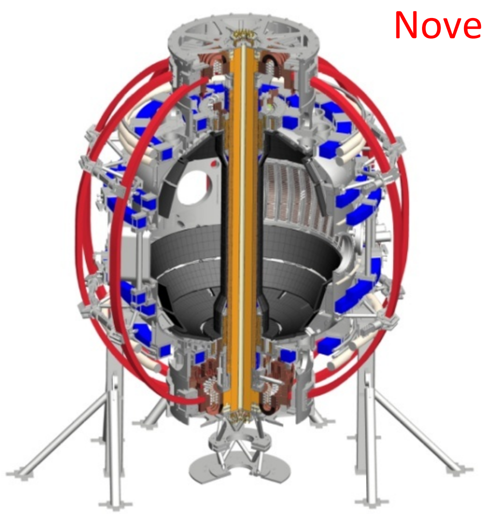
Evolution patterns and parameter regimes in edge localized modes on NSTX

Coll of Wm & Mary
Columbia U
CompX
General Atomics
FIU
INL
Johns Hopkins U
LANL
LLNL
Lodestar
MIT
Lehigh U
Nova Photonics
ORNL
PPPL
Princeton U
Purdue U
SNL
Think Tank, Inc.
UC Davis
UC Irvine
UCLA
UCSD
U Colorado
U Illinois
U Maryland
U Rochester
U Tennessee
U Tulsa
U Washington
U Wisconsin
X Science LLC

David R. Smith[1], G. McKee[1], R. Fonck[1],
A. Diallo[2], S. Kaye[2], and B. LeBlanc[2]
[1]*U. Wisconsin-Madison*
[2]*PPPL*

57th APS-DPP Meeting
Savannah, GA
November 16-20, 2015



Culham Sci Ctr
York U
Chubu U
Fukui U
Hiroshima U
Hyogo U
Kyoto U
Kyushu U
Kyushu Tokai U
NIFS
Niigata U
U Tokyo
JAEA
Inst for Nucl Res, Kiev
Ioffe Inst
TRINITI
Chonbuk Natl U
NFRI
KAIST
POSTECH
Seoul Natl U
ASIPP
CIEMAT
FOM Inst DIFFER
ENEA, Frascati
CEA, Cadarache
IPP, Jülich
IPP, Garching
ASCR, Czech Rep

Nonlinear ELM models are needed to predict ELM intensity, filament dynamics, saturation mechanisms, and transport properties

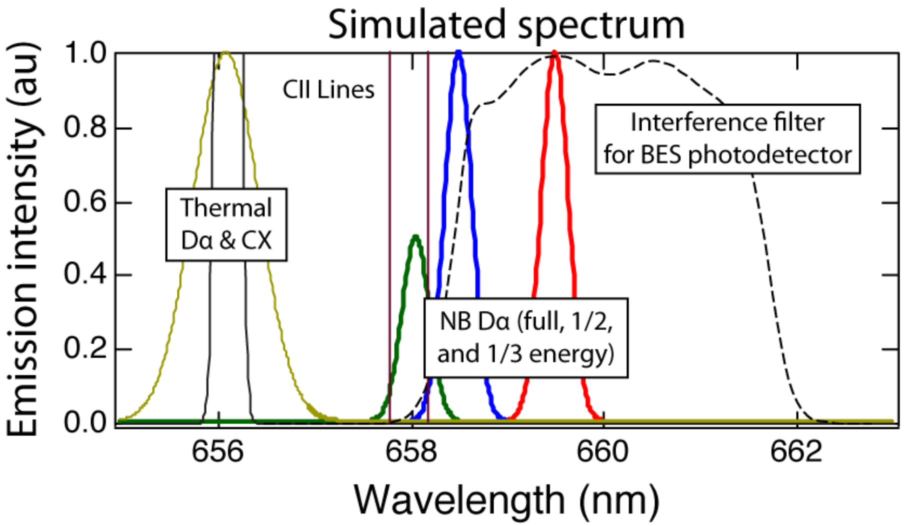
- Fast measurements show the nonlinear dynamics of ELM events, but linear models only capture ELM onset conditions

- Here, we investigate the nonlinear dynamics of NSTX ELM events with beam emission spectroscopy (BES) measurements
  - Machine learning techniques identify groups of ELMs with similar evolution characteristics
    - Unsupervised hierarchical and k-means clustering
    - Time series similarity metrics
  - The analysis provides an ELM classification scheme based on fast, direct measurements of ELM burst dynamics
  - The results illustrate an application of machine learning analysis to data-rich national fusion facilities

Outline

- Beam emission spectroscopy (BES) measurement principles
- BES measurements of ELM events
- Identifying ELM groups with machine learning analysis
    - Time series similarity metrics
    - Hierarchical and k-means cluster analysis
- Parameter regimes for identified ELM groups
- Leveraging large data volumes at national fusion facilities with machine learning techniques
    - Automatic pattern identification, data reduction, and hypothesis generation/testing
- Summary & BES upgrade for NSTX-U

Beam emission spectroscopy (BES) measures Doppler-shifted D_α emission from neutral beam particles

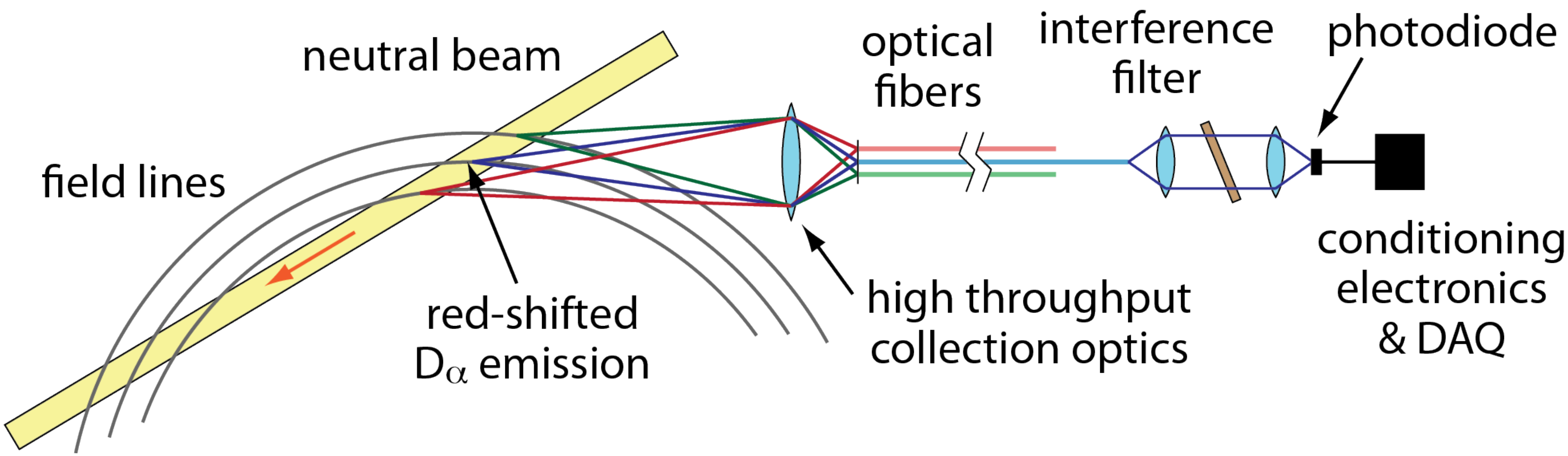


$$\frac{\delta I_{D\alpha}}{I_{D\alpha}} = \frac{\delta n}{n} \times C\left(E_{NB}, n, T_e, Z_{eff}\right)$$

