Real-Time Disruption Predictor Based on Anomaly Detection

J. Vega[1],
on behalf of
A. Pereira[1], R. Moreno[1], A. Murari[2], S. Dormido-Canto[3], S. Esquembri[4] and JET Contributors*

EUROfusion Consortium, JET, Culham Science Centre, Abingdon, OX14 3DB, UK
[1]Laboratorio Nacional de Fusión, CIEMAT, Madrid, Spain
[2]Consorzio RFX, Padua, Italy
[3]Departamento de Informática y Automática, UNED, Madrid, Spain
[4]Technical University of Madrid (UPM), Madrid, Spain
*See the Appendix of F. Romanelli et al., Proc. of the 25th IAEA Fusion Energy Conference 2014, Saint Petersburg, Russia

Acknowledgements
This work was partially funded by the Spanish Ministry of Economy and Competitiveness under the Projects No ENE2012-38970-C04-01, ENE2012-38970-C04-03 and ENE2012-38970-C04-04.

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement number 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

Motivation

- Experimental nuclear fusion devices like ITER need to explore advanced operation scenarios to achieve high performance plasmas
  - Plasmas near operational limits can produce disruptions

- Detrimental effects of disruptions to fusion devices
  - Mitigation actions

- A reliable real-time disruption predictor is a pre-requisite to any mitigation method
  - Reliability has to be understood in terms of high success rate, low false alarm rate and enough anticipation time

Motivation

- Machine learning methods can be used to create data-driven models to explain behaviours whose mathematical description from first principles is not possible
  - Example: disruption prediction

- Typically, disruption predictors have been implemented as classification systems
  - The more training samples the better
    - APODIS in JET (J. Vega et al. Fus. Eng. Des. 88 (2013) 1228-1231)
    - The need of large training datasets is a drawback for ITER and DEMO
  - Predictors from scratch are a potential alternative
    - Predictors learn from the first disruption in an adaptive way
      - S. Dormido-Canto et al. Nuclear Fusion. 53 (2013) 113001 (8pp)
      - J. Vega et al. Nuclear Fusion. 54 (2014) 123001 (17pp)
    - They also need past discharges (a small fraction) to be optimized

- Could disruption predictors be developed without the need of training with past discharges?
  - This means to learn the '*normal*' evolution of a discharge and to trigger an alarm when an anomalous behaviour is detected
  - How sure are we about the fact that the anomaly detected corresponds to a disruption?

Outline

- Anomaly detection concept
- Disruption prediction based on anomaly detection and the locked mode signal
- Results in JET with 2304 discharges with the ILW
- Conclusions

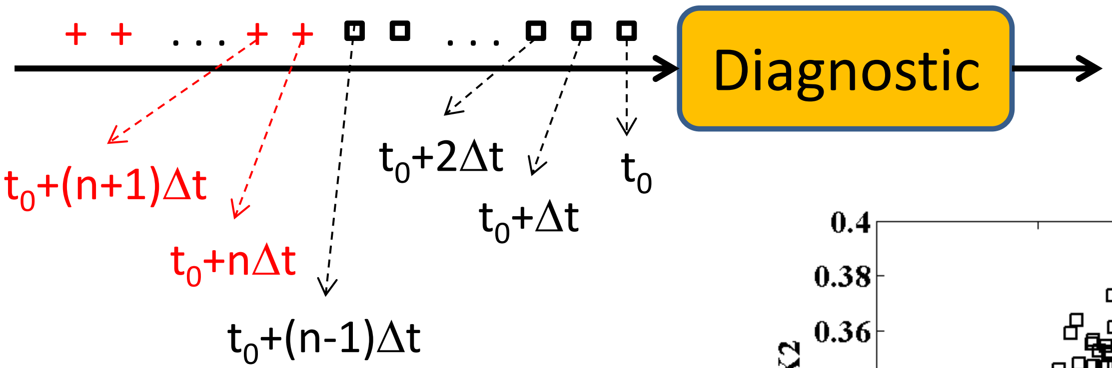
GOBIERNO DE ESPAÑA MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD
Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Anomaly detection: conceptual view

