

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

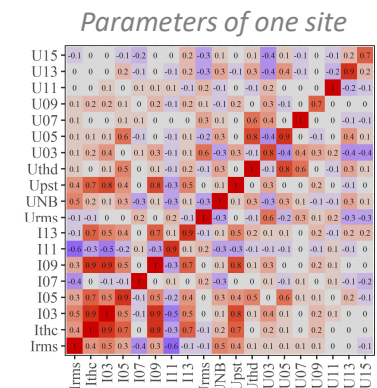
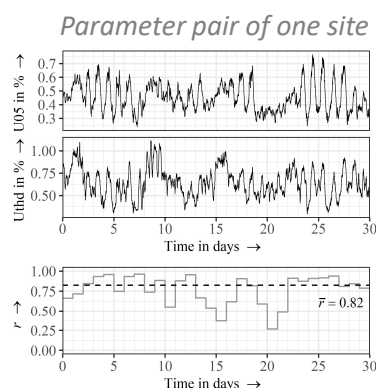
Max Domagk, Jan Meyer, Marco Lindner

Motivation

- Large PQ measurement campaigns: 85 sites, 20 parameters → 347,565 correlations
- Systematic dependencies not identifiable manually

1 Calculation

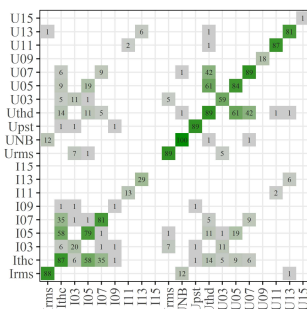
- Daily Spearman correlation coefficients for all parameter and site pairs
- Averaged over days and phases (Fisher-z transformation)
- One correlation matrix per site → 85 matrices hardly comparable



2 Aggregation

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

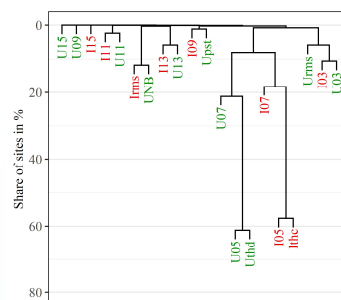
Share of sites with significant correlations



3 Visualization

- Correlation converted to distance
- Hierarchical clustering and 2D projection

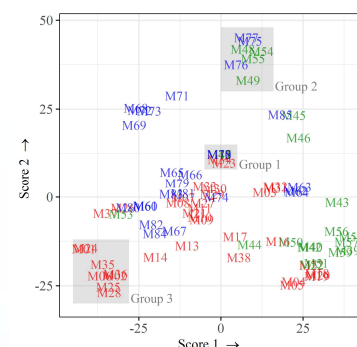
Dendrogram of hierarchical clustering



4 Results

- Three exemplary groups: isolated sites, 220/380 kV, 110 kV
- Highest shares: U03 (47 %) and Upst (33 %) → propagation across the network

2D projection via multidimensional scaling



Conclusion

- Averaging and aggregation make hundreds of thousands of correlations interpretable
- Flicker and 3rd voltage harmonic tend to correlate network-wide, higher orders stay local

Funded by

DFG Deutsche Forschungsgemeinschaft
German Research Foundation