neutral beam D_α emission

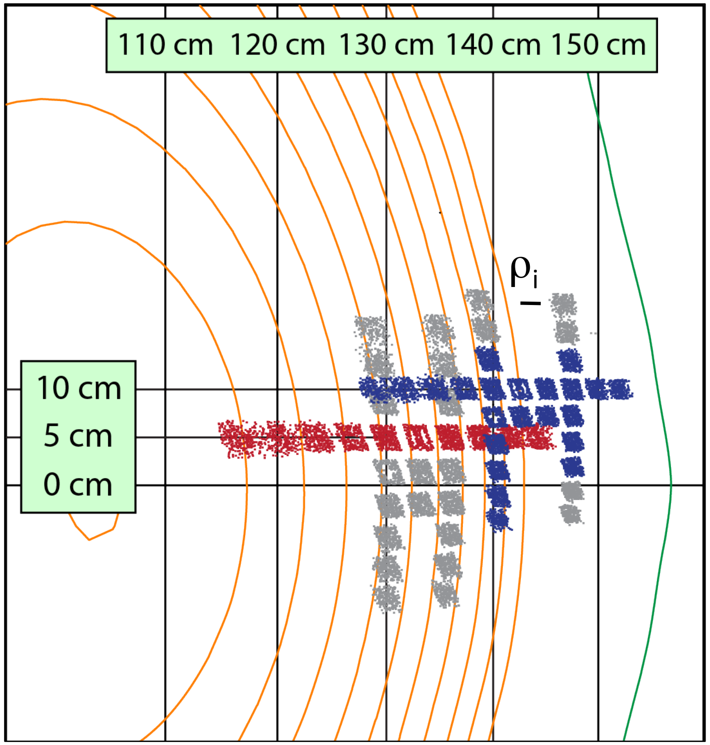
density fluctuation

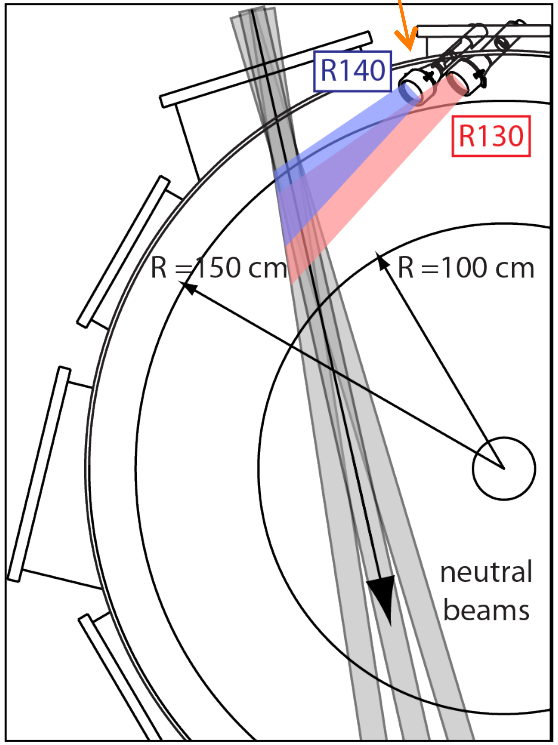
$C \approx 1/2$



The NSTX BES system can observe fluctuations on the ion gyroscale with $k_\perp \rho_i \leq 1.5$

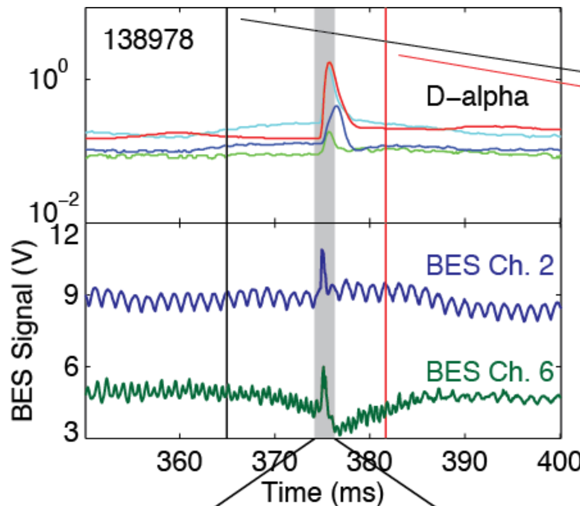
- Radial and poloidal arrays cover core/SOL
- 32 detection channels
- 2-3 cm spot size and $k_\perp \rho_i \leq 1.5$
- **Upgrade in progress: 2D layout and 16 additional detection channels**

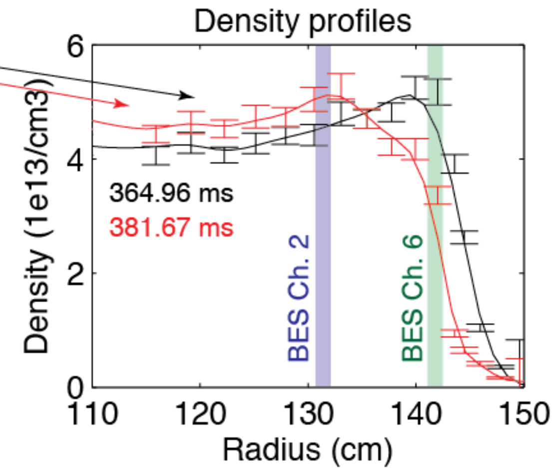




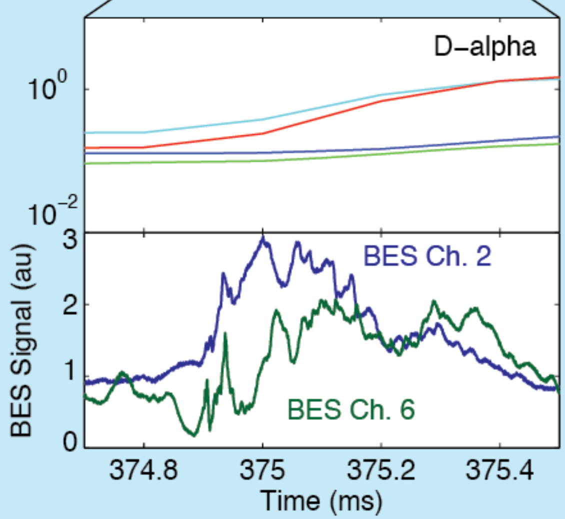
BES measurements with high time resolution show the μs evolution and radial profile of ELM events

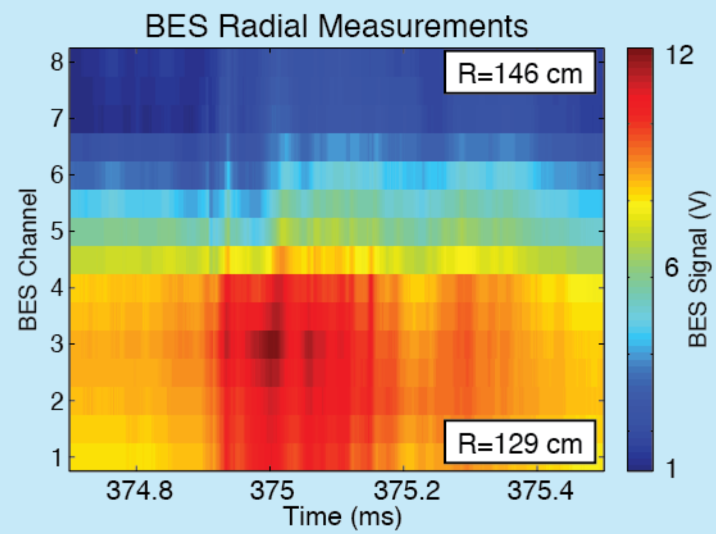
Slow measurements fail to capture the fast evolution of ELM events





BES captures the μs-scale evolution of ELM events

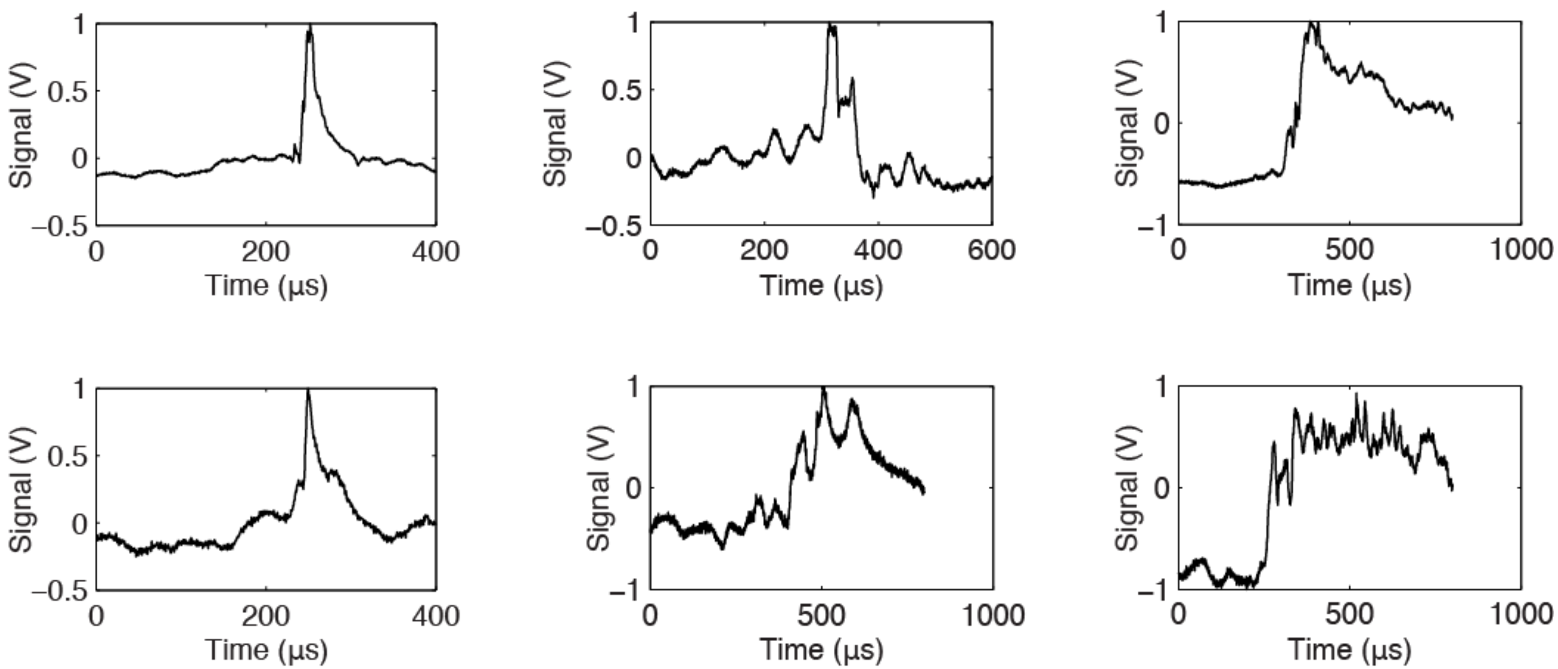




Measurements show significant variation in ELM evolution, but linear ELM models can not capture evolution dynamics