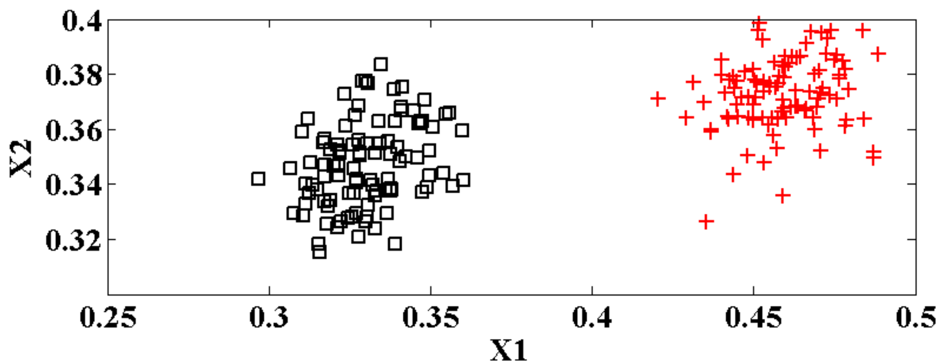
- Objective: to identify when the sample distribution changes (in the widest sense)
  - Possible changes: distribution parameters, the distribution itself, noise amplitude, …

- The aim **is not** to make hypothesis testing to determine the specific distributions but recognizing asap when the sample distribution is different

Algorithm for anomaly recognition



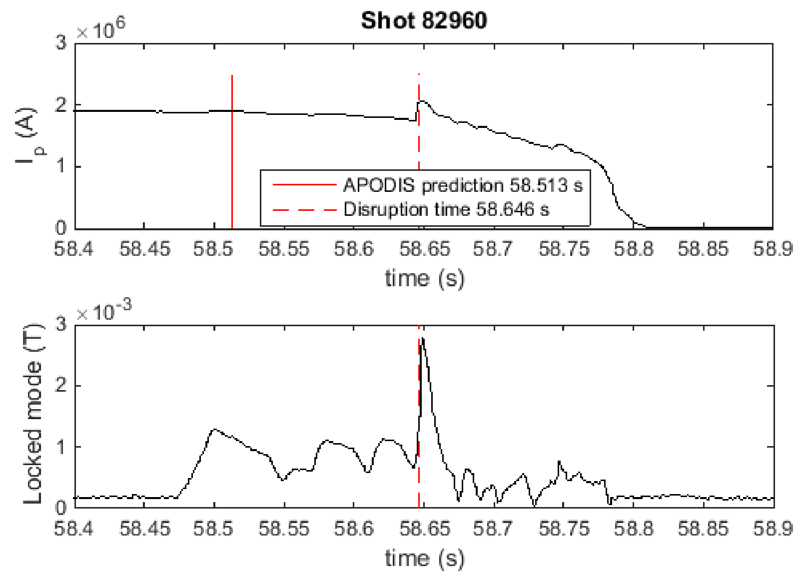
The detection process has to recognize that the samples start to be '*far*' from the initial cluster and, therefore, an alarm has to be triggered



Each particular application can define a specific mapping

Simple Predictor based on Anomaly Detection (SPAD): on-line setting

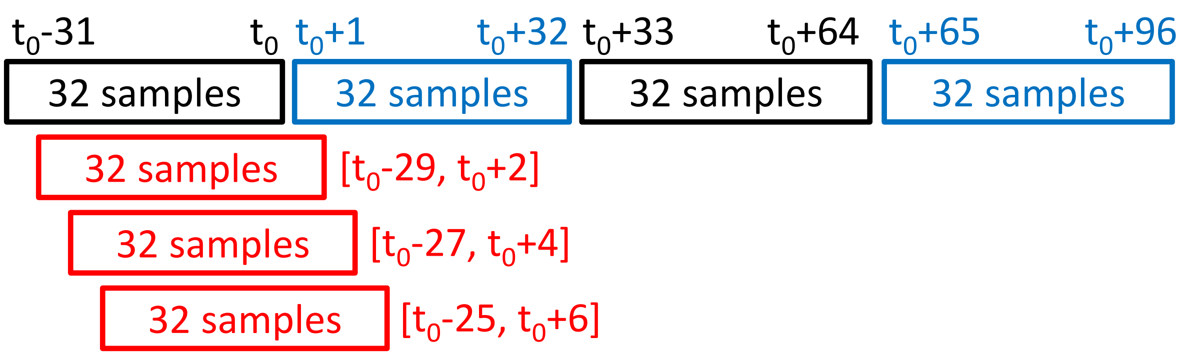
- Main advantage: no data from past discharges is needed
  - A new predictor is started with each new discharge
- Requirements
  - The delay between a true change and its detection should be minimal
  - The number of missed detections should be minimal
  - The number of false detections should be minimal
  - Data streams should be handled efficiently
  - The sequential data are read only once
- How sure are we about the fact that the anomaly detected corresponds to a disruption?
  - Feature selection to represent the plasma state is essential
  - The simplest SPAD predictor can be developed with a locked mode signal
    - Not basing the prediction only on an amplitude threshold



