Oscillation Modes Identification via SVD and PCA – CIGRE 2024 Paris Session Paper SC C4-11448

Presented by Dr Carlos Ferrandon

Presentation contents

- Introduction
- Test system utilised
- Data description
- Methodology
- Initial results
- Potential clustering approaches
- Future work

Introduction

- Power system oscillations can be expected within a large interconnected grid. Most of them are usually naturally damped.
- The sustained and poorly damped ones require our attention!
- The oscillation can be present in:
  - Frequency
  - Voltage
  - Current
  - MW
- Objective: to leverage on the available data from Phasor Measurement Units (PMUs) by utilising practical-and-ready dimensionality reduction techniques such as Principal Component Analysis (PCA) and Singular Value Decomposition (SVD), in order to determine the existence of the oscillatory phenomena.
- Oscillation Detection and Monitoring in real-time: out of scope… for now

Post-mortem tool?

- Pre-processing:
  - Bad data identification needed. Sometimes simple eye-screening checks can reveal bad data traces! Voltage traces for example
  - Use PMUs data quality checks
  - Special care with ambient data and internal PMUs clocks[1]

[1] Mishra, C., Vanfretti, L., Delaree, J., & Jones, K. D. (2024). Internal clock errors in synchrophasor ambient data: Effects, detection, and a posteriori estimation-based correction. International Journal of Electrical Power and Energy Systems, 161. https://doi.org/10.1016/j.ijepes.2024.110208

Post-mortem tool?

- Two dimensionality reduction techniques[2] are utilized for a matrix of observations $\mathbf{A}$:
- Singular Value Decomposition (SVD). Contains the singular values, representing the strength or impact of each mode.
- PCA (Principal Component Analysis) uses the foundations of SVD, with the difference of scaling the original data by shifting it to the origin and standardizing it, making it is visually easier to see representations of the projections from it.
- Both methods look to extract relevant information of the data observed and project it in the most significant representation, that although reduced, still captures the variability of them as much as possible.
- **Main strength: off-the-shelf, quick screening method**
- Although straightforward engineering judgement is still required

[2] . R. Messina, Wide area monitoring of interconnected power systems, 2nd Edtion. IET, 2022.

Test system and data available

- Wide area ISO-NE oscillation phenomena in 2017
- Assessment of available PMU data [3]
  - Voltage
  - Current
  - Frequency
- Data available in certain substations, but not in all of them.
- Sampling rate: 0.033 s, 30 samples per second.
- 350 seconds of data analysed.
- **Goal:** prove that the set of measurements carry a certain correlation, and that the cumulative variance can be explained a minimum of Principal Components.

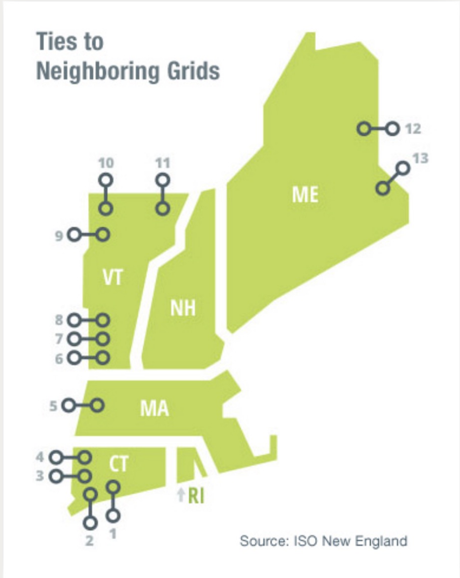


Figure 1 – ISO-NE map[4]

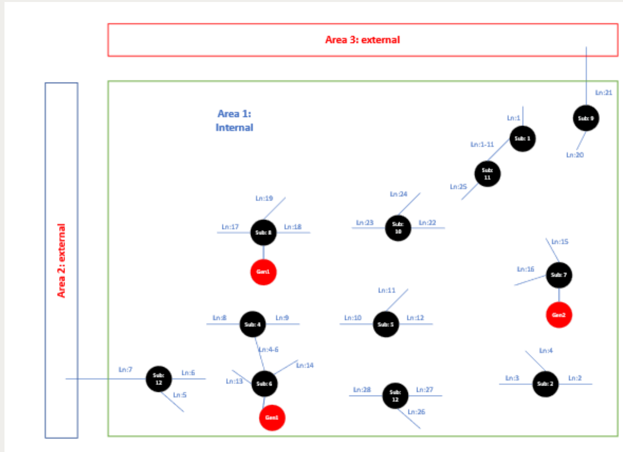


Figure 2 – ISO-NE observed substations

[3] . S. Maslennikov et al., “A Test Cases Library for Methods Locating the Sources of Sustained Oscillations,” in 2016 IEEE Power and Energy Society General Meeting (PESGM), Boston,MA, USA, 2016, pp. 1–5. doi: 10.1109/PESGM.2016.7741772.