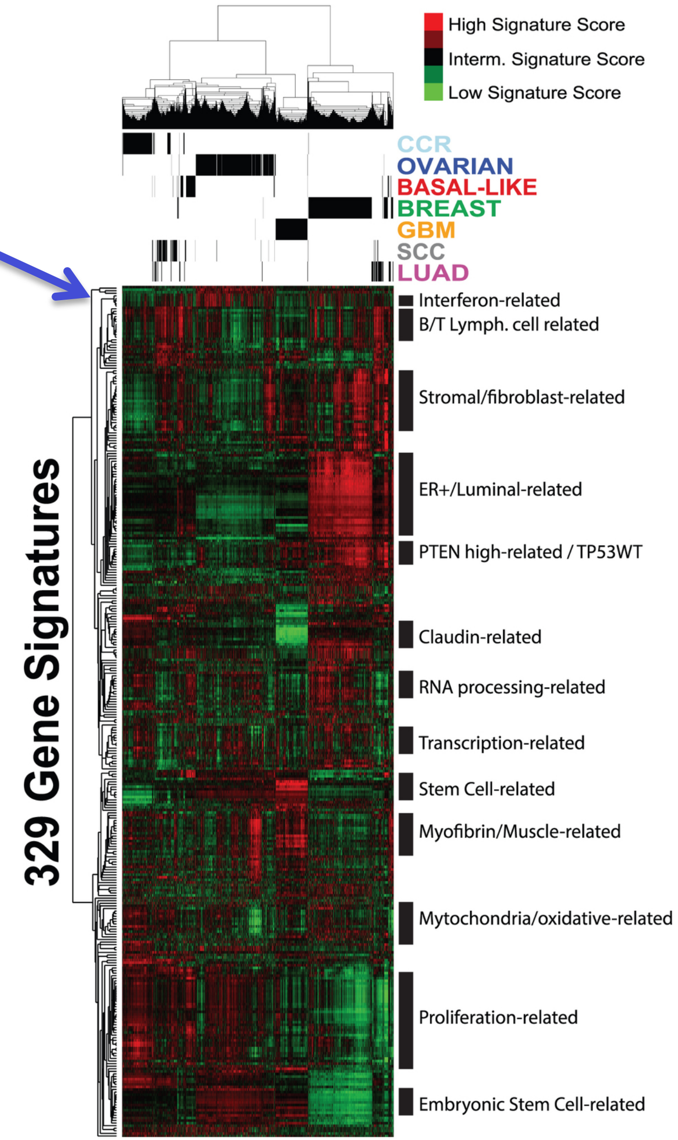
- Database of 51 ELM events measured with BES
  - 8 radial BES channels spanning pedestal region
  - Sampled from 34 NSTX discharges from 8 run days spanning 4 months
  - 1%-16% stored energy loss and observable pedestal collapse

Examples from the ELM database



To identify groups of ELMs with similar evolution characteristics, we apply cluster analysis to ELM time series data

- Hierarchical clustering
  - Popular in data-rich genomics to connect genes and functional pathways
  - Output: a multilevel hierarchy that merges the most similar objects
  - Requires a **similarity metric**
- Time series similarity metrics
  - Time-lag correlation
  - Euclidean distance
  - Dynamic time warping (DTW)
  - Wavelet decomposition
- K-means clustering
  - Output: data partitioned into k mutually exclusive clusters
  - Requires distance metrics in an external coordinate system
  - Optimum cluster number found by trial-and-error

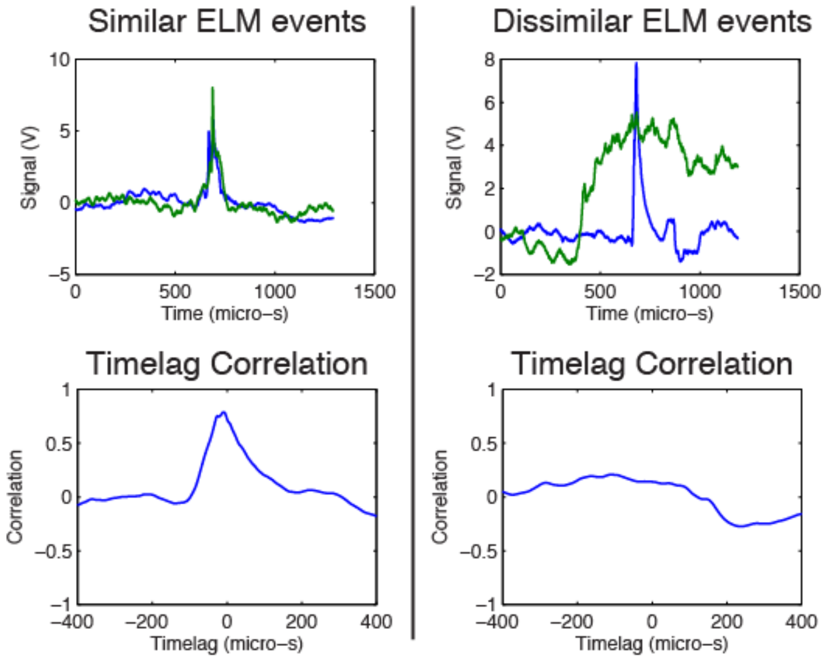


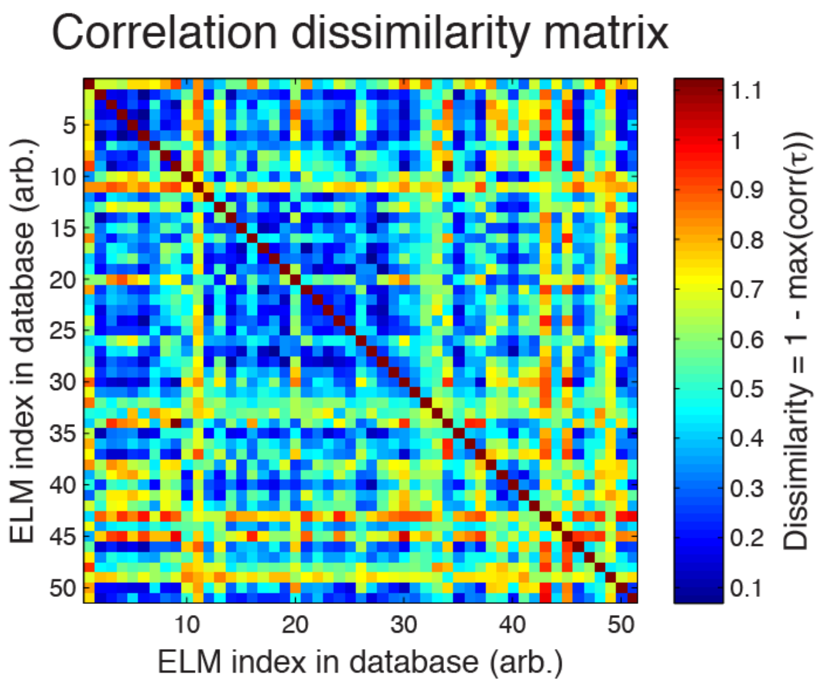
Prat et al, Scientific Reports 3, 3544 (2013)

For hierarchical clustering of time-series data, assemble similarity metrics into a dissimilarity matrix