The JET LM predictor based on amplitude threshold misses the alarm

SPAD implementation

- Outlier recognition criterion
  - Data have to be processed in time windows to avoid ignoring the frequency domain
- Temporal resolution
  - Time windows: 32 ms
  - Sampling rates: 1 kS/s
  - A sliding window mechanism can be used to achieve a resolution of ms without increasing sampling rates



- In each time window, the data are compressed into two components by means of the Haar wavelet transform (approximation coefficients)

Haar wavelet transform

- A wavelet transform is a filter
  - Low pass/high pass

S(t) (N samples)

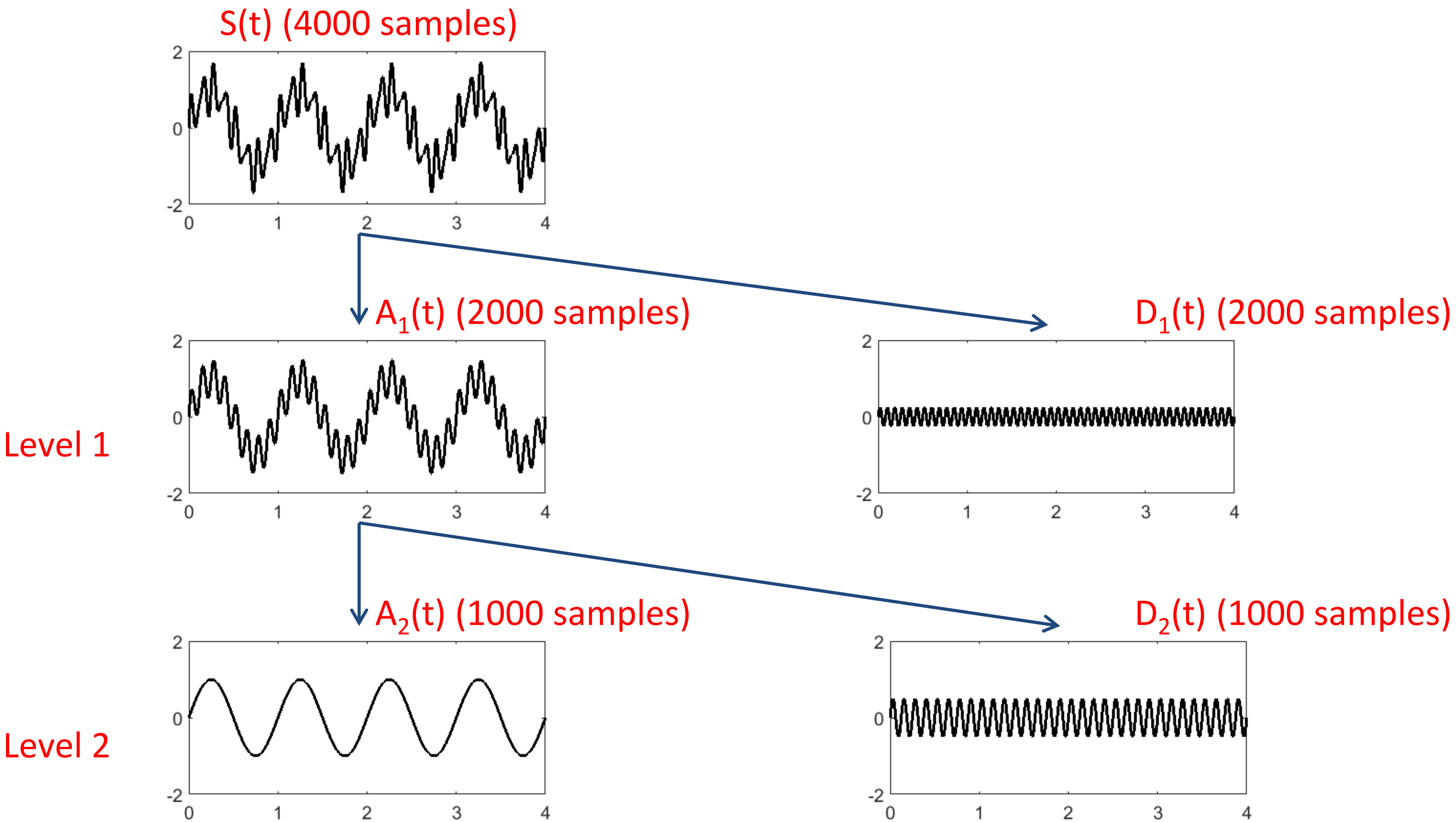
Level 1 $A_1(t)$ (N/2 samples) $D_1(t)$ (N/2 samples)

