ICHQP 2026
22nd International Conference on Harmonics and Quality of Power
Dresden, Germany

#50

Identification and Visualization of Correlation Structures in Large-Scale Power Quality Data

Max Domagk, Jan Meyer, Marco Lindner

Motivation

Current developments

- Increase in renewables and new technologies (e.g. HVDC) → more power electronics
- Potential impact on PQ levels (e.g. harmonics) → extended PQ monitoring campaigns
- Large amounts of data (often only compliance assessment) → underutilized potential

Correlation Analysis

- Identification of systematic and recurring correlations
- Analysis of dependencies and propagation
- Open challenge: multivariate data (sites, parameters, phases) exceeds manual analysis

Motivation

State of the art

- Mostly pairwise relationships, few sites and PQ parameters
- Manual selection and interpretation

Gap

- No automated, scalable analysis for full measurement campaigns

This paper

- Automated analysis and visualization of correlation structures
- Daily Spearman coefficients, aggregated across phases, parameters and sites
- Distance-based visualization (clustering and 2D projection)

Measurement Data

Dataset

- German transmission system
- 85 measurement sites (110-380 kV)
- 10-min values acc. to IEC 61000-4-30
- 20 voltage and current parameters
  - RMS,
  - THD/THC,
  - harmonics (order 3–15, odd),
  - flicker,
  - Unbalance

→ 58 parameter-phase combinations

Correlations

- Between parameters:

  140,505 = 1,653 pairs per site × 85 sites

- Between sites:

  207,060 = 3,570 pairs per parameter × 58

→ 347,565 correlations in total

Methodology

- Hundreds of thousands correlations → interpretable structures

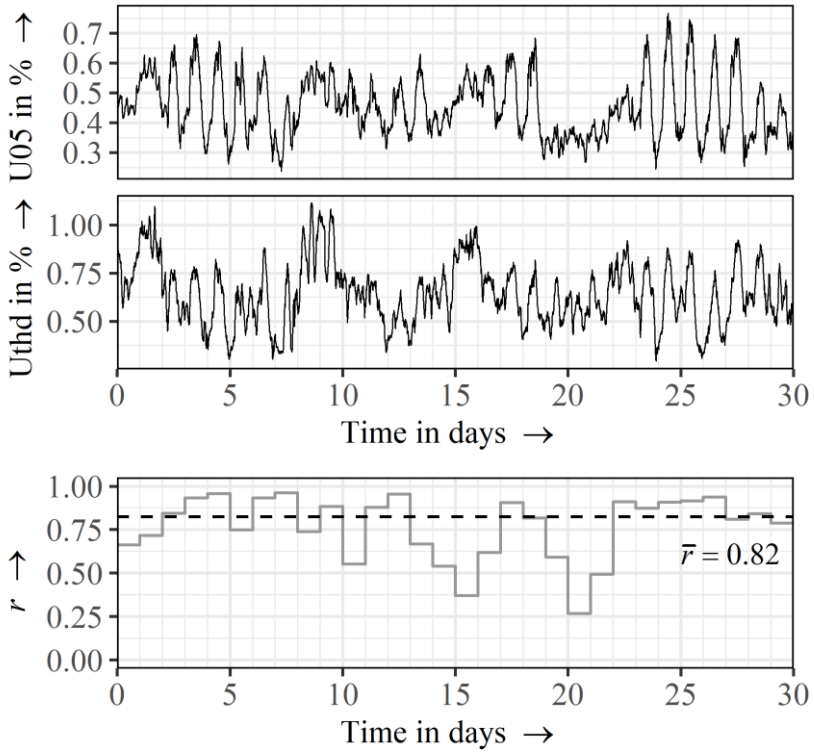
Calculation of daily correlation coefficients

