Jet Propulsion Laboratory
California Institute of Technology

Stellar variability classification of TESS data using (un)supervised learning

Second NASA AI and Data Science Workshop for Earth and Space Sciences
February 9 - 11, 2021, virtually hosted by JPL

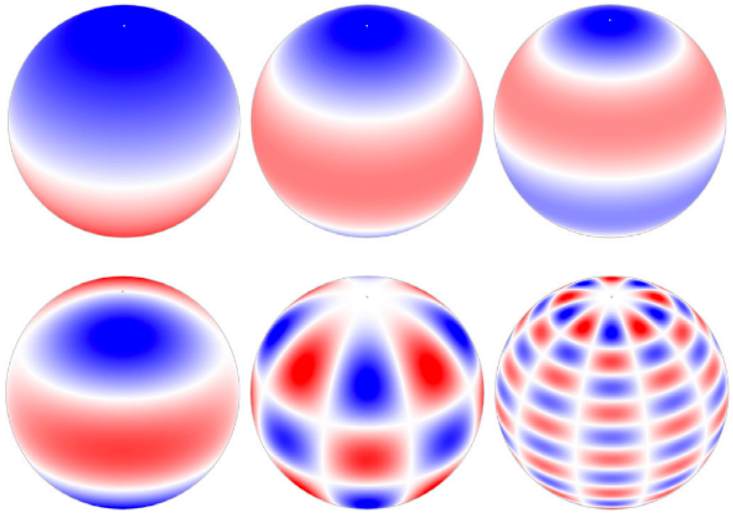
Jeroen Audenaert[1], A. Tkachenko[1] & C. Aerts[1,2,3]

[1] Institute of Astronomy, KU Leuven, Belgium
[2] Department of Astrophysics/IMAPP, Radboud University Nijmegen, Netherlands
[3] Max Planck Institute for Astronomy, Germany

jeroen.audenaert@kuleuven.be

KU LEUVEN

Stellar variability

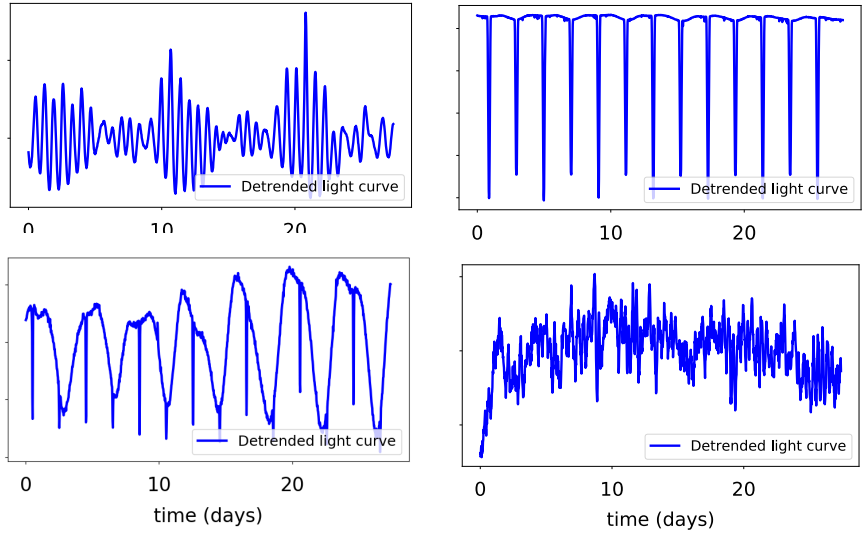


Credit: Aerts (2021)

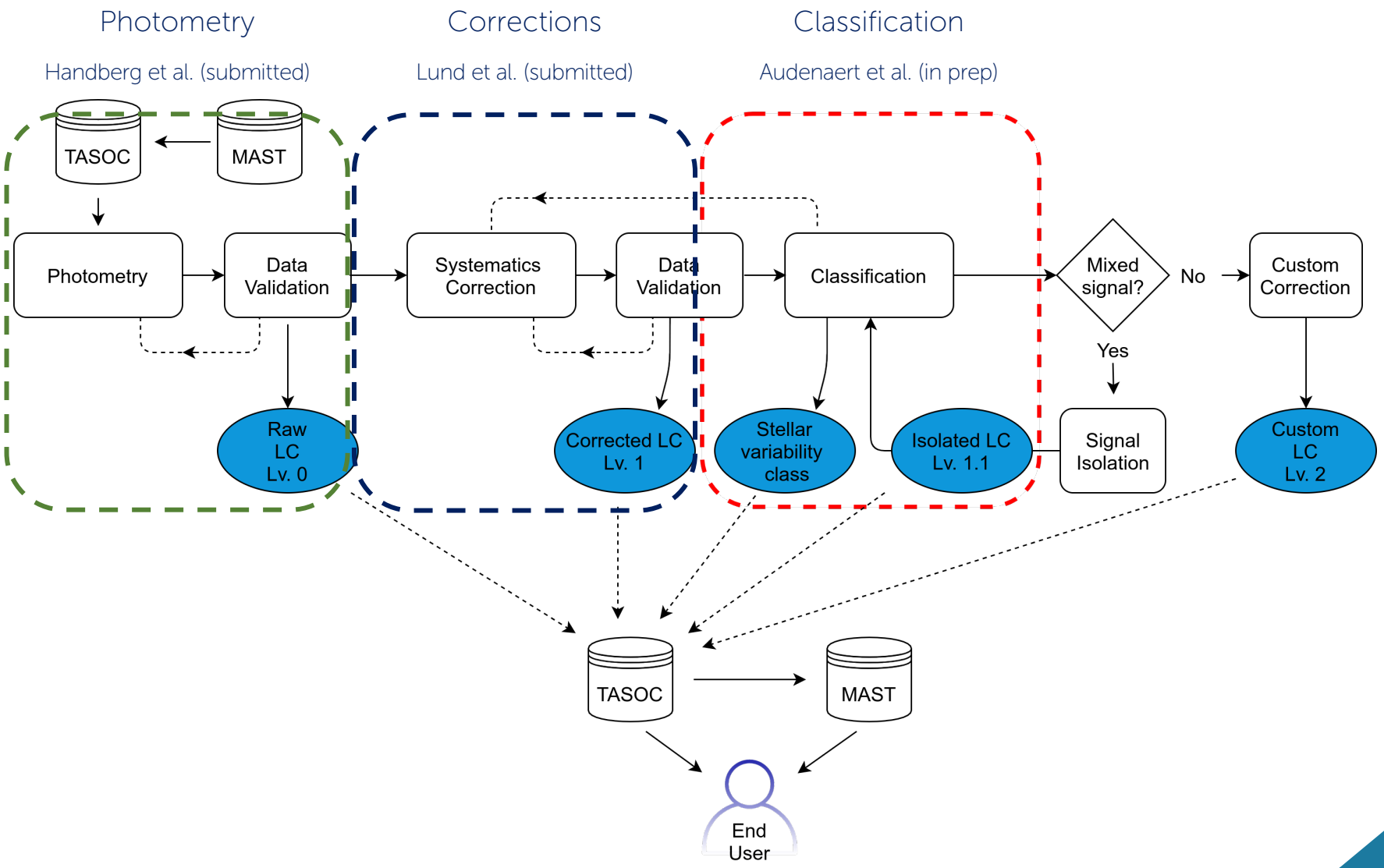
Surface of a **pulsating** star:
red and blue are hotter and cooler regions
alternating during the pulsation cycle