Level 2 $A_2(t)$ (N/4 samples) $D_2(t)$ (N/4 samples)

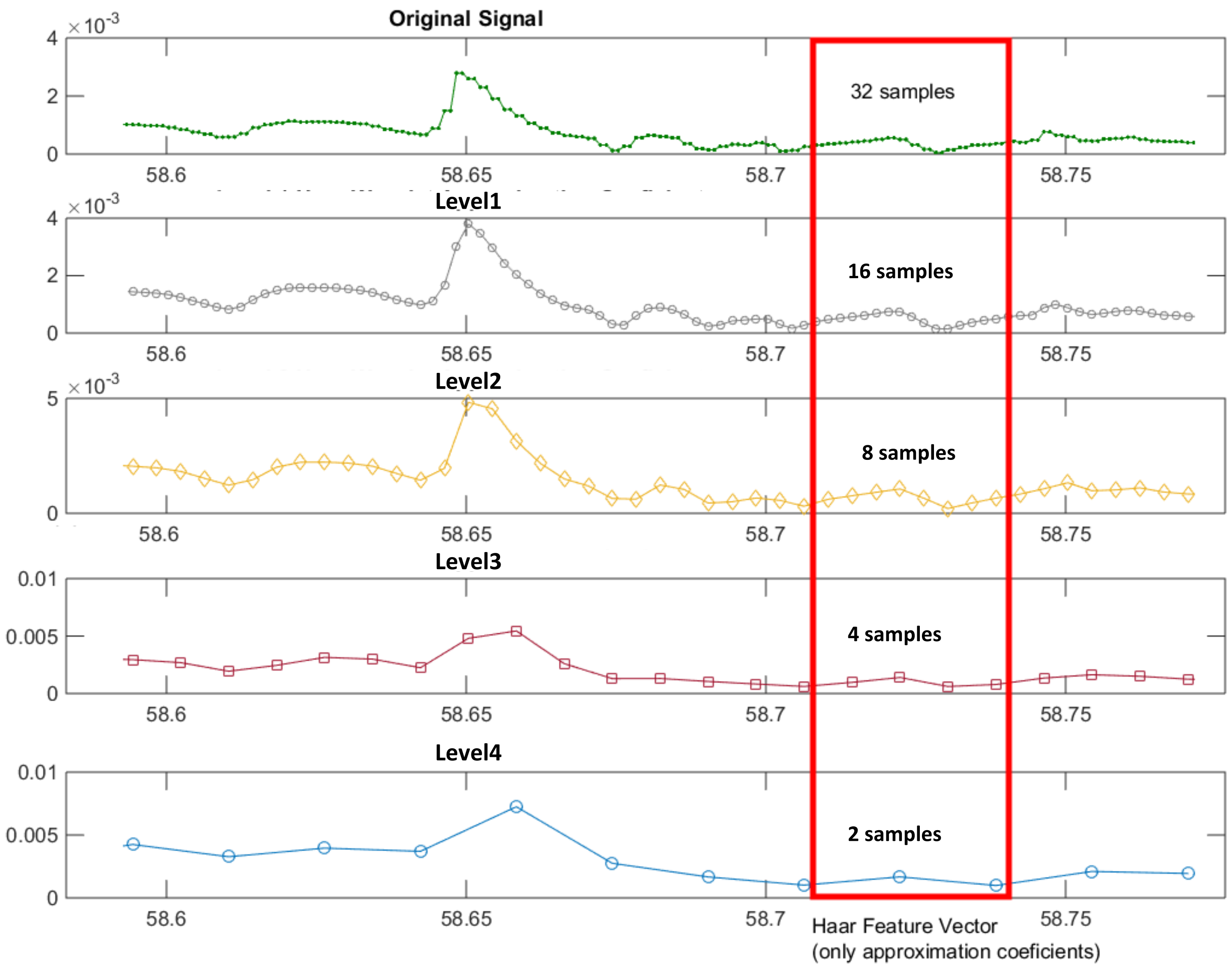
Level 3 $A_3(t)$ (N/8 samples) $D_3(t)$ (N/8 samples)

. . . . .