- Daily Spearman coefficients
- Per site and parameter pair

Averaging of coefficients

- Averaging over all days/phases
- Fisher-z transformation

Parameter pair of one site



Methodology

- Hundreds of thousands correlations → interpretable structures

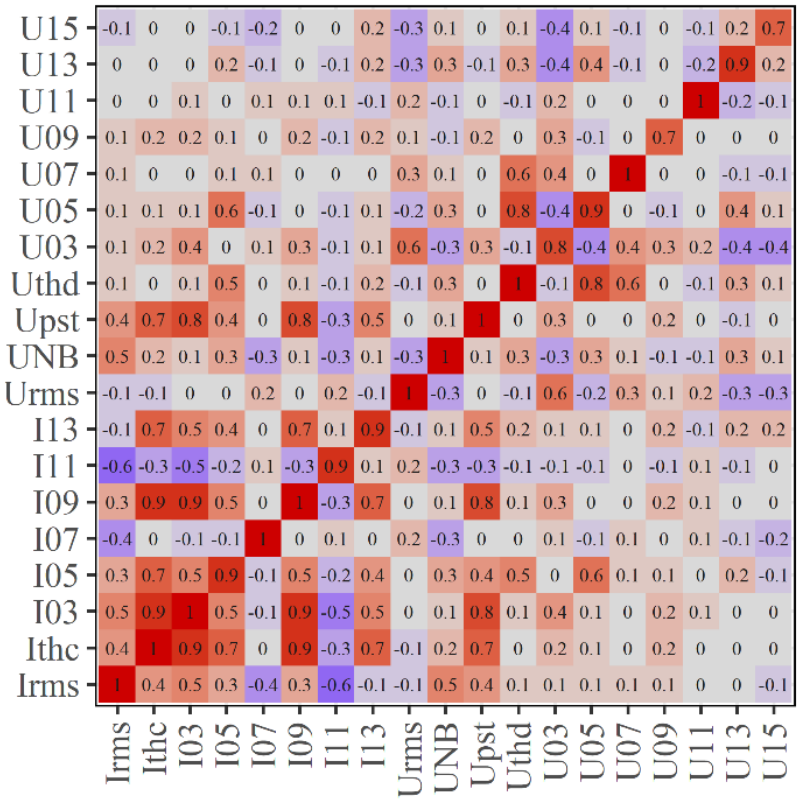
Calculation of daily correlation coefficients