Credit: NASA Goddard Space Flight Center

Light curve: surface brightness over time



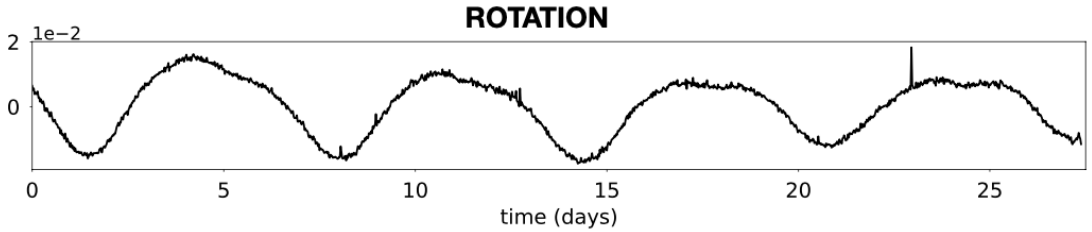
TASOC/T'DA Pipeline

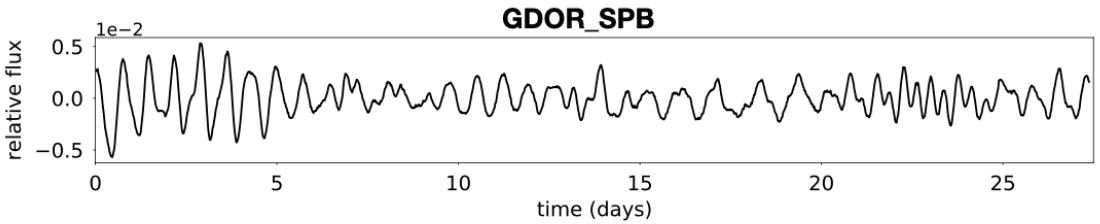


Training data

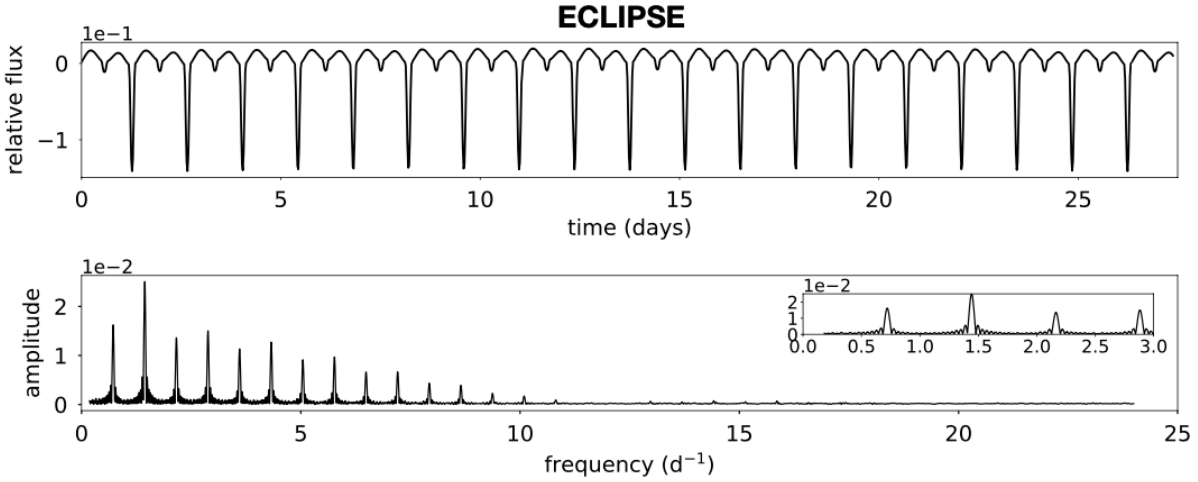
Kepler Q9

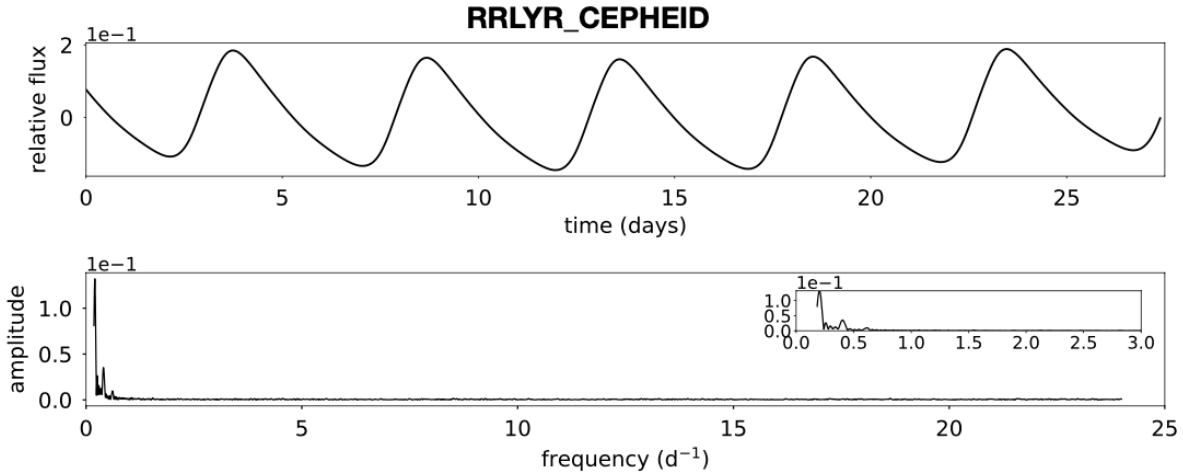
Solar-like pulsators	RR Lyraes & Cepheids	γ-Doradus & Slowly Pulsating B stars	δ Scuti & β Cephei stars	Eclipsing & Transiting	Contact binaries & rotational variables	Aperiodic stars	Constant stars
1800	62	630	772	974	2,260	830	1000