[4] ISO New England, “ISO New England,” https://www.iso-ne.com/about/key-stats/maps-anddiagrams Accessed: July 18, 2024. [Online]. Available: https://www.iso-ne.com/about/keystats/maps-and-diagrams

Test system and data available

- Previously known event:

ISO-NE-case2_MASLENNIKOV_ISO-NE.txt

```
Case: ISO-NE_case2

-------Description----------
Power system: ISO New England is a North-East part of the Eastern Interconnection in the USA
Peak load is about 26,000 MW

Date of event: October 3, 2017

Type of oscillations: multi-frequency, wide-spread oscillations

Frequency: dominant modes 0.08Hz, 0.15Hz and 0.31Hz

Peak-to-peak magnitude: up to 130 MW

Is it a resonance condition case?: No

Location of the source: outside of ISO-NE; in Area 3, see map ISO-NE_map.pdf

Details of the source: large generator in Area 3

Confidence level on the known location of the source: 100%

Duration of time interval of submitted PMU data: 6 min

Comments:
An issue in the governor of a large generator outside of ISO-NE has created a multi-frequency process
with growing magnitude during 5 minutes. Oscillations with significat MW magnitude were observed in
multiple locations of the New England power system.
```

Figure 3 – Known data

Methodology and results

1) Observed data:

26 rows of **voltage** observatios for 10800 samples in time.

$$\mathbf{A} = \begin{pmatrix} v_{sub1(1)} v_{sub2(1)} \cdots v_{subm(1)} \\ v_{sub1(2)} v_{sub2(2)} \cdots v_{subm(2)} \\ \vdots \quad \vdots \quad \vdots \quad \vdots \\ v_{sub1(t)} v_{sub2(t)} \cdots v_{subm(t)} \end{pmatrix}$$

2) Mean-centered data

Mean-centered data of observations

$$\mathbf{B} = \mathbf{A} - \overline{\mathbf{A}}$$

Standardize each element of $\mathbf{B}$

$$Z_{ij} = \frac{B_{ij}}{\sigma_j}$$

Better comparability between variables is achieved, allowing to spot inputs with highest variability and potential anomalies in the data.

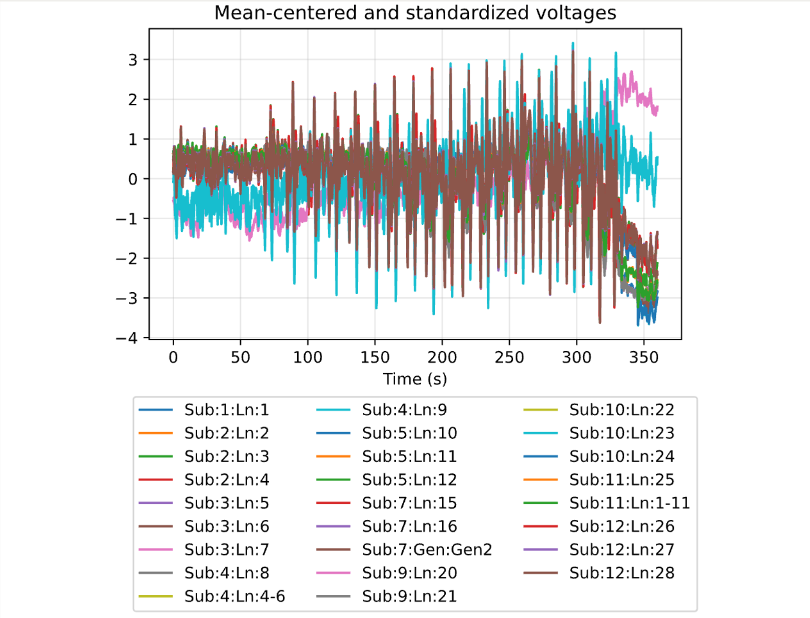


Figure 3 – Mean-centered and standardized voltages

PCA: Cumulative variance explained

- The 3 main modes of oscillation identified by the three most significant Principal Components represent a hierarchical coordinate system based on data that will represent the statistical variation in the data set[5].