- Daily Spearman coefficients
- Per site and parameter pair

Averaging of coefficients

- Averaging over all days/phases
- Fisher-z transformation

Parameters of one site



Methodology

- Hundreds of thousands correlations → interpretable structures

Calculation of daily correlation coefficients

- Daily Spearman coefficients
- Per site and parameter pair

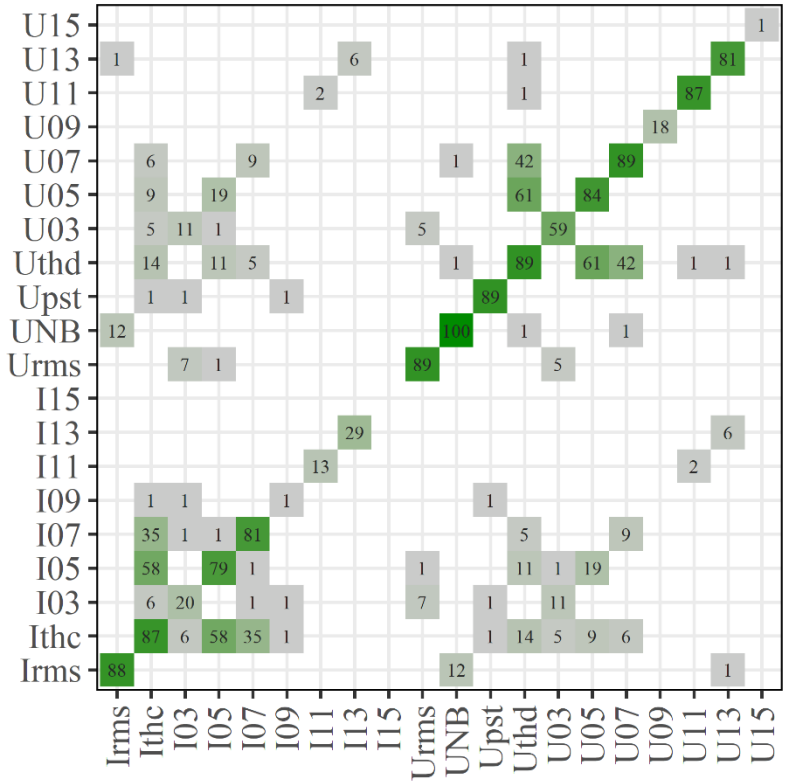
Averaging of coefficients

- Averaging over all days/phases
- Fisher-z transformation

Aggregation of significant correlations

- Significant correlations ($r \geq 0.7$)
- Share across all parameters/sites

Share of sites with significant correlations



Methodology

- Hundreds of thousands correlations → interpretable structures

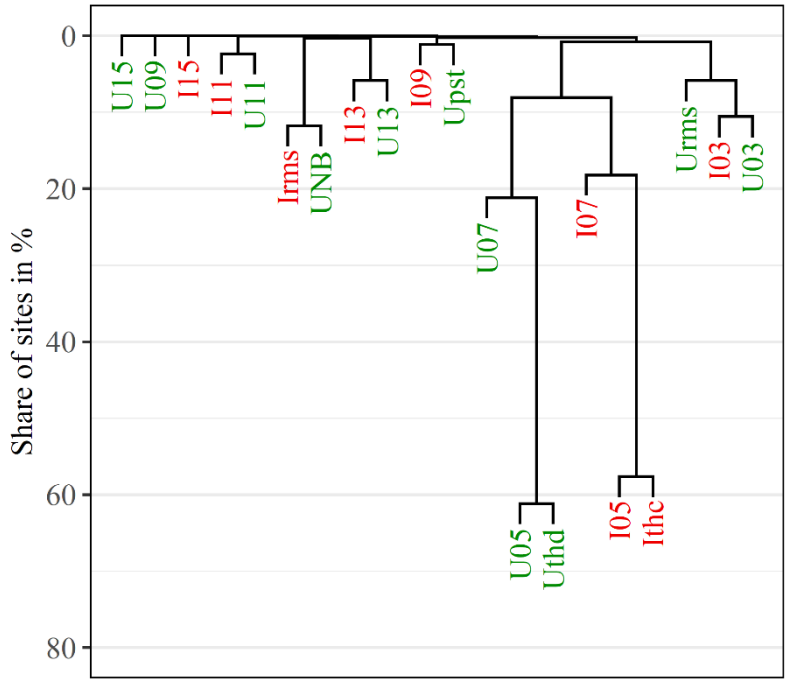
Calculation of daily correlation coefficients
- Daily Spearman coefficients
- Per site and parameter pair

Averaging of coefficients
- Averaging over all days/phases
- Fisher-z transformation

Aggregation of significant correlations
- Significant correlations ($r \geq 0.7$)
- Share across all parameters/sites