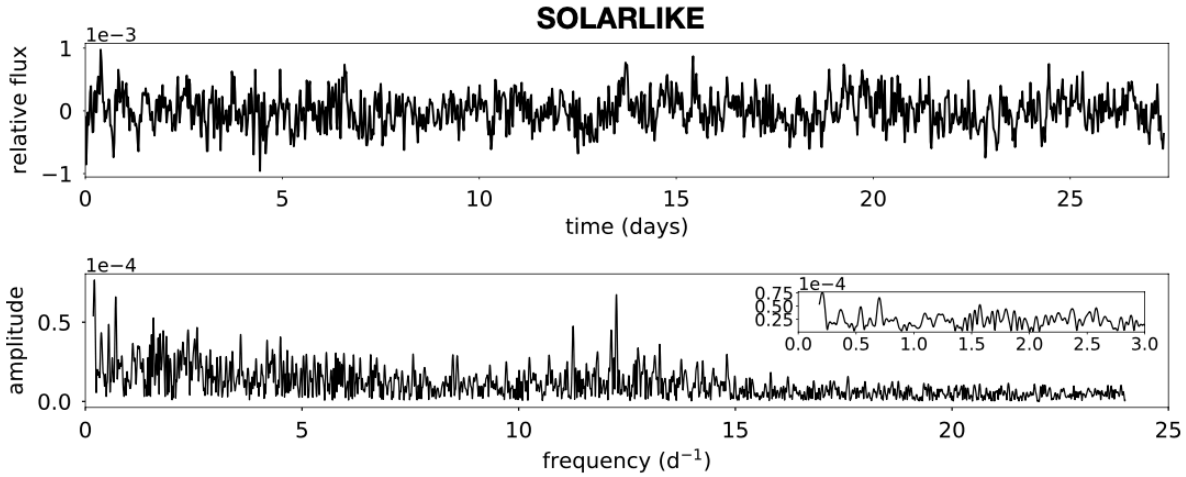


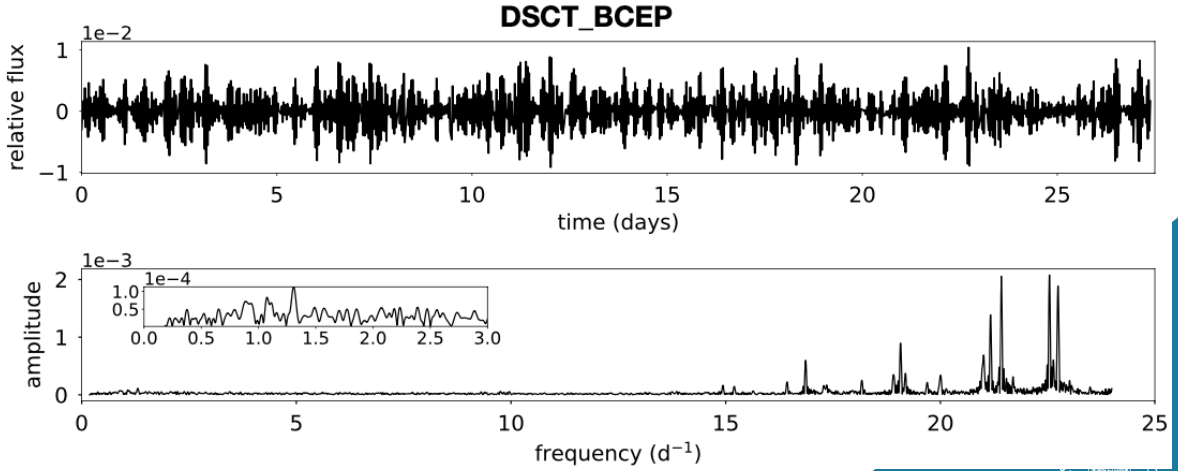


Training data

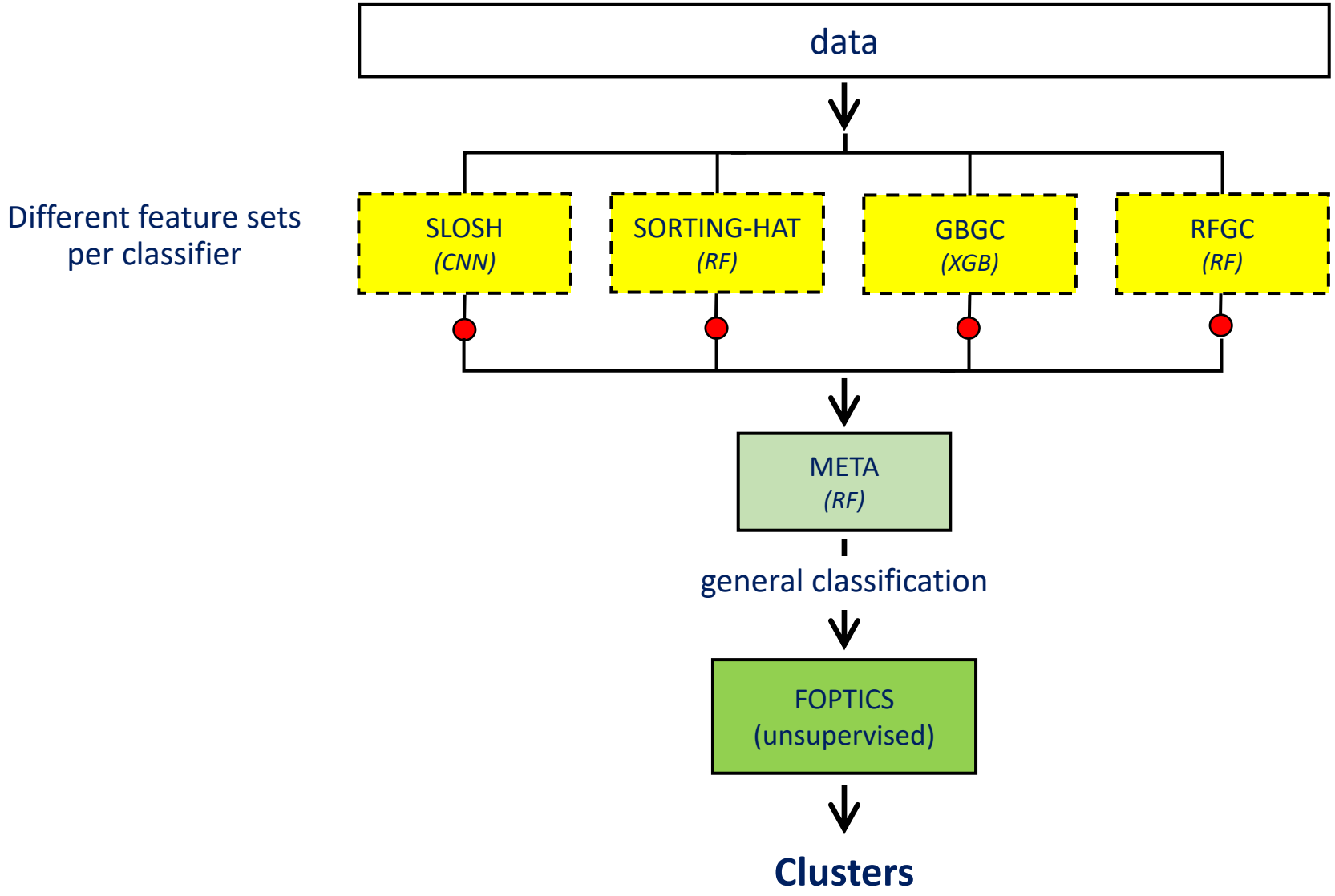