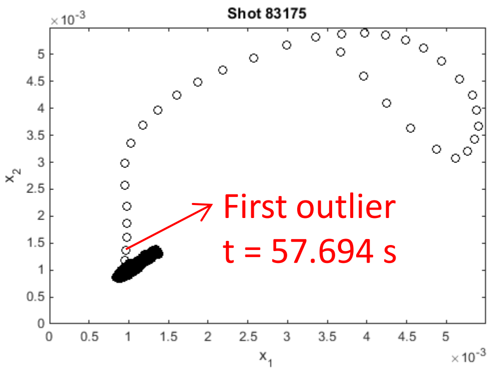
Wavelet transform

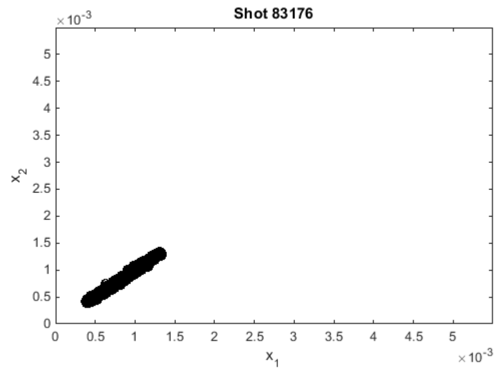


Haar wavelet: approximation coefficients



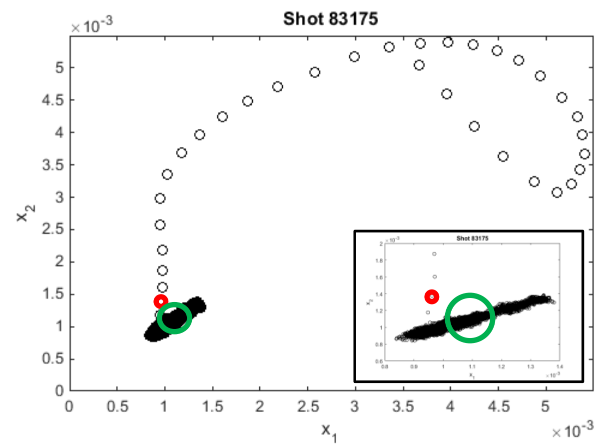
SPAD implementation





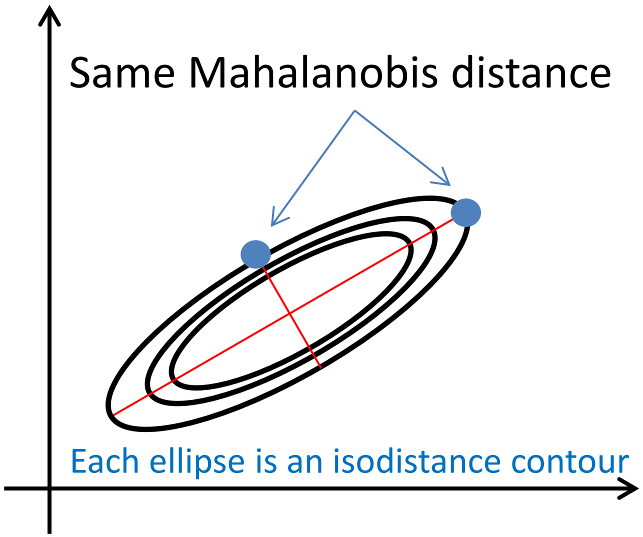
Scatterplots in the bi-dimensional space. Points are represented every 2 ms

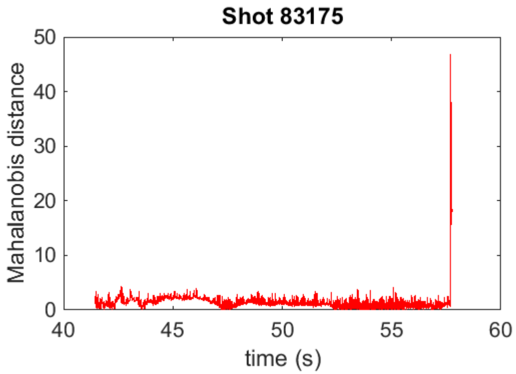
- In the non-disruptive phases of the discharges, the points in this bi-dimensional space show a compact cluster structure
- The alarm has to be triggered the first time that a point is 'far enough' from the cluster center
  - The Euclidean distance does not seem to work
  - The distances defined by the Euclidean metric take no account of any patterns of covariance that exist in the data
- A simple inspection of the cluster data shows a positive covariance in the data