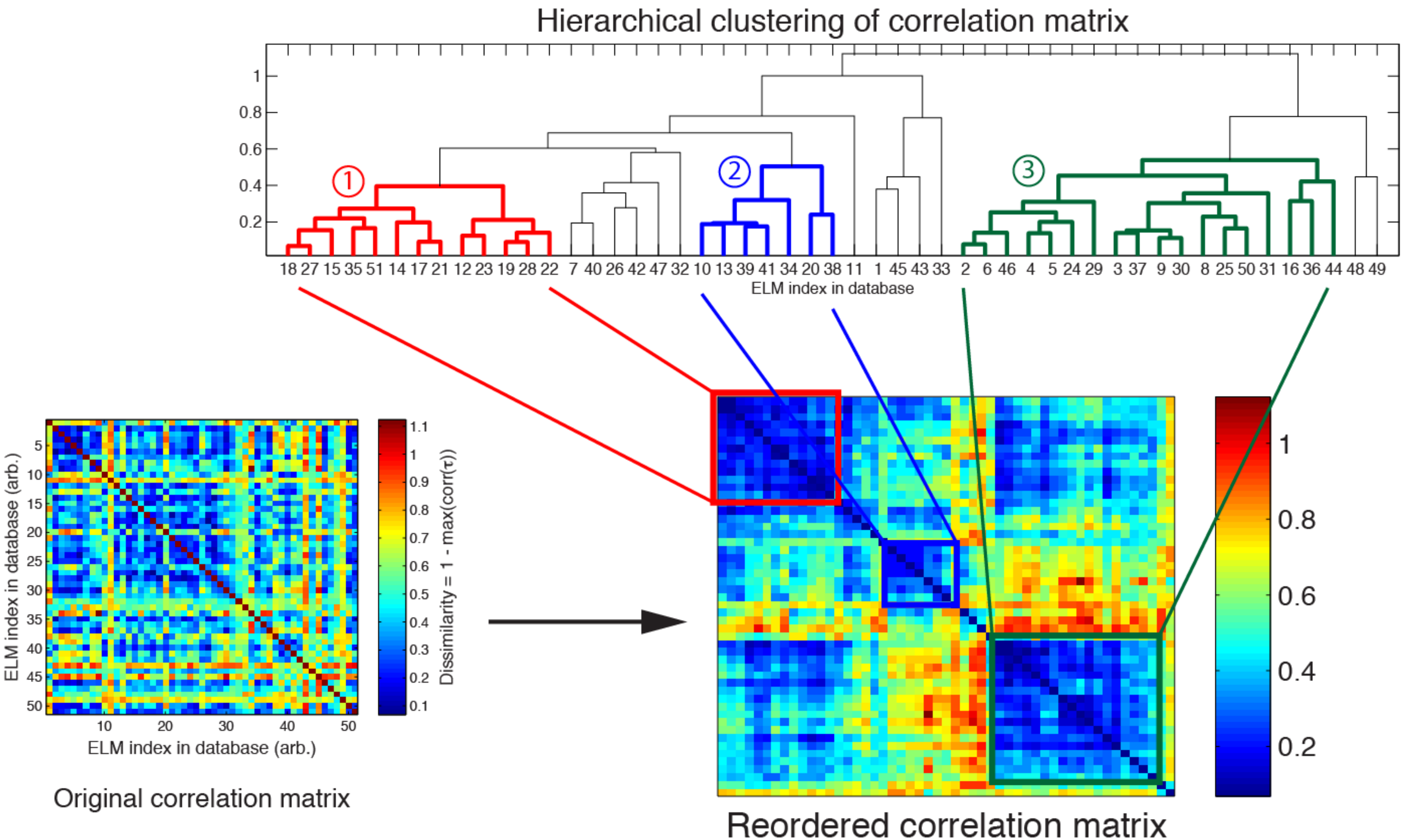
1) Time-lag correlation is an effective similarity metric for time series data

2) Assemble pair-wise metrics into a dissimilarity matrix

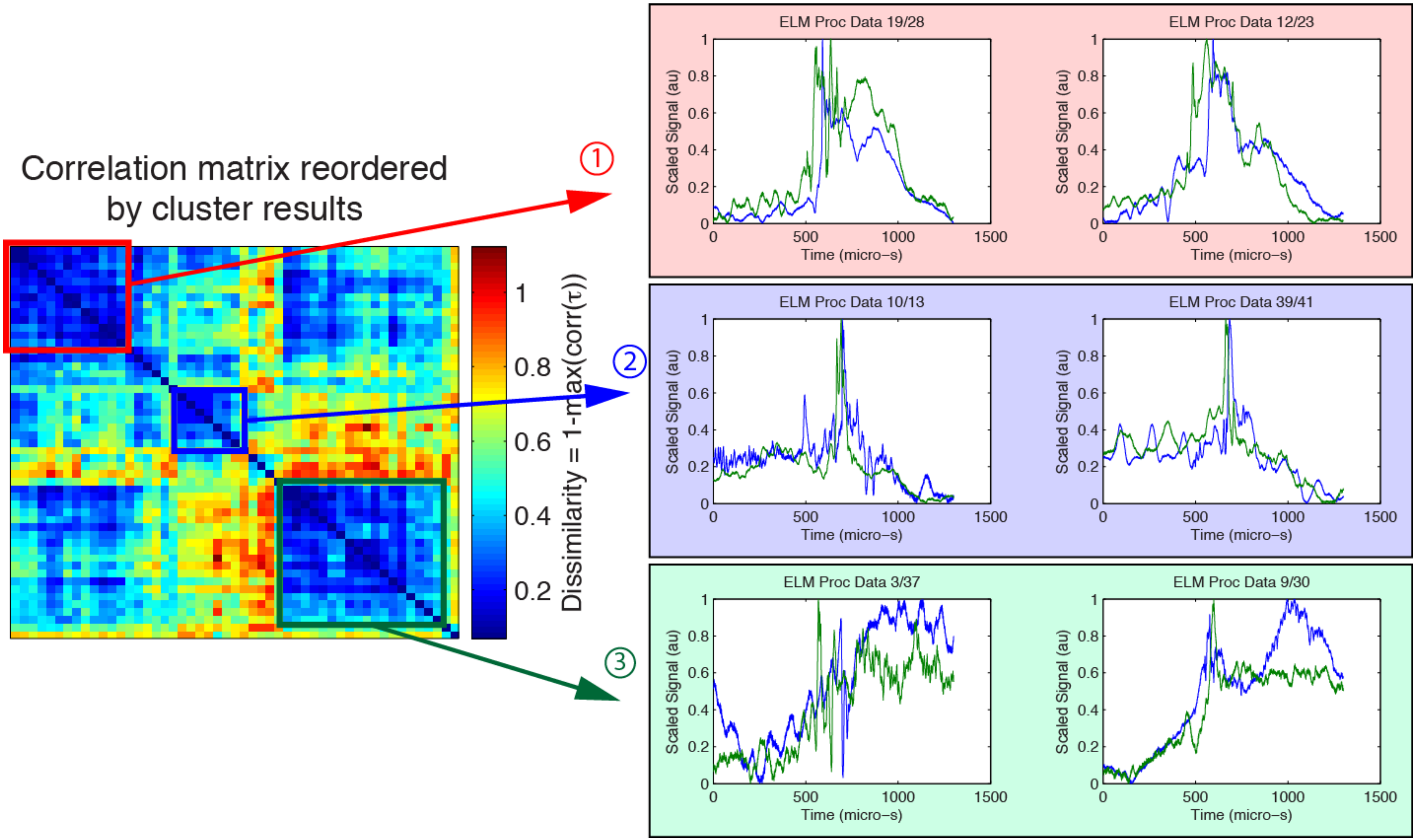




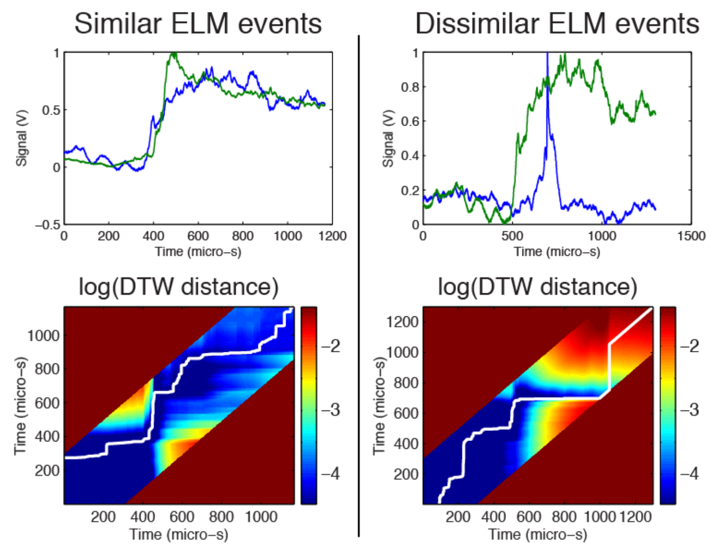
Hierarchical clustering applied to dissimilarity matrix identifies groups of similar ELM events



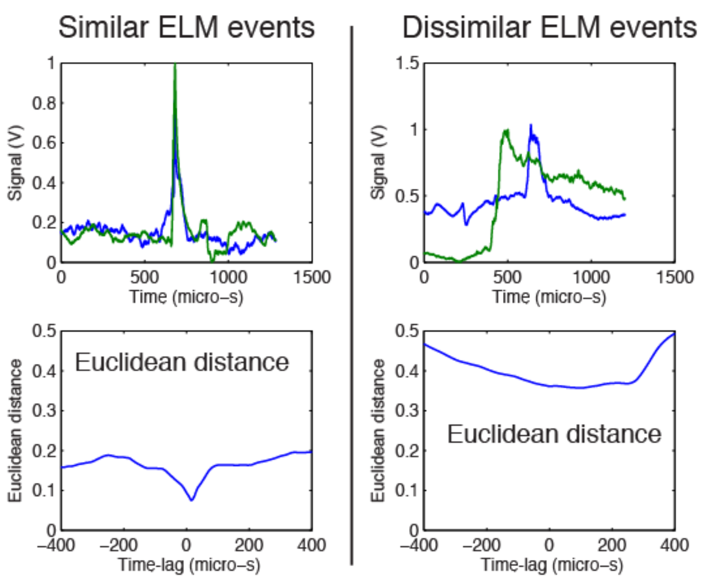
ELM groups identified by hierarchical clustering exhibit similar evolution characteristics



Other effective similarity metrics include Euclidean distance, dynamic time warping, and wavelet decomposition