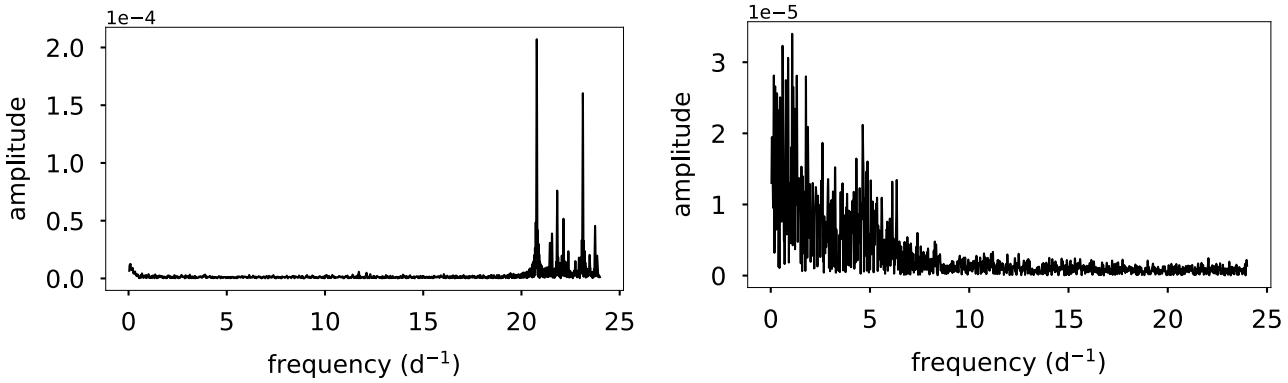


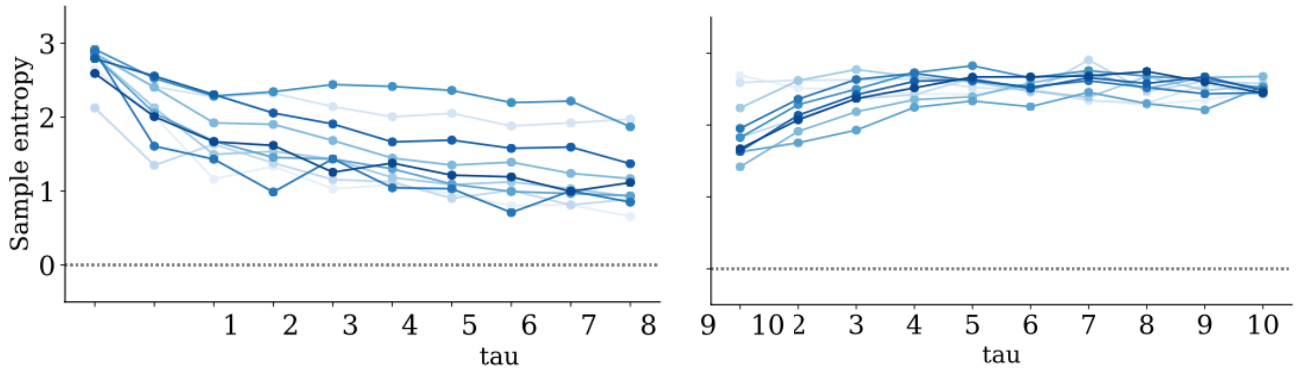
Classification scheme



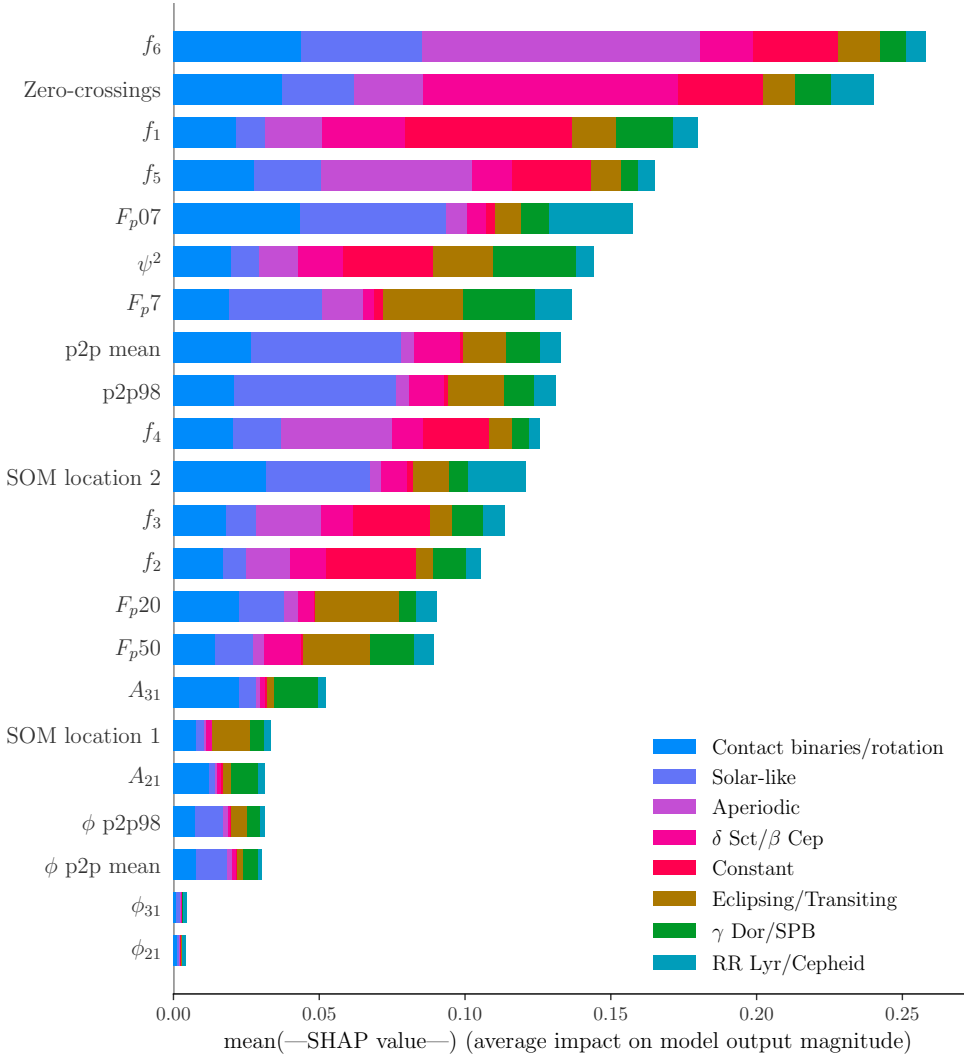
Feature engineering