Visualization
- Correlation → distance
- Clustering and 2D projection

Dendrogram of hierarchical clustering



Methodology

- Hundreds of thousands correlations → interpretable structures

Calculation of daily correlation coefficients

- Daily Spearman coefficients
- Per site and parameter pair

Averaging of coefficients

- Averaging over all days/phases
- Fisher-z transformation

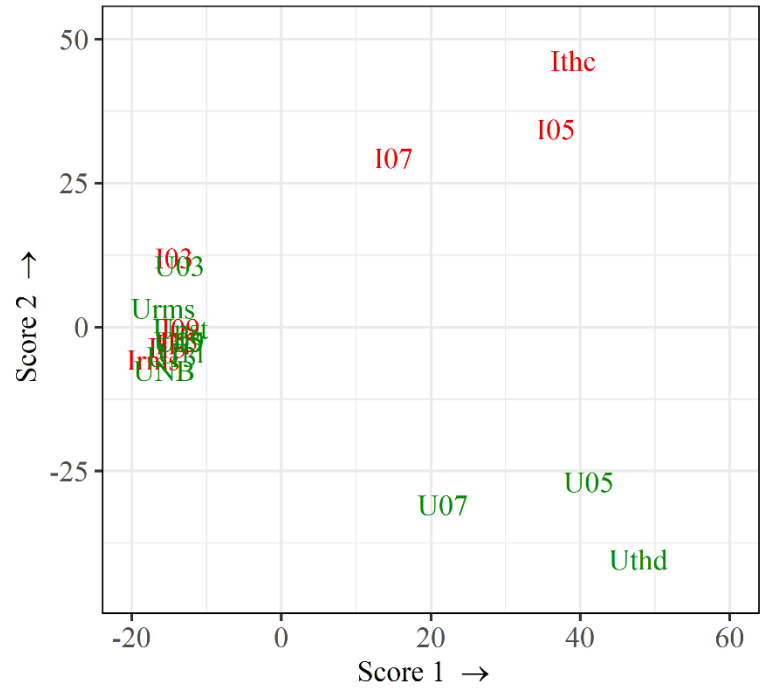
Aggregation of significant correlations

- Significant correlations ($r \geq 0.7$)
- Share across all parameters/sites

Visualization

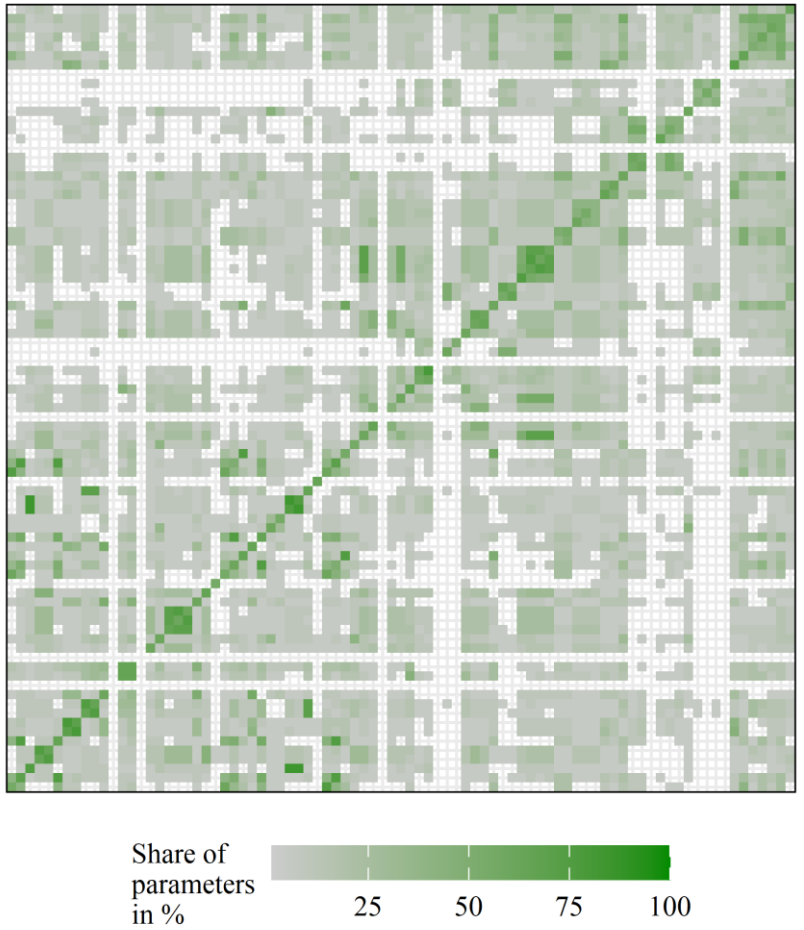
- Correlation → distance
- Clustering and 2D projection