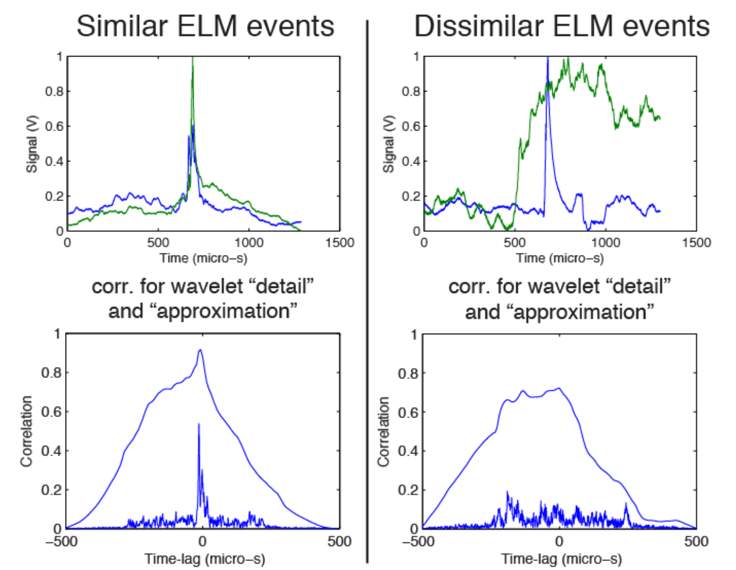


Euclidean distance

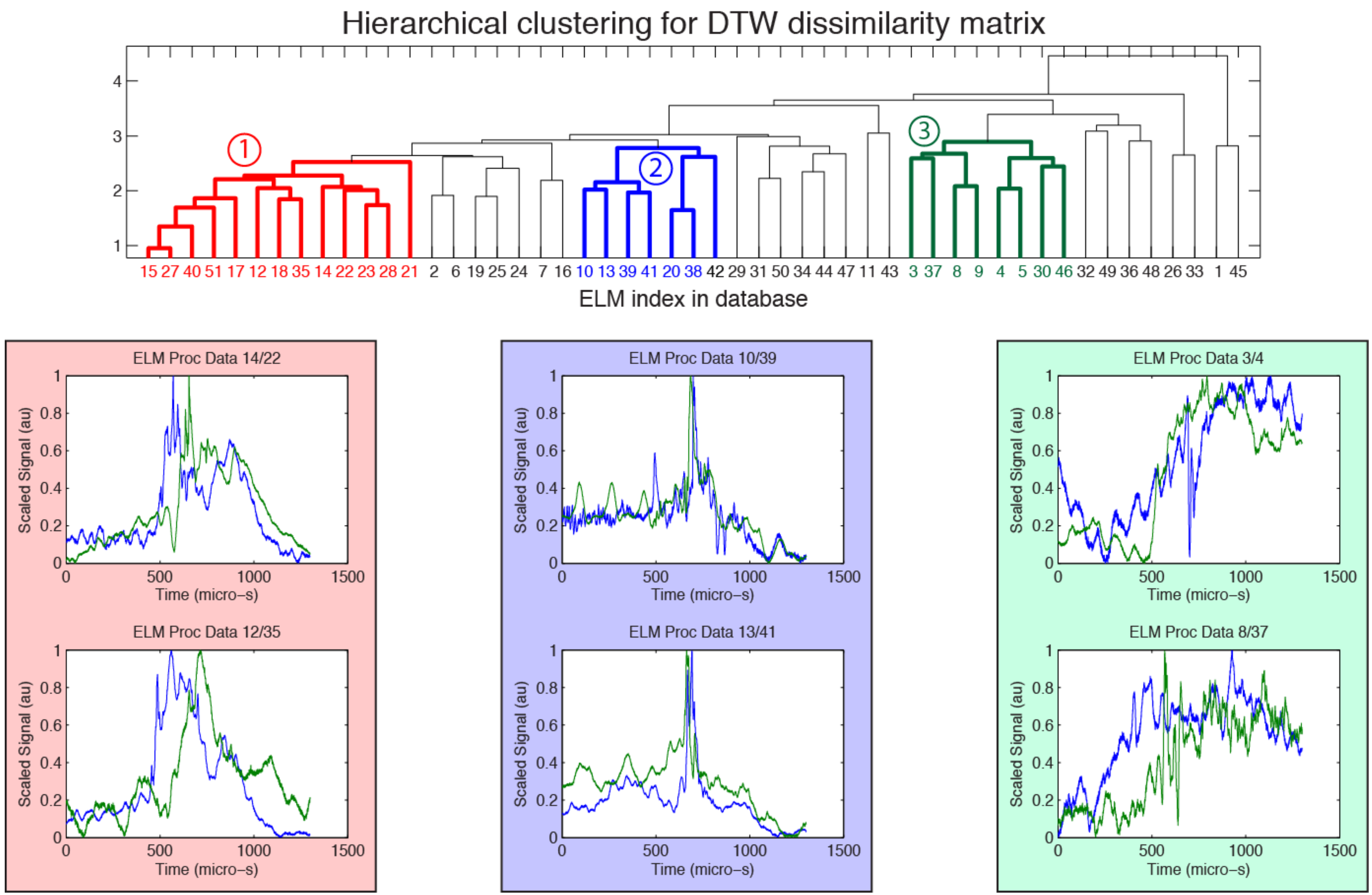


Dynamic time warping

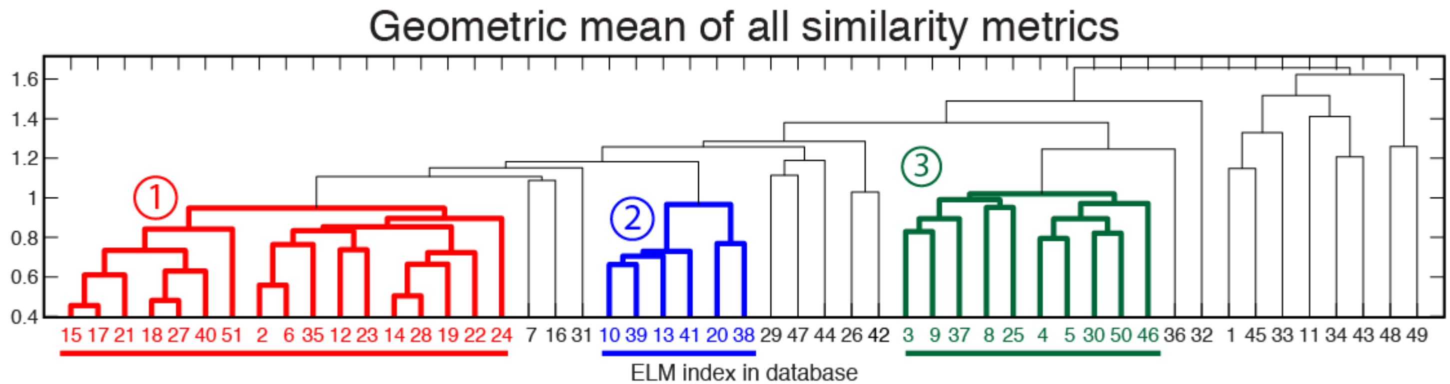
Wavelet decomposition

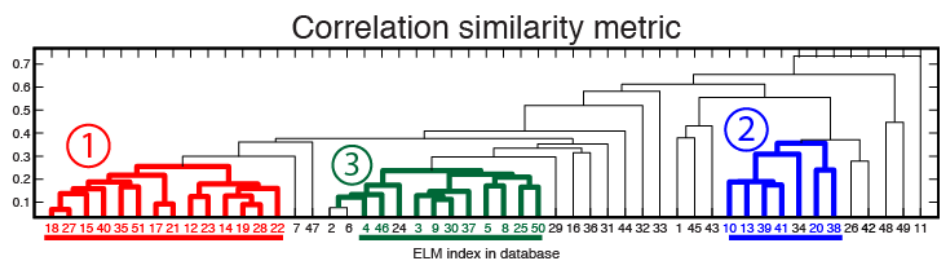


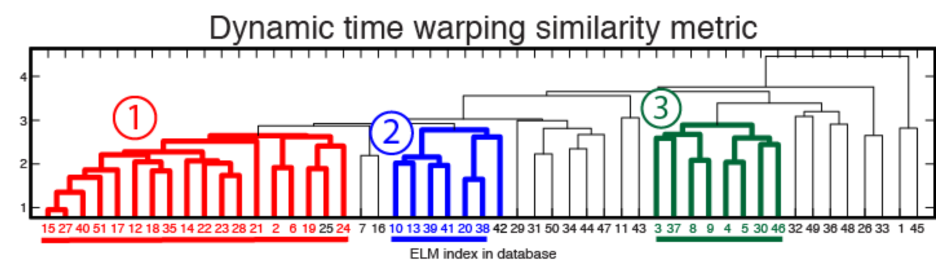
For instance, dynamic time warping (DTW) yields cluster results similar to time-lag correlation (cf. pg. 11)