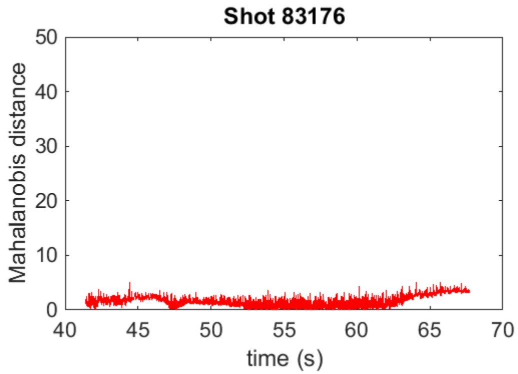


SPAD implementation

- The metric proposed by Mahalanobis does adjust for covariance
- The computation of the Mahalanobis distance is carried out every 2 ms through the wavelet transform of the latest 32 samples of the locked mode signal





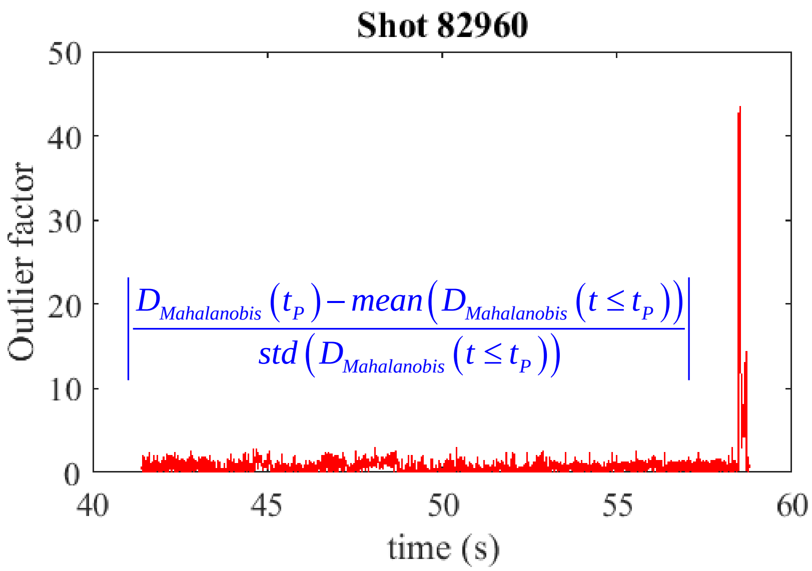


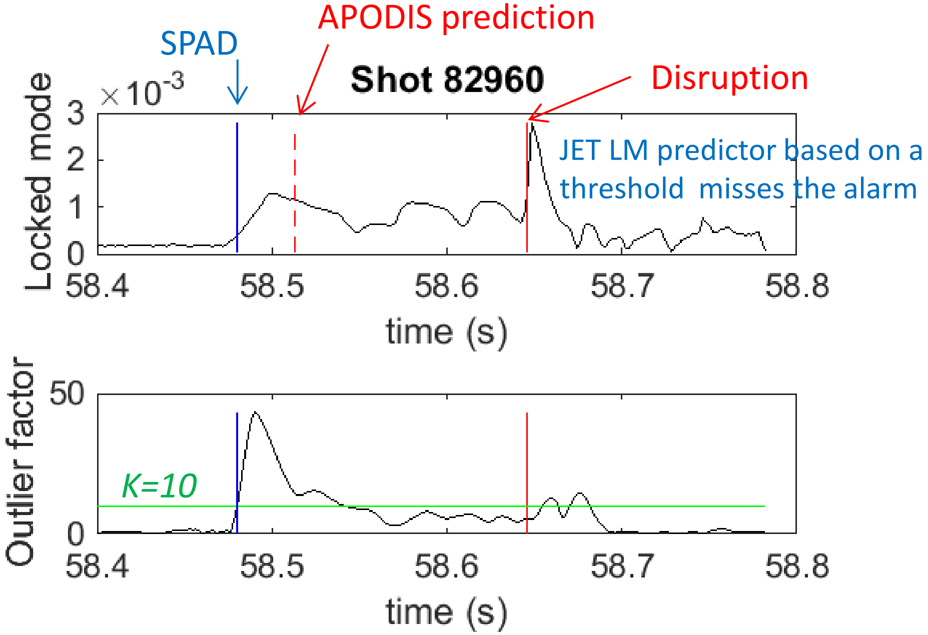
- Where is the distance limit to recognise $X(t_P)$ as outlier?
  - Outlier criterion

$$\left| \frac{D_{Mahalanobis}(t_P) - mean\left(D_{Mahalanobis}(t \le t_P)\right)}{std\left(D_{Mahalanobis}(t \le t_P)\right)} \right| \ge K$$