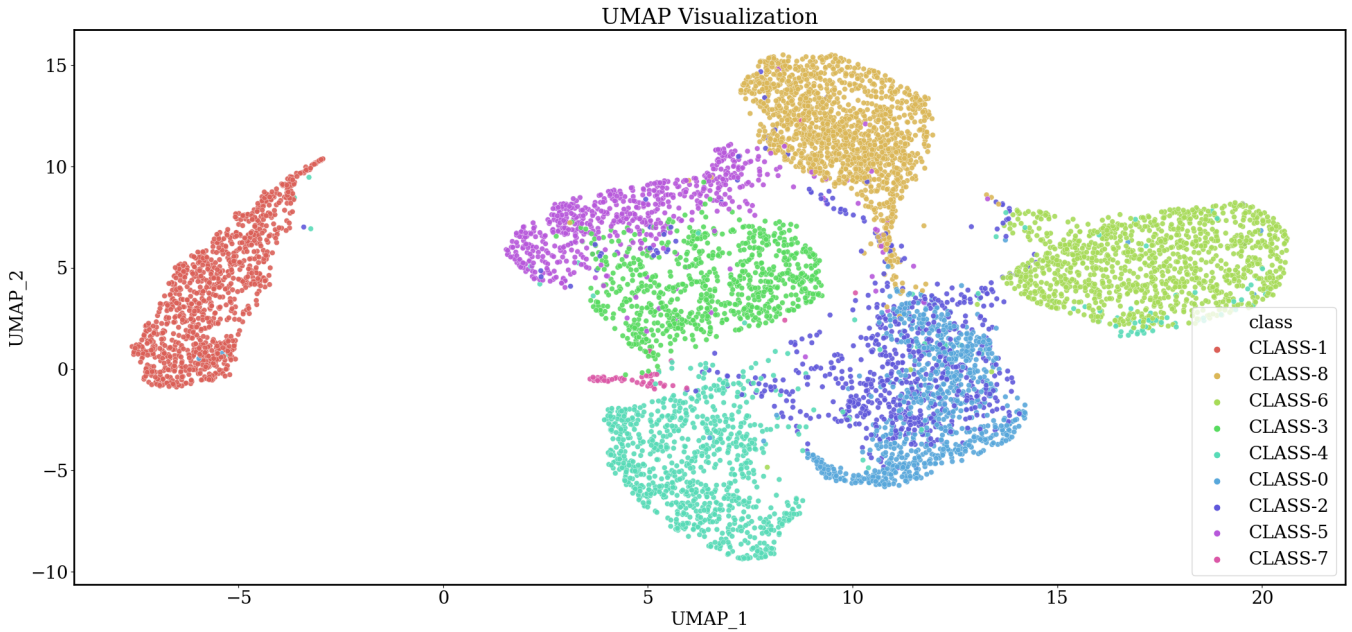
- Fourier analysis
- Statistical attributes
  - Skewness, MAD…
- Entropy analysis
  - Differential entropy
  - Multiscale entropy
- SOM