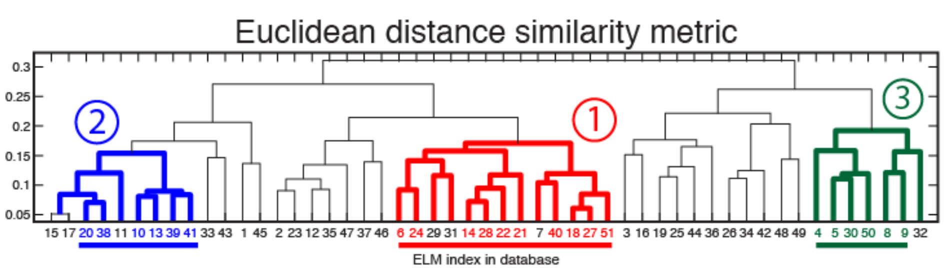


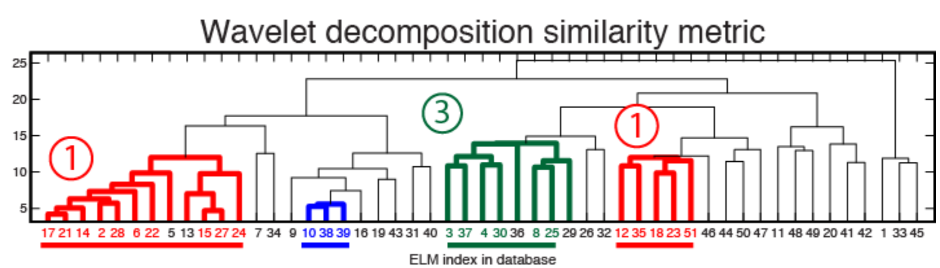
Hierarchical cluster results are largely consistent for all metrics, including the geometric mean of metrics











K-means clustering is feasible by designating a set of benchmark ELMs

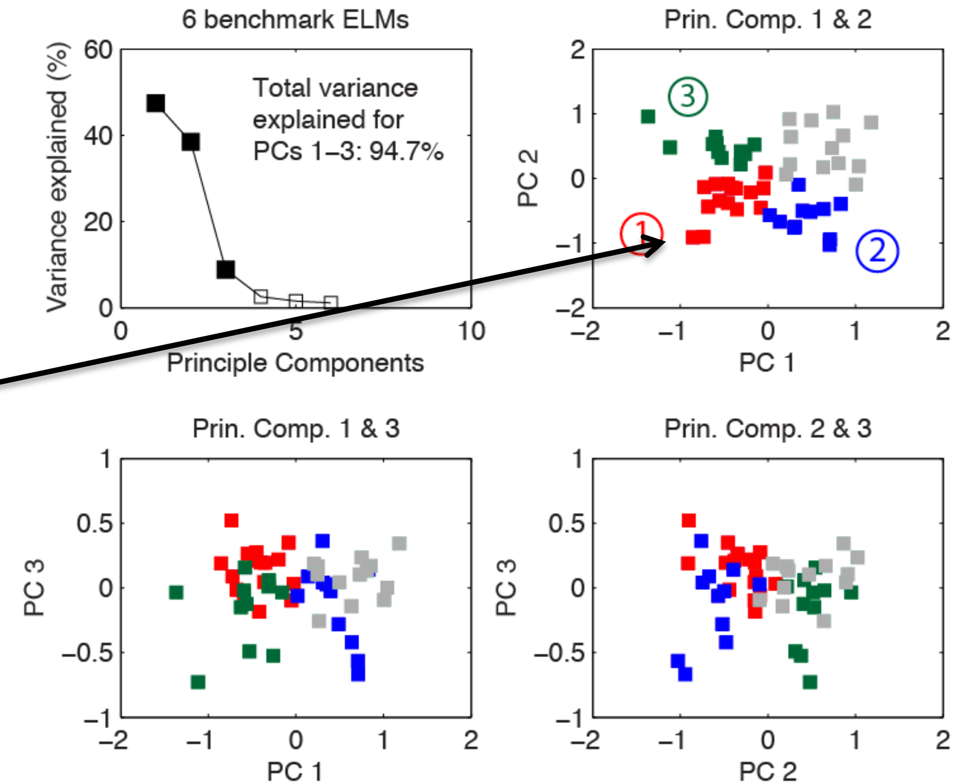
- K-means clustering requires distances in an external coordinate system (hierarchical clustering operates on pair-wise similarity metrics)
- A set of benchmark ELMs can function as the external coordinate system
- Visualize k-means results by plotting groups in principle component space

K-means clustering
with 6 benchmark ELMs

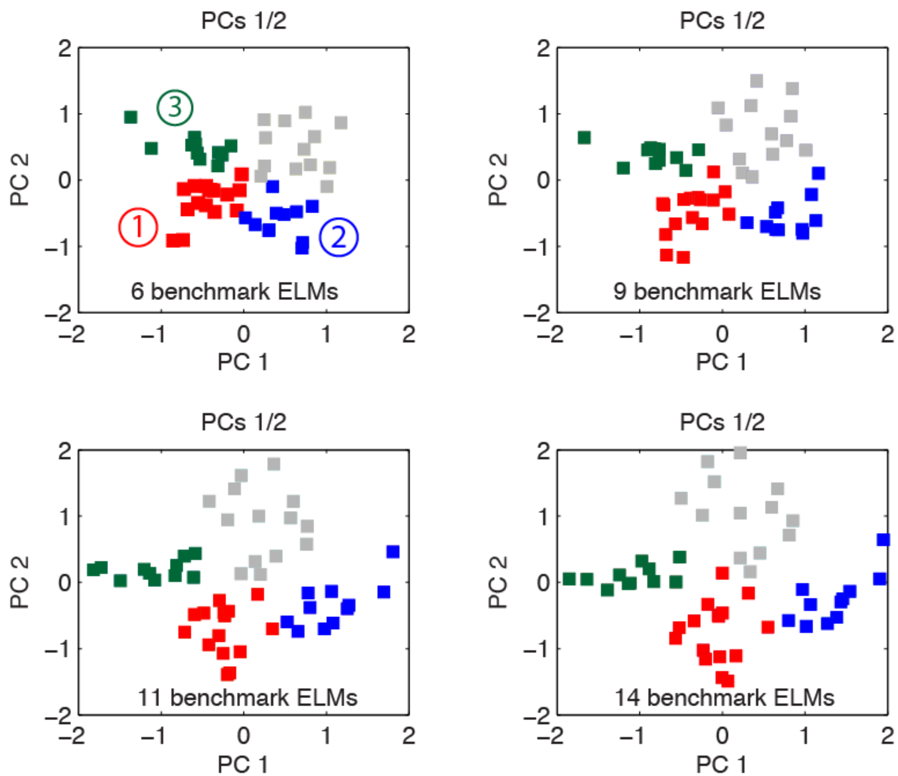
# of clusters	Mean ratio*
2	0.49
3	0.51
4	0.52
5	0.48
6	0.46
7	0.45

Optimal

*Out-of-cluster/in-cluster distance ratio



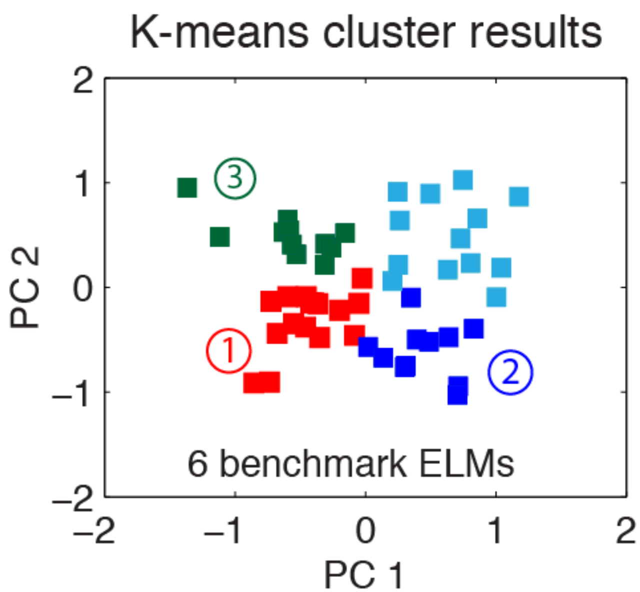
K-means cluster results are consistent for different sets of benchmark ELMs



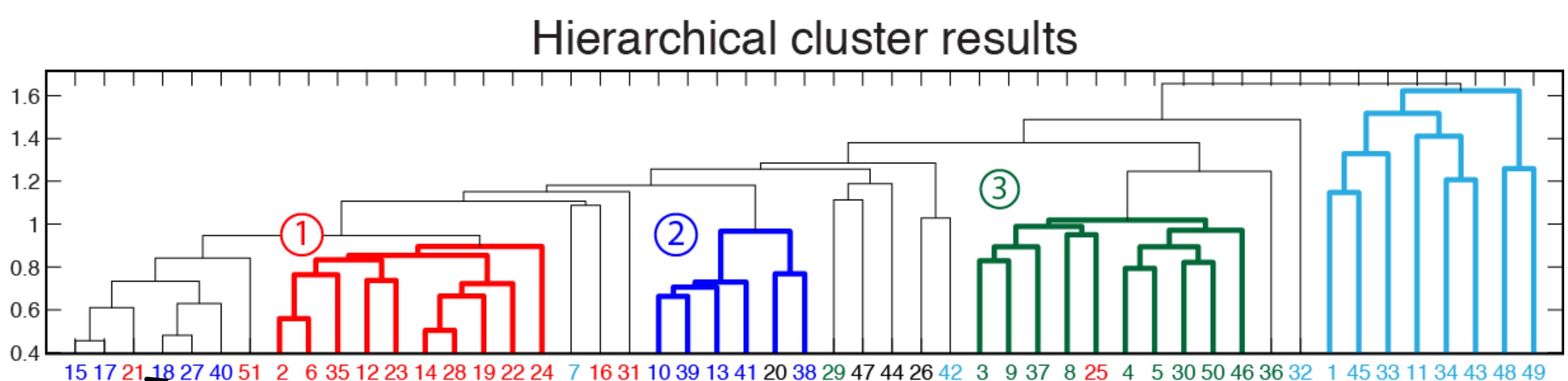
# of benchmark ELMs	Optimal cluster #	Mean ratio
6	4	0.52
9	4	0.52
11	4	0.53
14	4	0.52

