage Spike

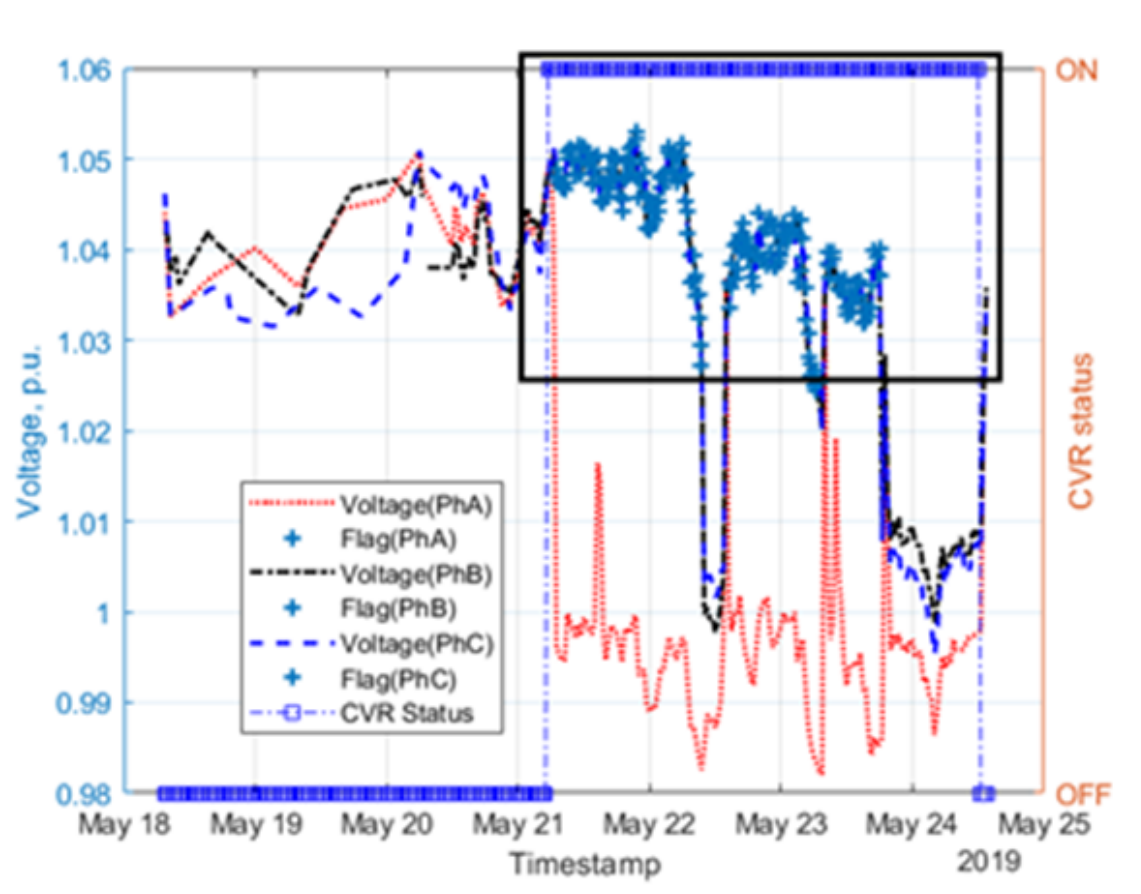
---

- 1 Define the time range of data cleaning as  $T_{Range}$ ;
  - 2 Extract the voltage data within  $T_{Range}$  and store as  $V_{T_{Range}}$ ;
  - 3 Find the moving mean and standard deviation based on the CVR status:
    - 4 CVR OFF and ON moving mean are calculated over  $K$  time horizon as, respectively  $V_{MM}(CVR_{status} == 0) = Movemean[V_{T_{Range}}(CVR_{status} == 0), k]$  and  $V_{MM}(CVR_{status} == 1) = Movemean[V_{T_{Range}}(CVR_{status} == 1), k]$ ;
    - 5 CVR OFF and ON moving std deviation are calculated over  $K$  time horizon as, respectively  $V_{MStd}(CVR_{status} == 0) = Movestd[V_{T_{Range}}(CVR_{status} == 0), k]$  and  $V_{MStd}(CVR_{status} == 1) = Movestd[V_{T_{Range}}(CVR_{status} == 1), k]$ ;
  - 6 Find the indices of time instants where voltage spike occurred based on moving mean, standard deviation, CVR status, and ANSI boundary
 