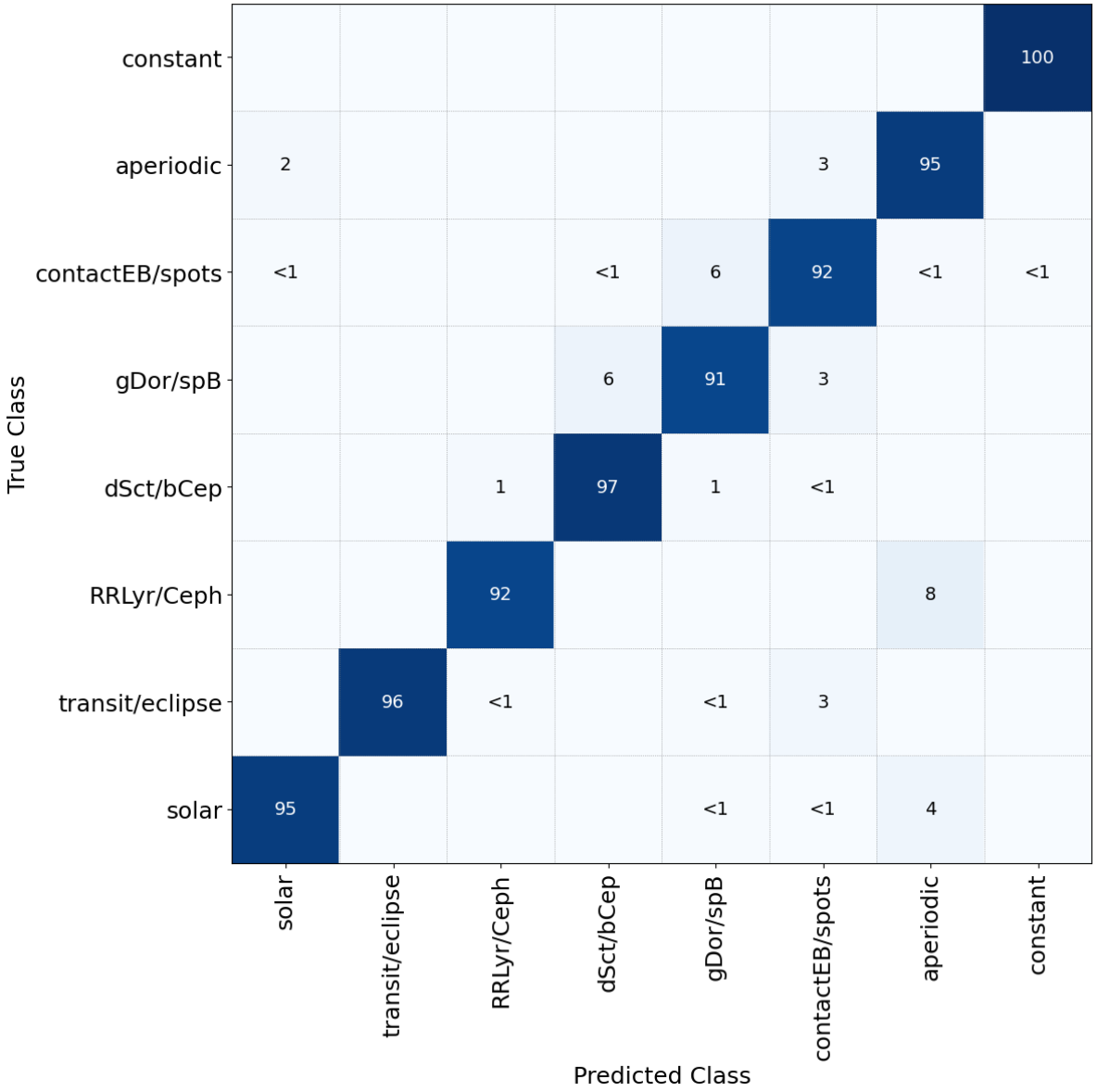


Feature engineering