- Clusters are highly consistent for calculations with different benchmark ELMs
- Red cluster ELMs: 2, 35, 23, 19, 22, 12, 28, 14, 51, 24
- Blue cluster ELMs: 13, 15, 10, 39, 41, 40, 38, 17, 27, 18
- Green cluster ELMs: 30, 4, 50, 3, 5, 36, 29, 9, 8, 46

Results from k-means clustering and hierarchical clustering are largely consistent



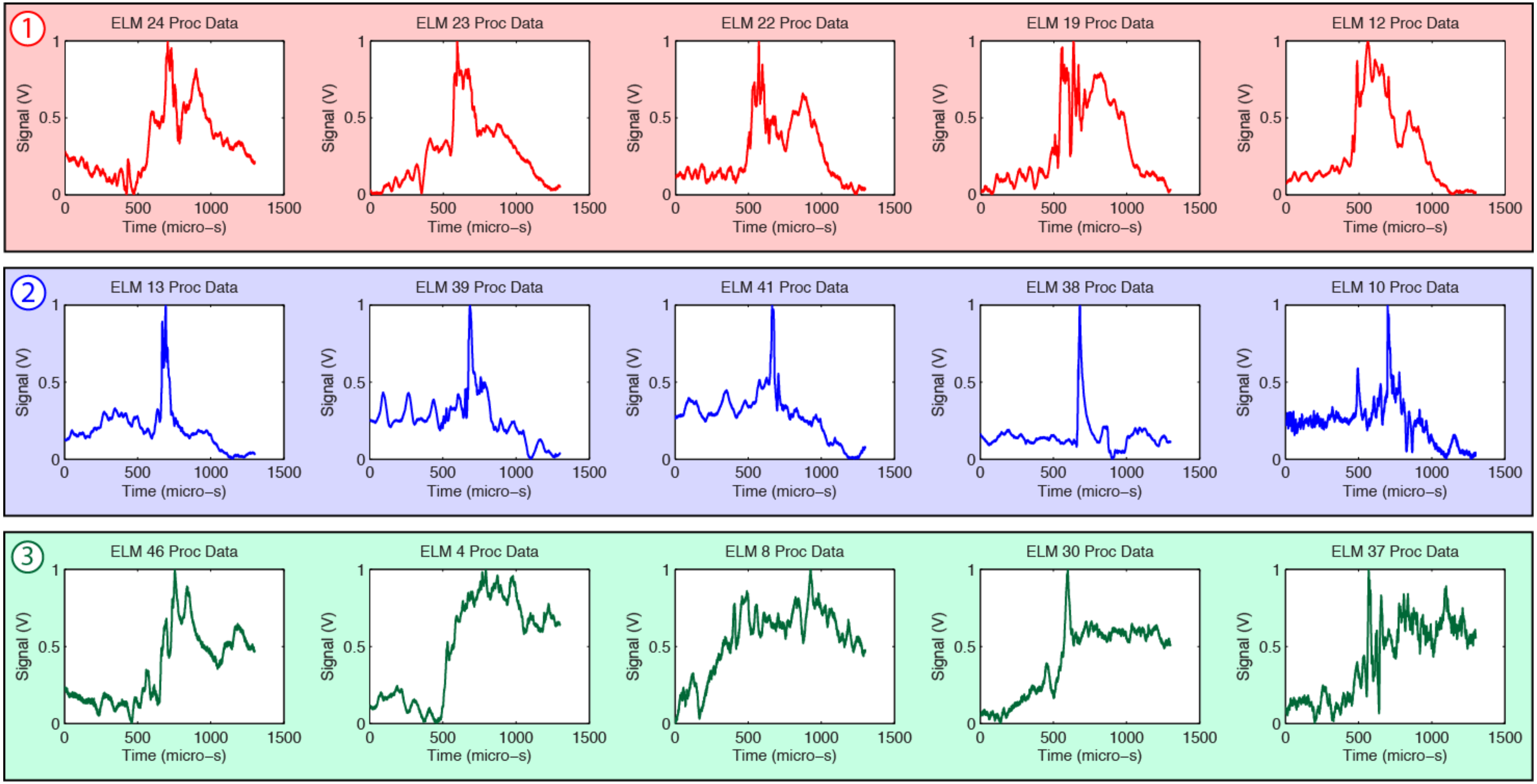
- Red, Blue, and Green groups in k-means results are **largely consistent** with previous hierarchical cluster results
- The Cyan group in k-means corresponds to poorly linked ELMs in the hierarchical cluster



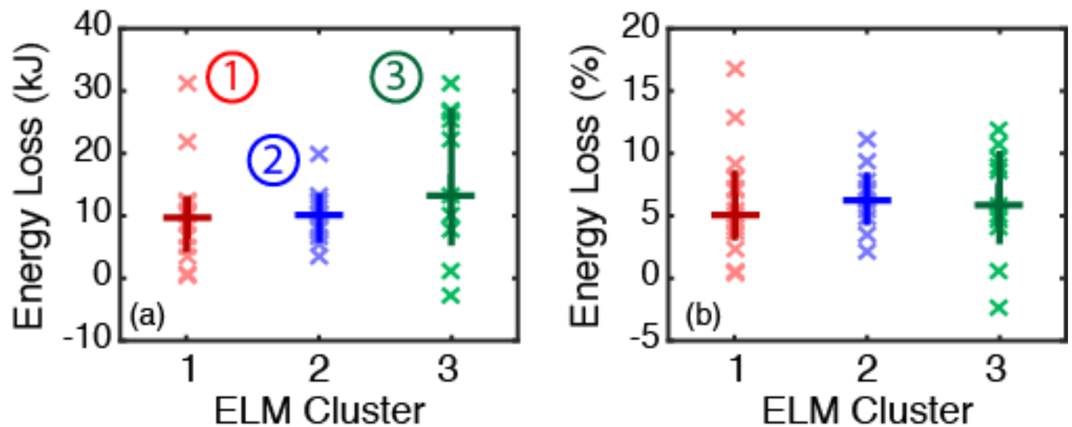
Colors correspond to k-means clusters

ELMs with similar evolution characteristics

This analysis demonstrates that machine learning techniques can independently identify patterns and similarities in time-series data.



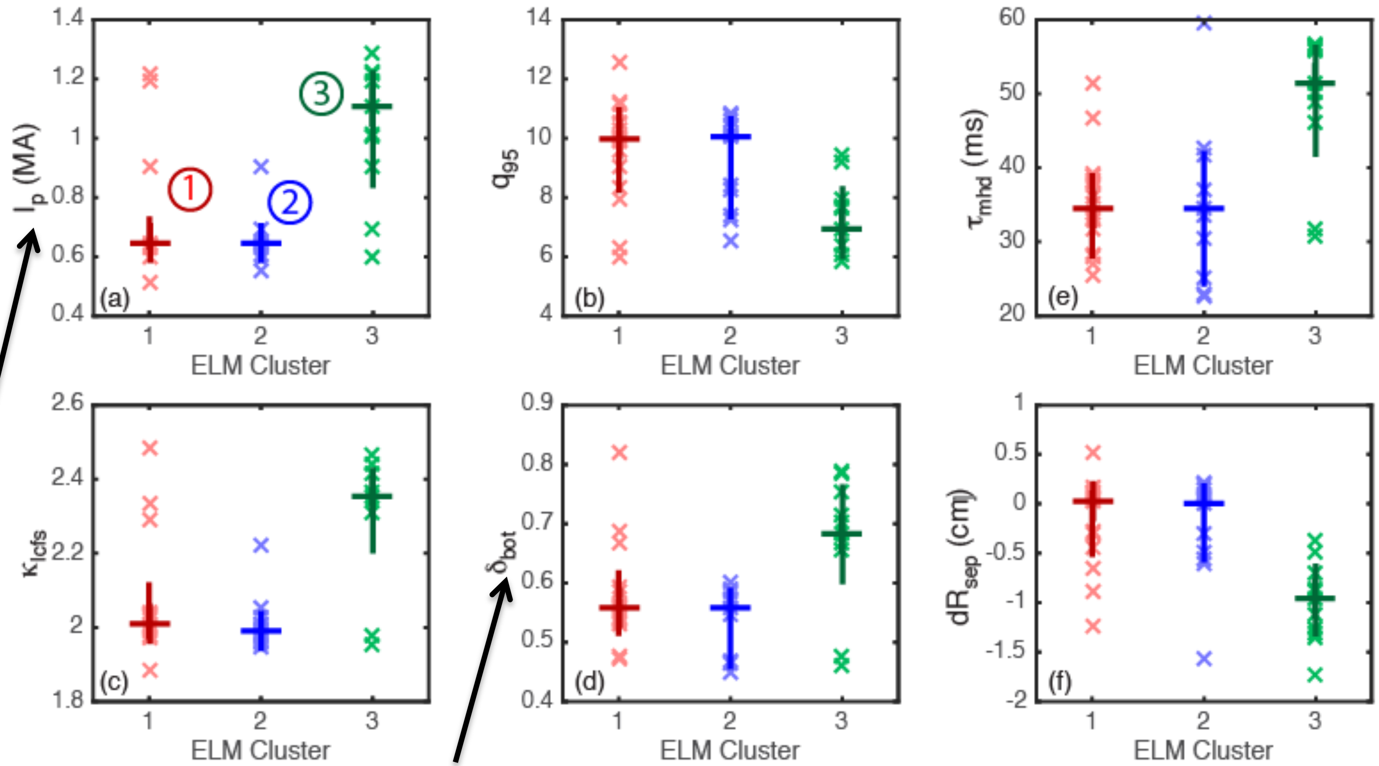
Identified ELM groups exhibit similar stored energy losses