2D projection via multidimensional scaling



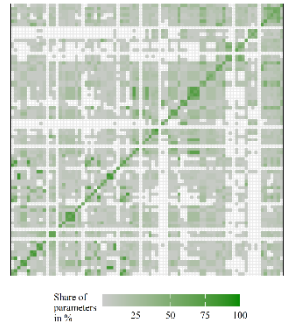
Results – Correlations between sites

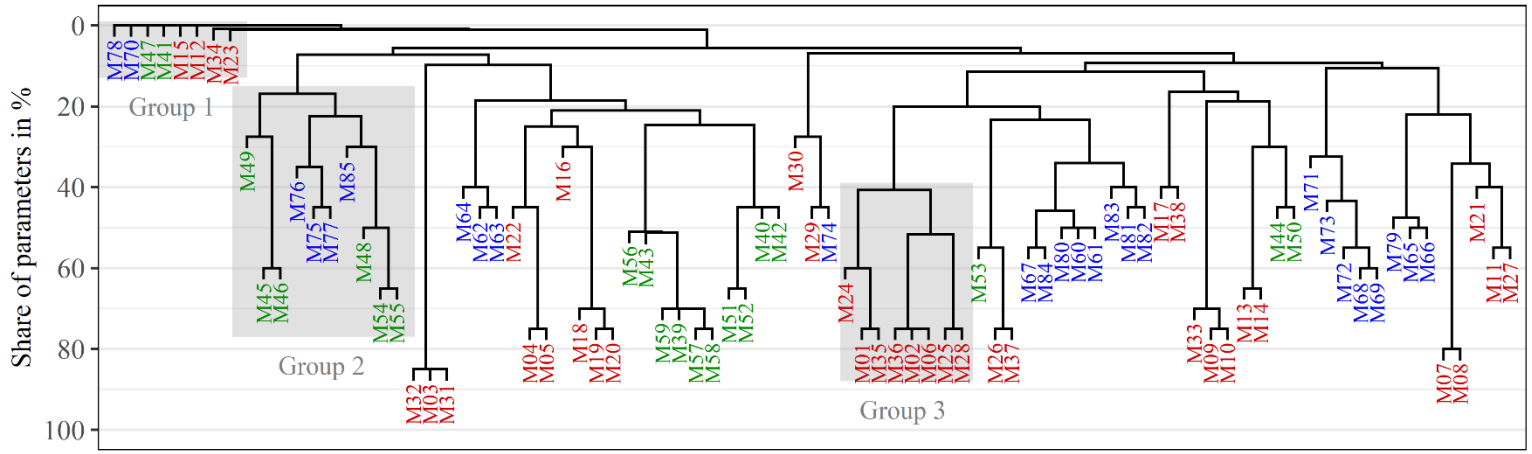
- Aggregation matrix for 85 sites:
  too large for direct interpretation



Results – Correlations between sites

- Aggregation matrix for 85 sites: too large for direct interpretation
- Visualizations make the correlation structure interpretable
- Three exemplary groups:
  - **Group 1:** no or very few correlations with other sites
  - **Group 2:** 220/380 kV, 20–60 % of parameters correlated
  - **Group 3:** 110 kV, 40–75 % correlated, weak links to other groups





110 kV, 220 kV, 380 kV

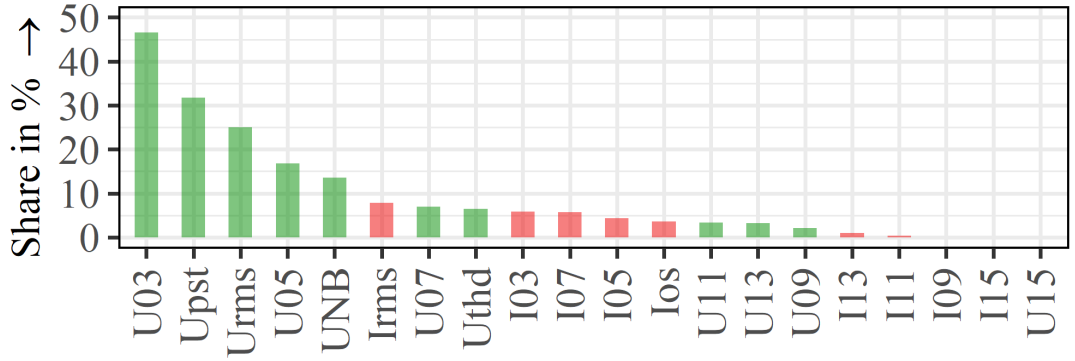
Results – Correlations between sites

- Most sites correlate for only 1-2 parameters

- Highest shares: U03 (47 %) and Upst (33 %)