SPAD implementation

- In this first version *K* = 10
- In future versions, *K* could be determined ‘on-the-fly’ during each running discharge





J. Vega, R. Moreno, A. Pereira et al. “Advanced disruption predictor based on the locked mode signal: application to JET”. 1st EPS Conference on Plasma Diagnostics. April 14-17, 2015. Frascati, Italy

SPAD results in JET

The information content in the 32 samples of the time windows can be compressed with the wavelet transform in 2, 4, 8 or 16 points

Data compression	False alarms (%)	Missed alarms (%)	Tardy detections (%)	Valid alarms (%)	Premature alarms (%)
2	7.13	13.43	3.53	81.45	1.59
4	7.31	11.48	3.36	83.22	1.94
8	7.42	11.84	3.00	83.39	1.77
16	+18%	12.37	3.71	81.80	2.12

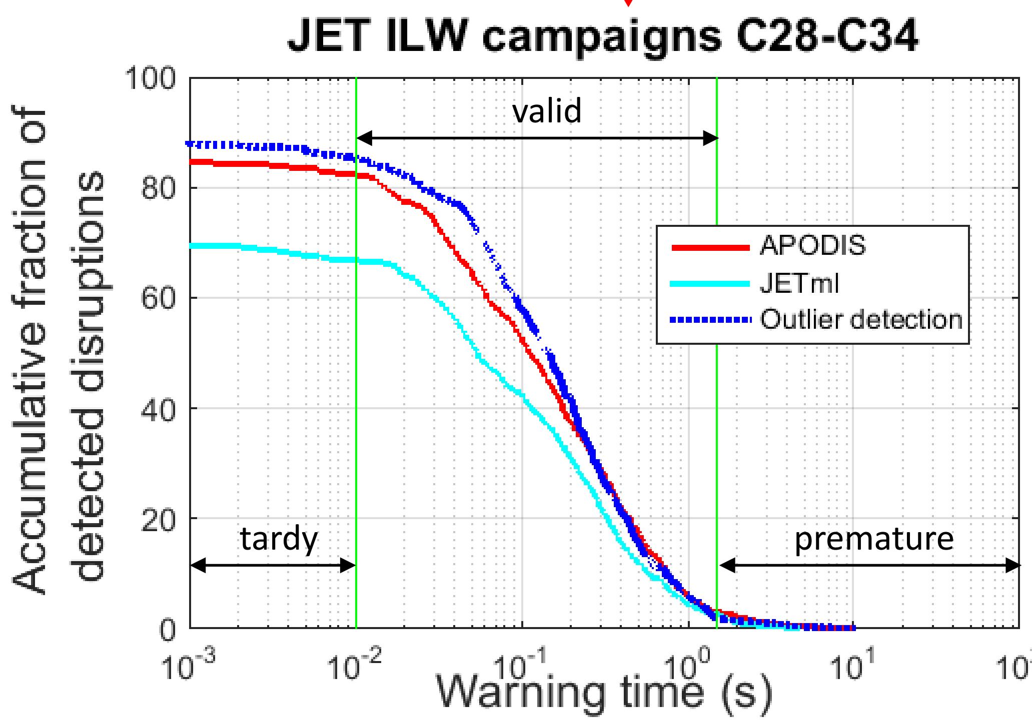
Missed alarms: no alarm or alarm triggered after the disruption