$$T_{anom\_spike} = [V_{T_{Range}} > V_{MM} + 3 * V_{MStd}] \mid [V_{T_{Range}} < V_{MM} - 3 * V_{MStd}] \mid [V_{T_{Range}} > 1.1 * V_{MM}] \mid [V_{T_{Range}} < 0.9 * V_{MM}] \mid [V_{T_{Range}} < 1.1 * nominalV \mid V_{T_{Range}} > 0.9 * nominalV];$$
  - 7 Declare voltage spike associated time instants,  $T_{anom\_spike}$  as  $NaN$   $V_{T_{Range}}[T_{anom\_spike}] = NaN;$
-

## Detection of Data Associated with Phase Imbalance Concerns – Case Study

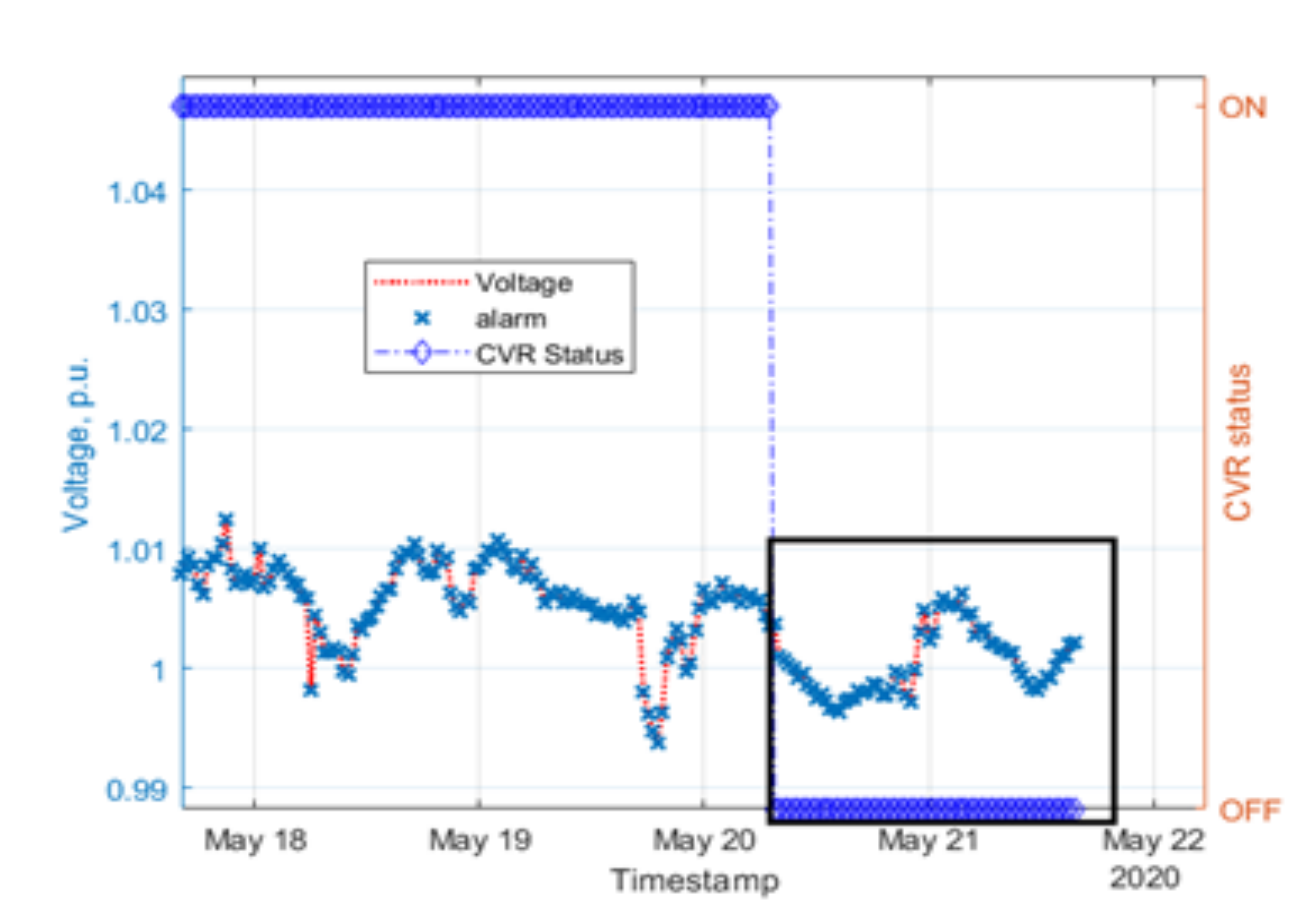