- Stored energy loss can't differentiate the identified ELM groups

Identified ELM groups correlate with distinct parameter regimes (I)

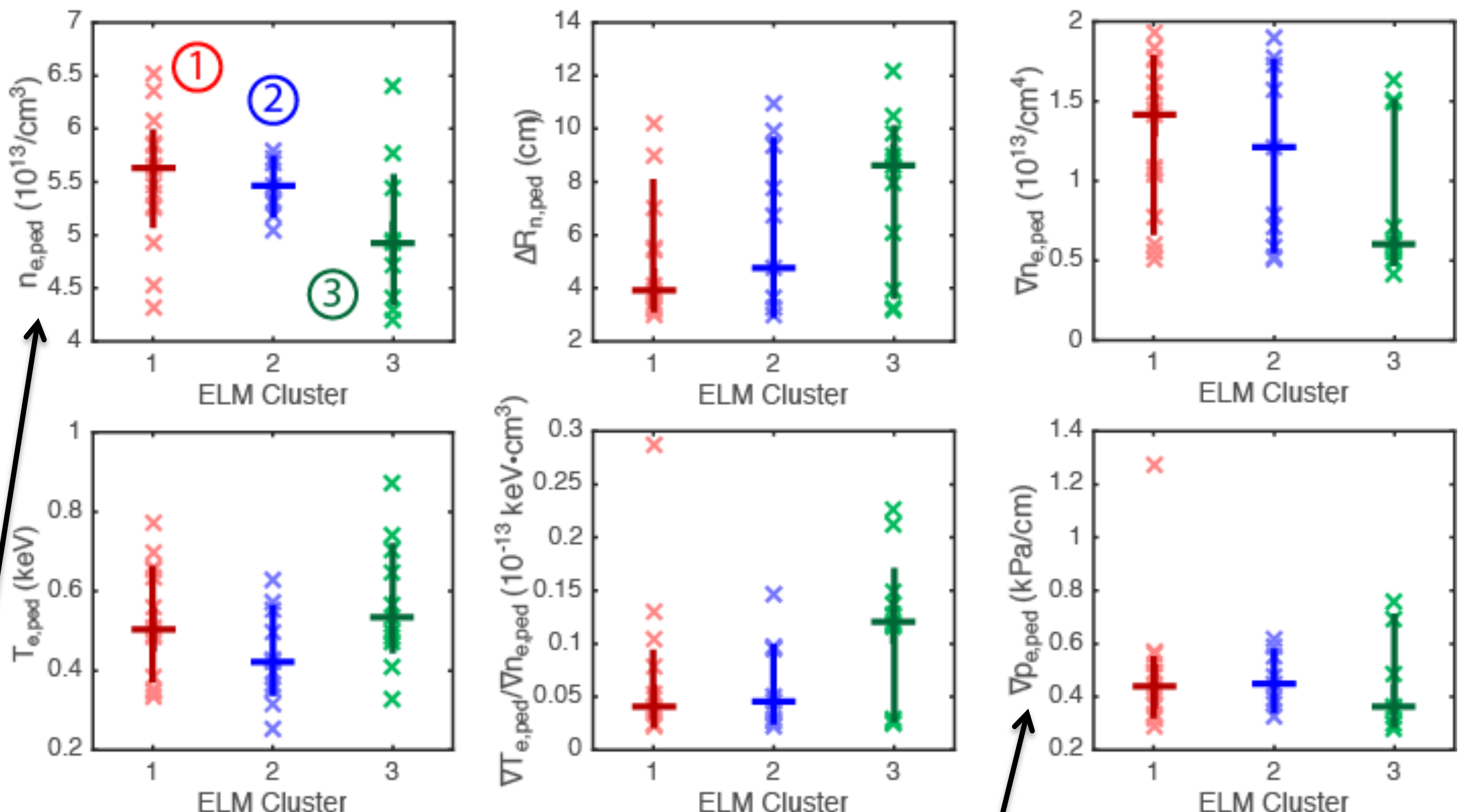
- Red and Blue ELM groups exhibit similar parameter regimes
- Observed evolution patterns and associated parameter regimes suggest genuine variations in underlying nonlinear dynamics



Higher triangularity is stabilizing for PB modes (Snyder, PPCF, 2004)

At higher Ip, most unstable mode shifts to lower n (Liu, PoP, 2014)
Higher Ip is stabilizing for PB (Snyder, PPCF, 2004)

Identified ELM groups correlate with distinct parameter regimes (II)



Pedestal density parameters appear to differentiate the ELM groups

The pressure gradient does not appear to differentiate the ELM groups

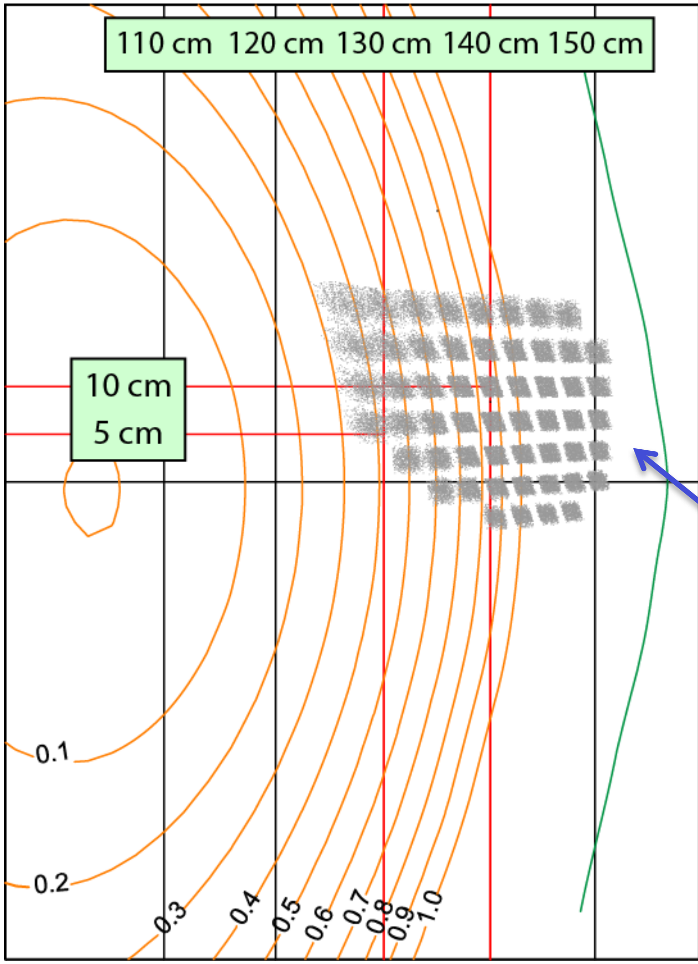
Machine learning techniques can distill insights from large datasets at national fusion facilities

- Machine learning techniques are designed for large data volumes with high dimensionality or complex dependencies
  - Improve the scientific productivity of large datasets
  - Automate analysis activities like pattern recognition, data grouping, and data reduction (scalings, classification, hypothesis testing, etc)
  - Large, diverse datasets are less susceptible to false positive results
- Data-rich fields like astronomy, neurology, seismology, and genomics have successfully leveraged machine learning
  - Cancer genomics (cf. pg. 8)
  - Exo-planet detection
  - Seismic phase/waveform classification
  - Seizure onset prediction

Summary

- Unsupervised machine learning algorithms identified groups of ELMs with similar evolution characteristics
  - BES measurements with Alfvenic time resolution capture the nonlinear evolution of ELM events
  - The identified ELM groups correlate with parameter regimes relevant to ELM physics
    - Ip, q-95, δ, dR_{sep}, $n_{e,ped}$, $\Delta R_{n,ped}$
    - Observed ELM groups can motivate nonlinear models or validation benchmarks
  - Machine learning techniques are broadly applicable to analysis activities at data-rich national fusion facilities
    - The techniques are highly scalable for automatic data classification, pattern recognition, or relationship quantification
    - Several data-rich scientific fields have successfully applied machine learning techniques to large datasets

Expanded, 2D BES for NSTX-U



- Expansion from 32 to 48 detection channels
  - New detectors, electronics, and DAQ delivered to NSTX-U
- 2D pedestal/SOL imaging
  - 54 fiber bundles in approx. 9×7 arrangement