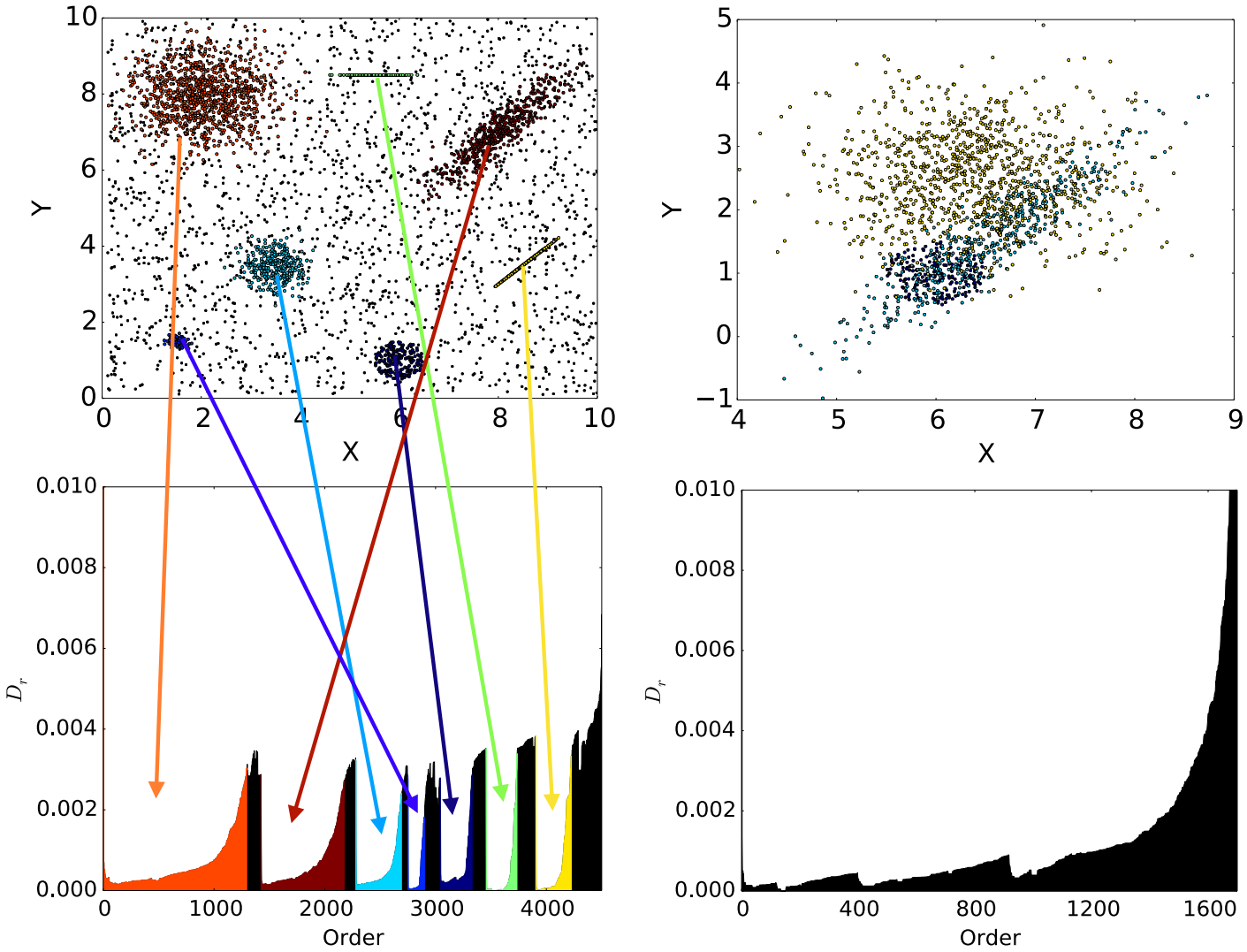




Supervised results (meta)