10

- Only Phase A voltage was lowered during CVR ON
- Phase B and C data within that period are detected as anomalous data



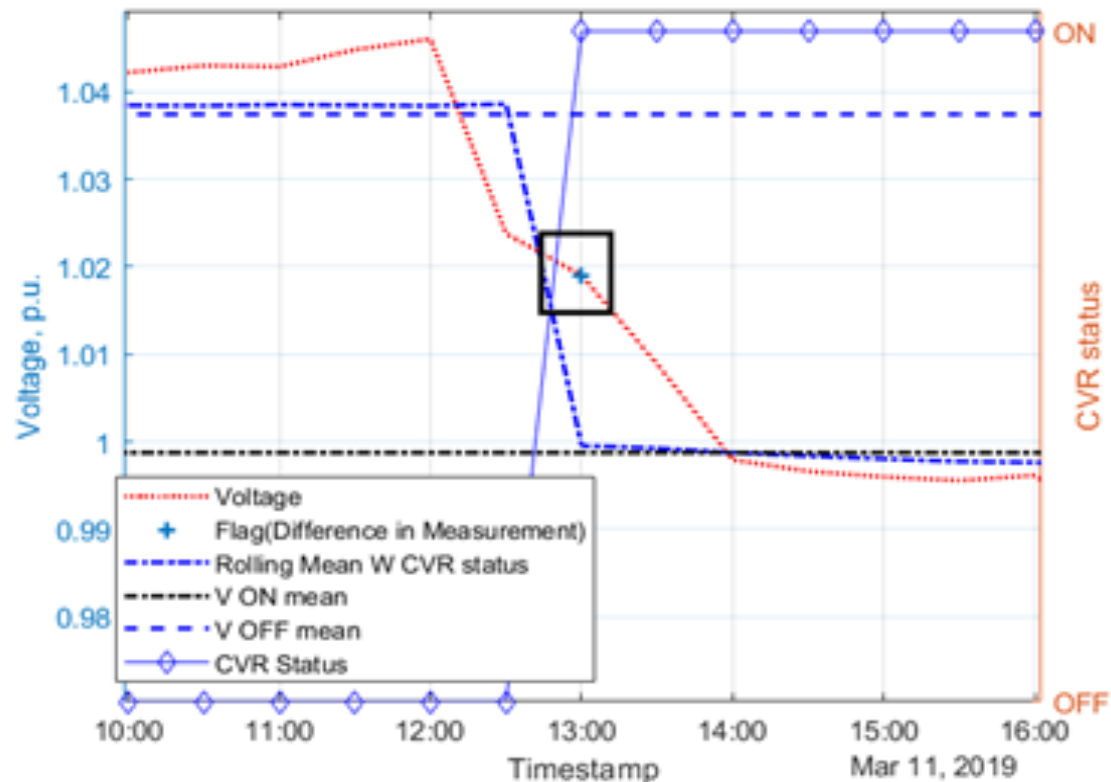
## Detection of Invalid CVR Data – Case Study

- While CVR is OFF, voltage is in CVR on range, and LOLM is active which clearly defines some anomaly and ensures that these data are opposite to the natural CVR operation