Tardy detection: warning time < 10 ms

Premature alarms: warning time > 1.5 s

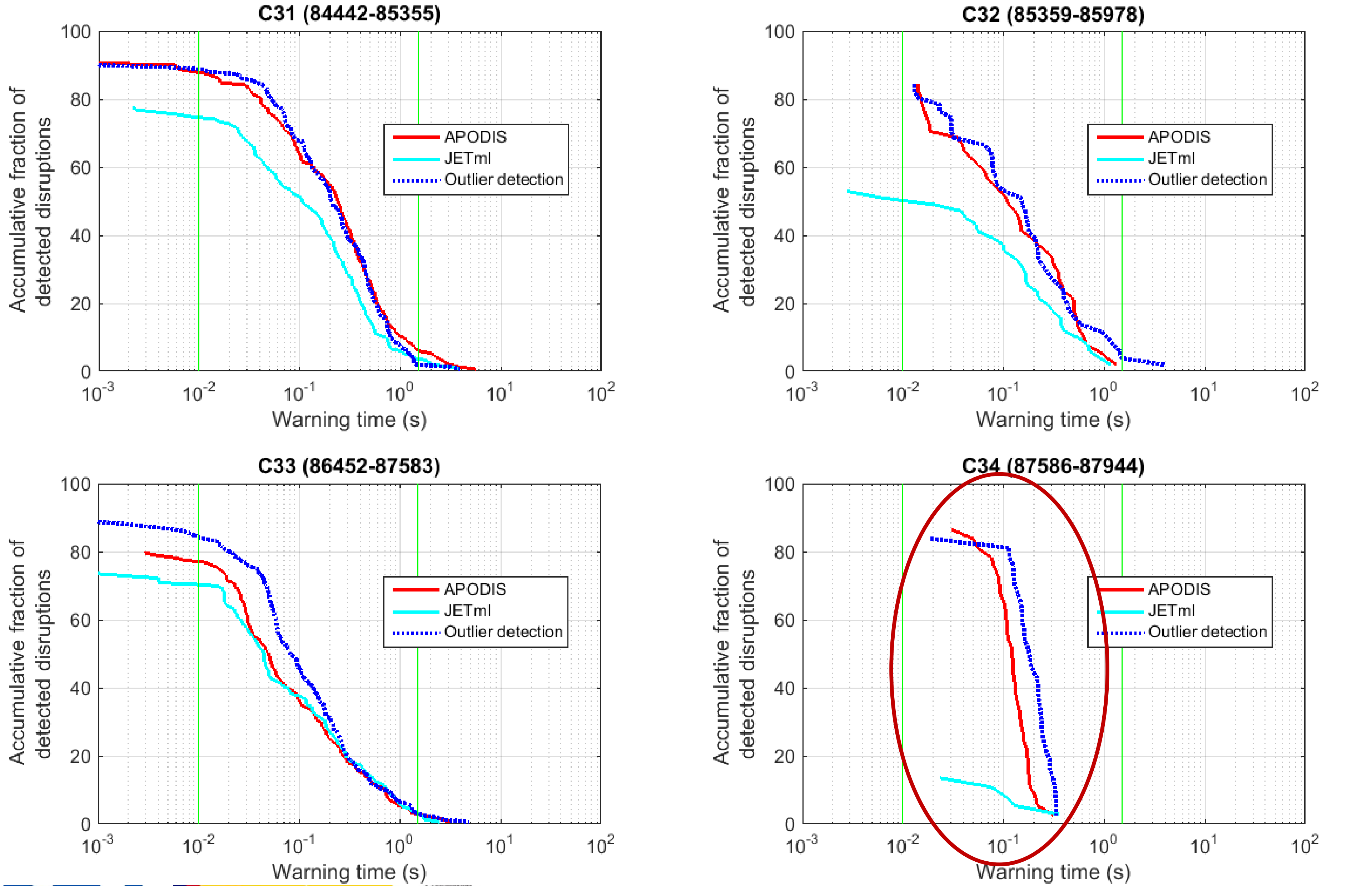
Valid alarms: 0.01 s ≤ warning time ≤ 1.5 s

- All safe discharges and all unintentional disruptions in the range 82460-87918 (all ILW experimental campaigns) have been considered
  - 1738 non-disruptive discharges
  - 566 unintentional disruptive discharges



JET GOBIERNO DE ESPAÑA MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

SPAD results by campaigns



Conclusions (1/2)

- A disruption predictor based on anomaly detection has been developed and tested
  - ILW campaigns: C28-C34 (+2300 discharges)

- The anomaly detection has been implemented within each discharge
  - No previous data are needed for training purposes
  - Discover anomalies through outliers
  - To be sure that the anomaly corresponds to a disruption: locked mode

- It is only based on the locked mode signal
  - Simplicity
  - It requires a specific data processing and it is not based on an amplitude threshold

Conclusions (2/2)

- The computations required are fast enough to install the predictor under real-time requirements
  - SPAD can coexist with other predictors

- SPAD outperforms both JET APODIS and the JET LM predictors
  - However, improvements are possible

- SPAD is a predictor that uses a single signal in the time domain linked to physical phenomenology, which is well understood
  - In line with ITER requirements

Does this predictor work in NSTX?