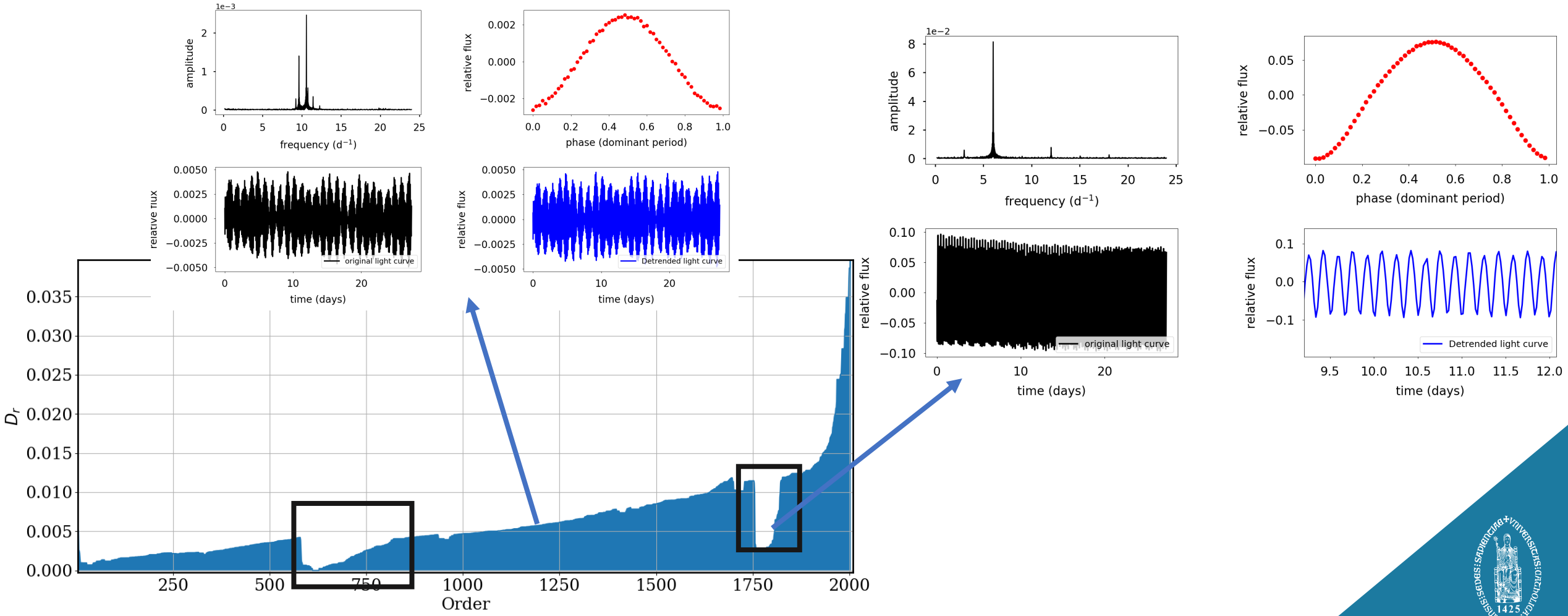


FOPTICS Clustering (*Fuzzy* Ordering Points To Identify the Clustering Structure)

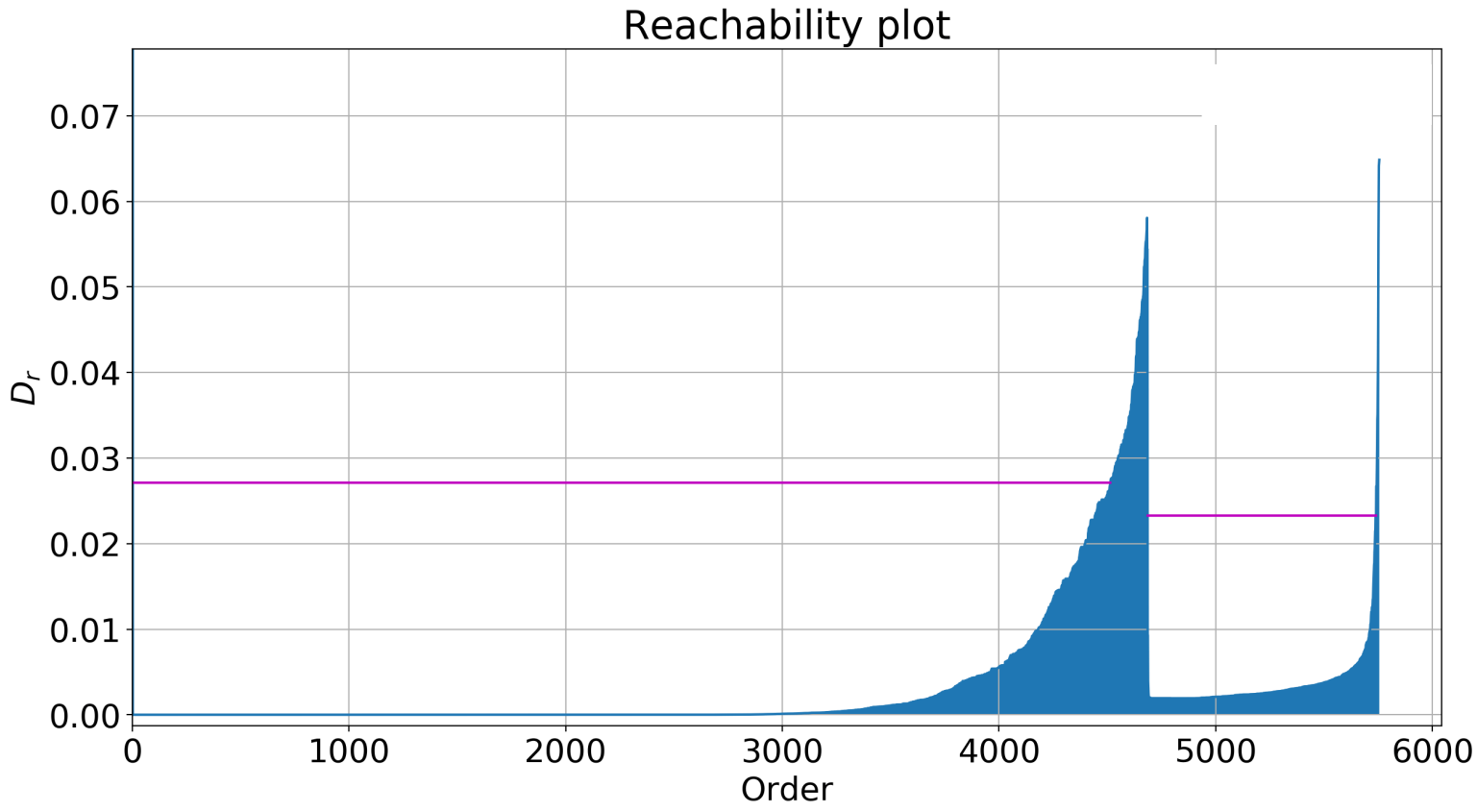


Ankerst+ (1999), Kriegel+ (2005)

FOPTICS - example



FOPTICS - cluster extraction



Brecheisen+ (2004)

Conclusion

- Combine power of supervised & unsupervised learning in one framework
  - High-level classification through ensemble of supervised models
    - General labelled classification of stars
  - Detailed clustering through FOPTICS
    - Characterize substructure
    - Discover new classes
    - Identify misclassifications

Jeroen Audenaert
jeroen.audenaert@kuleuven.be
KU Leuven, Institute of Astronomy
Celestijnenlaan 200D bus 2401
3001 Leuven
Belgium

KU LEUVEN