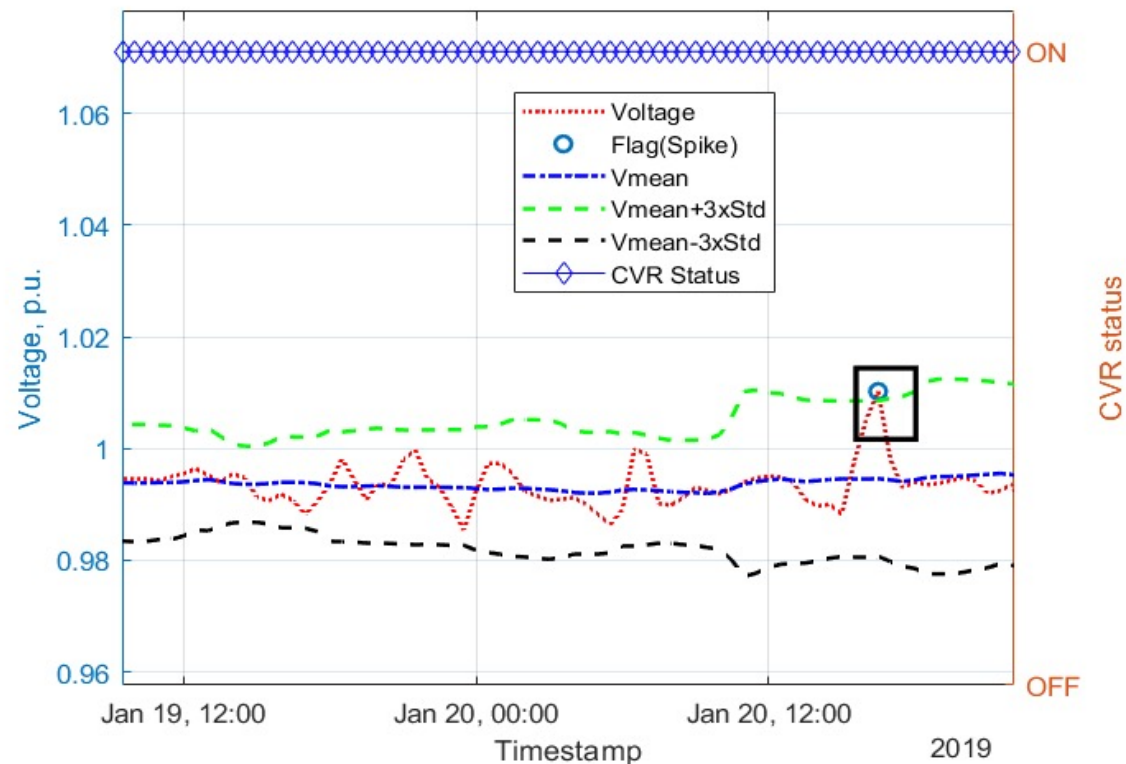
## Detection of Primary and Secondary Measurement Lagging - Case Study

- CVR status is determined based on the SLV and SP data collected from the secondary side
- The feeder-head voltage data is collected from the primary side



## Detection of Voltage Spike(s) – Case Study

- The check considers the ANSI range to ensure that it is not a typical outlier.
- Then, it checks whether the data lies with the 3-standard deviation of the rolling mean.



- An example of voltage reduction on three transformers where yearly dataset went through data cleaning using the developed algorithms

---

<b>Xfmr</b>	<b><math>\Delta V</math> w/o cleanin g (%)</b>	<b><math>\Delta V</math> w cleanin g (%)</b>	<b>Alg1(%)</b>	<b>Alg2(%)</b>	<b>Alg3(%)</b>	<b>Alg4(%)</b>
<b>TR1</b>	4.68	4.72	0.0	0.0	0.10	0.06
<b>TR2</b>	3.99	4.51	0.0	1.68	0.14	0.53
<b>TR3</b>	3.57	3.59	0.01	0.0	0.03	0.23

- This paper developed several algorithms to detect the anomaly data
- By excluding these anomalies, it can maintain the assessment accuracy related to CVR benefits.
- This is especially true if a utility is running a large scale CVR program where thousands of feeders are running on CVR
- The developed algorithms are flexible and can be adopted easily by utility engineers based on the system condition