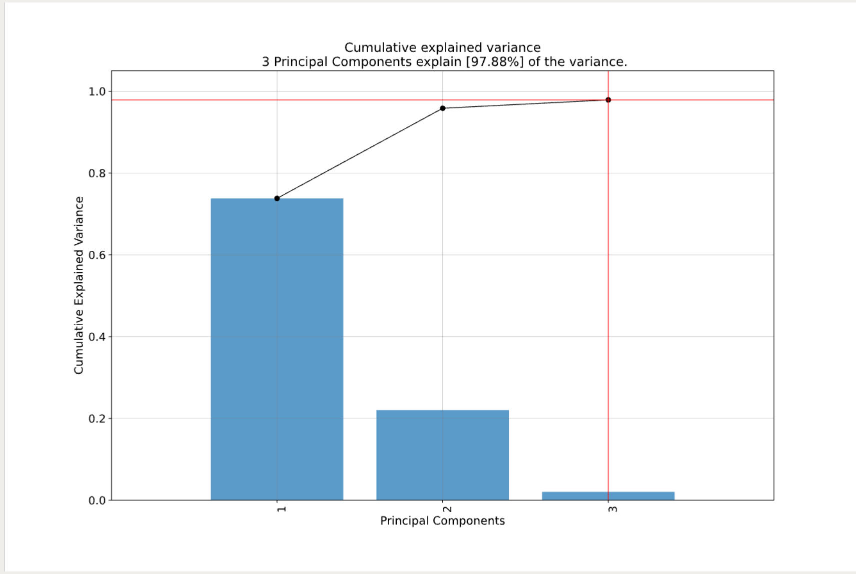


Figure 4 – Cumulative variance explained by the first three principal components

Can we cluster the 3 Principal Components?

By graphing the three most significant principal components, certain clusters can be visualised in the data. By applying a clustering technique, such as k-means clustering, these clusters can be shown.

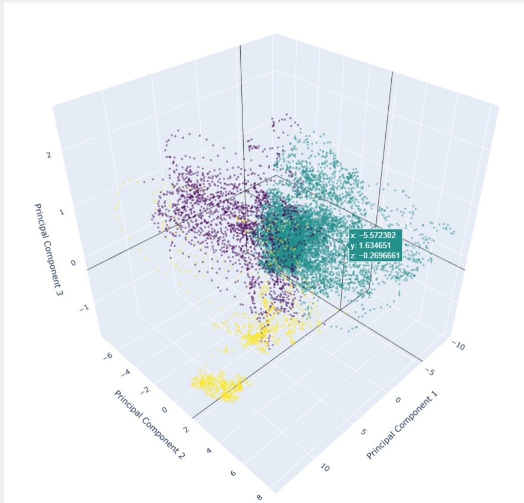


Figure 5 – Principal component space of PMU voltage measurements - clustered via k-means clustering

Clustering can be improved!

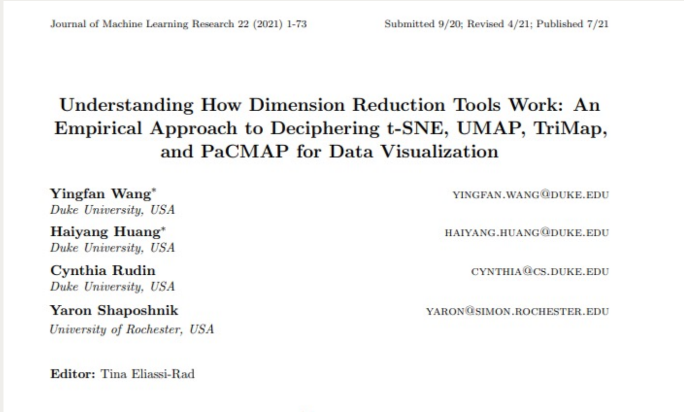
[5] S. L. Brunton and J. N. Kutz, Data-Driven Science and Engineering, 2nd edition. Seattle, WA: Cambridge University Press, 2022. doi: https://doi.org/10.1017/9781009089517.

Conclusions and further work

- SVD and PCA can effectively being utilized as dimensionality reduction techniques of large datasets, such as the ones found in PMU input.
- Readiness and usefulness of the methods that are easily used in Python.
- Pre-processing of data is important, in order to avoid anomalies that can affect the techniques, specially for PCA.
- Both techniques, SVD and PCA, are linear methodsi.e., they identify clusters from measured data but they are not able to capture nonlinearities. On the example of this work, inter-area oscillation modes occurred, i.e. linear modes of motion. PCA was effectively able to capture most of the variability in the first three principal components. Should the event under analysis included highly non-linear dynamics, particularly under stressed conditions, PCA on its own may have not captured the essence of non-linear interactions in its duration.

Thank you!

Other techniques to test?



Journal of Machine Learning Research 22 (2021) 1-73 Submitted 9/20; Revised 4/21; Published 7/21

Understanding How Dimension Reduction Tools Work: An Empirical Approach to Deciphering t-SNE, UMAP, TriMap, and PaCMAP for Data Visualization

Yingfan Wang* YINGFAN.WANG@DUKE.EDU
Duke University, USA
Haiyang Huang* HAIYANG.HUANG@DUKE.EDU
Duke University, USA
Cynthia Rudin CYNTHIA@CS.DUKE.EDU
Duke University, USA
Yaron Shaposhnik YARON@SIMON.ROCHESTER.EDU
University of Rochester, USA

Editor: Tina Eliassi-Rad