→ propagation across large parts of the network

- Higher-order harmonics remain local



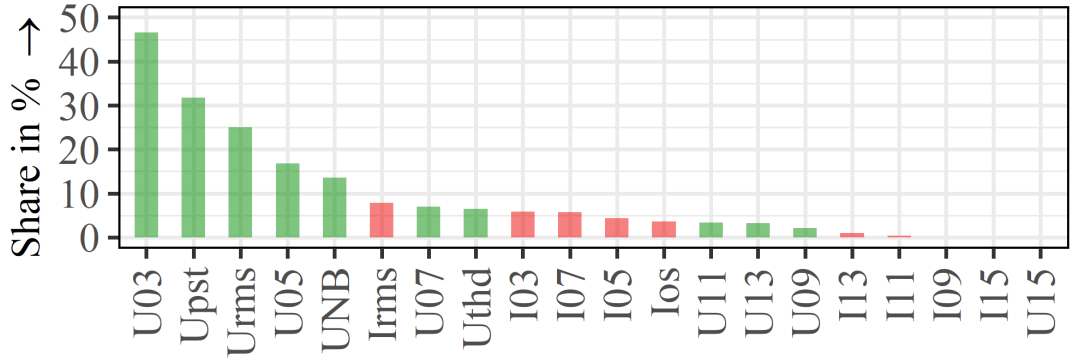
Results – Correlations between sites

- Most sites correlate for only 1-2 parameters

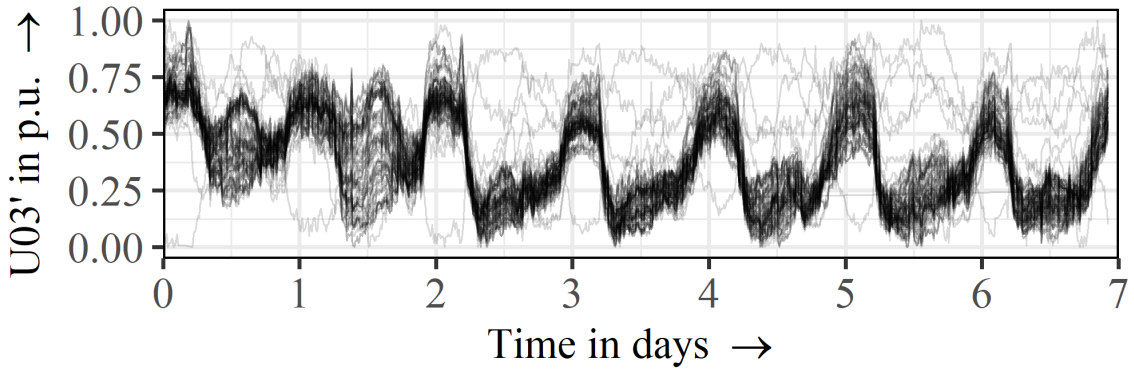
- Highest shares: U03 (47 %) and Upst (33 %)

→ propagation across large parts of the network

- Higher-order harmonics remain local



- Time series of U03 across all 85 sites confirm the temporal similarity



Conclusion

Findings

- Averaging and aggregation make hundreds of thousands of correlations interpretable
- Clustering and 2D projections reveal correlation structures
- Flicker and 3rd voltage harmonic tend to correlate network-wide, higher-orders remain local
- Methodology scales to full measurement campaigns (85 sites, 20 parameters)

Outlook

- Correlation-based time segmentation → temporal changes and operational regimes

Thank you for your attention!

Technische Universität Dresden

max.domagk@tu-dresden.de
+49 351 463 35223
maxdomagk.de