Comparison of unsupervised methods to determine common patterns in the termination phase of disruptive discharges in JET

J. Vega[1], R. Dormido[2], S. Dormido-Canto[2], G. Rattá[1], D. Gadariya[1], A. Murari[3] and JET Contributors[4]

[1]Laboratorio Nacional de Fusión, CIEMAT, Madrid, Spain
[2]Dpto. Informática y Automática - UNED, Madrid, Spain
[3]Consorzio RFX (CNR, ENEA, INFN, Universitá di Padova, Acciaierie Venete SpA), Corso Stati Uniti 4, 35127 Padova, Italy
[4]See the author list of 'Overview of JET results for optimising ITER operation' by J. Mailloux et al. to be published in Nuclear Fusion Special issue: Overview and Summary Papers from the 28th Fusion Energy Conference (Nice, France, 10-15 May 2021)

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

This work was partially funded by the Spanish Ministry of Economy and Competitiveness under the Projects No. PID2019-108377RB-C31 and PID2019-108377RB-C32

Motivation

Even today, in fusion, most databases are built manually, which can have various drawbacks (manpower, human errors, …)

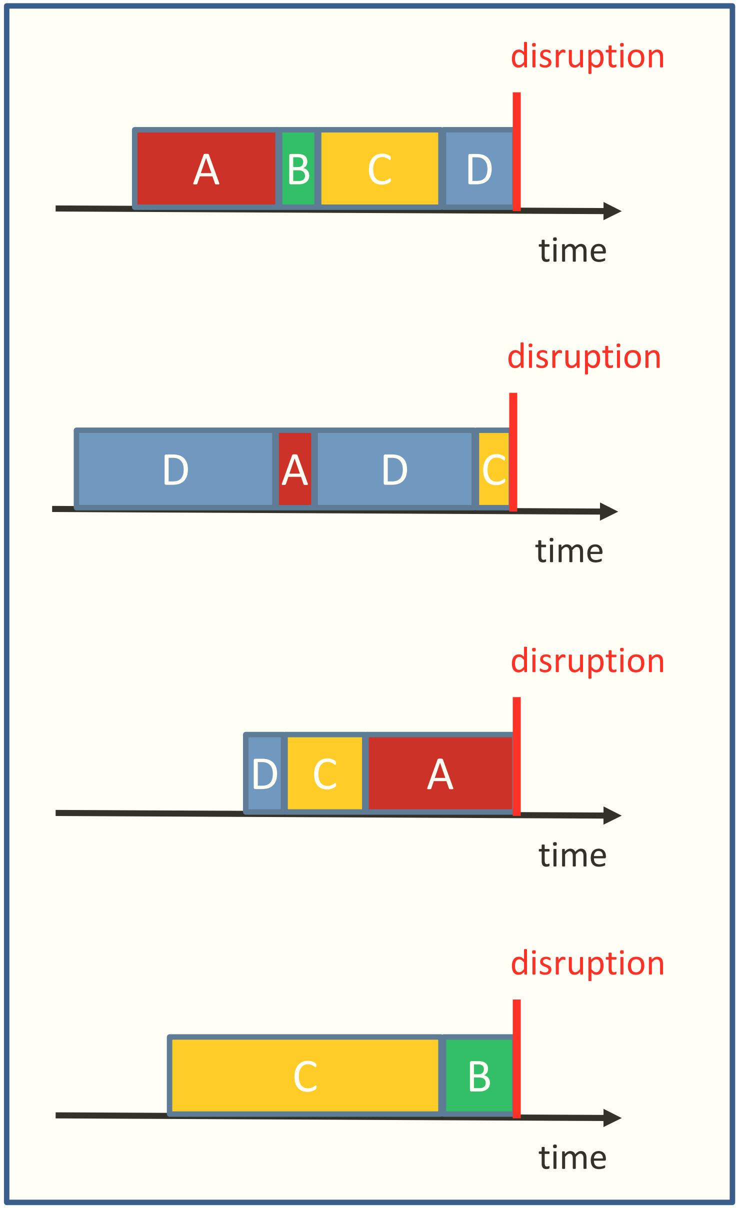
- Determining the root cause of a disruption can be a difficult and time consuming task
  - If the root cause is not very clear, in depth visual data analysis is required to look for specific signal patterns
- Can we characterise the '*termination phase*' of disruptive discharges by the evolution of signal patterns?
  - This would allow the automatic classification of disruption types
- Machine learning methods by means of unsupervised classification techniques can contribute significantly
  - The objective is to discuss two different approaches based on two different mathematical foundations that provide coherent results

GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA E INNOVACIÓN Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Motivation

Final goal: label evolution in the last part of disruptive discharges

- Finding consecutive temporal segments that show different behaviours close to the disruption
  - The concept of '***temporal segment***' is the key in this respect
- How are the segments defined?
  - How do the segments differ between them?
- After segment labelling, it is possible to find common sequences at the end of the discharges
  - Labels are assigned by means of an unsupervised clustering
- Are the several sequences of labels intrinsically related to the physics root cause of the disruption?



GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN
Ciemat
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Overview

- Conceptual view of the method
- Two different approaches to the implementation
- Results
- Conclusions

Steps

- Segment definition (discharge by discharge) in a multidimensional parameter space
- Recognition of temporal segments related to disruptions
- Unsupervised classification of patterns with the dataset of discharges and segment fusion
- Grouping of similar temporal sequences to identify different plasma behaviours leading to disruptions

Parameter space

Plasma Current
ModeLock
Plasma Internal Inductance
Plasma Density
Poloidal Beta
Total Input Power
Plasma Vertical Position
Radiated Power
Stored diamagnetic energy derivative
Bolometry signal for Core plasma lower section
Bolometry signal for Core plasma upper section
Bolometry signal for lower section
Bolometry signal for upper section
Vertical Soft X-Ray Core plasma measurement signal
Vertical Soft X-Ray High Field Side measurement signal (1/2)
Vertical Soft X-Ray High Field Side measurement signal (2/2)
Vertical Soft X-Ray Low Field Side measurement signal (1/2)
Vertical Soft X-Ray Low Field Side measurement signal (2/2)
High Field Side Line Integrated Density Signal
Core Plasma Line Integrated Density Signal (1/2)
Core Plasma Line Integrated Density Signal (2/2)
Low Field Side Line Integrated Density Signal

- Signals typically related to disruption prediction
- Are these 22 signals enough?
  - This is part of the analysis
- $\mathbf{x}(t) = \left(Ip(t), ML(t), \ldots \right) \in \mathbb{R}^{22}$
  - Combination of parameters in lower dimensional spaces have to be tested
- Off-line analysis
- The data flow of feature vectors is analysed from plasma start to extinction

Step 1: segment definition (discharge by discharge)

- How to define segments?

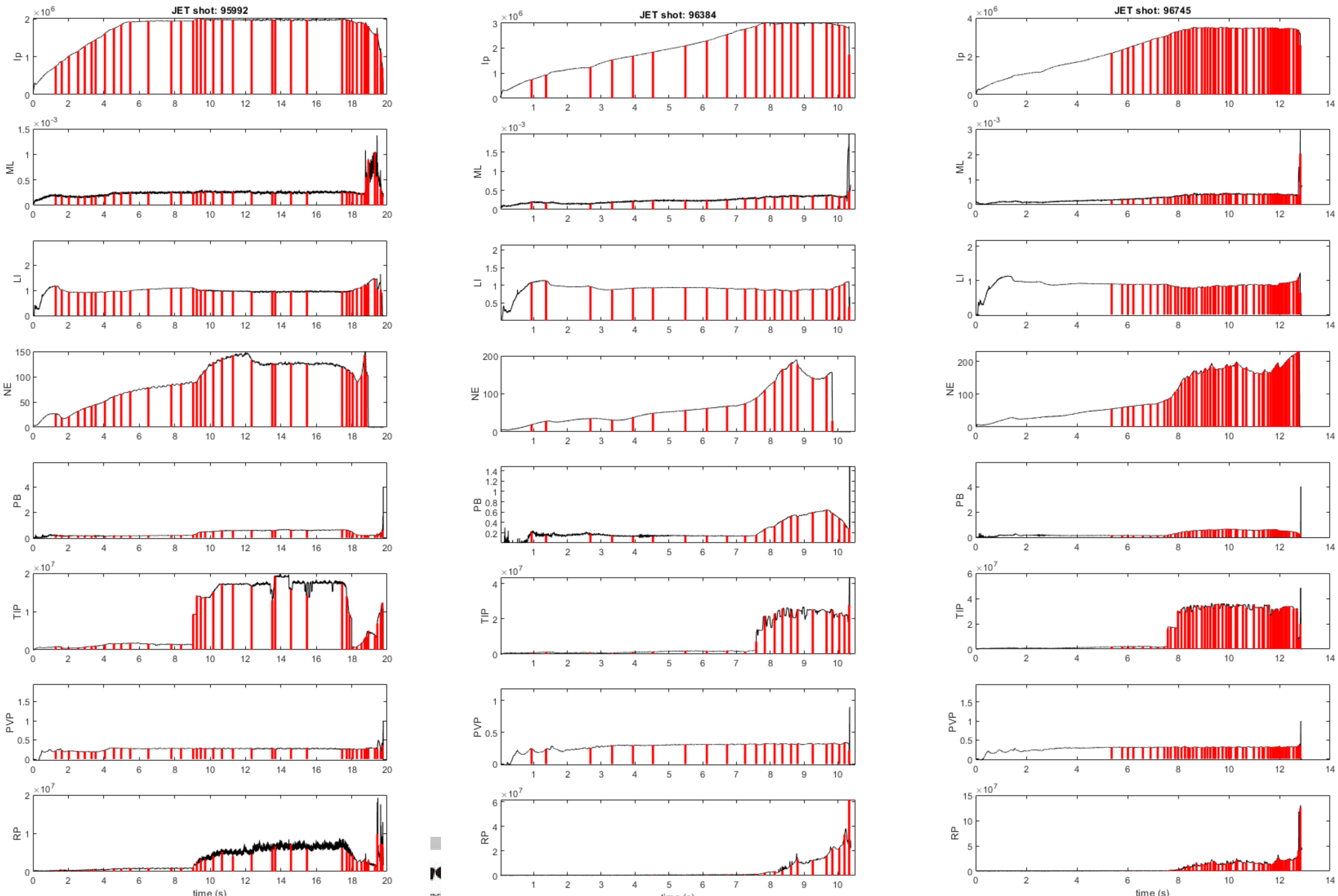
- The available information is a continuous data flow that provides the temporal evolution of feature vectors in multi-dimensional spaces
  - Machine learning methods can be used to extract knowledge from this temporal evolution

- Mathematical foundation
  - The main assumption in learning from data is that the examples are drawn independently from a fixed but unknown probability distribution function (PDF)
  - In any system that generates a continuous data flow (data streaming setting), the PDF may change as the data are streaming
    - The new PDF is also unknown
  - Such changes in the data may convey interesting time-dependent information and knowledge and, in general, the changes can be seen as ***anomalies*** in the system evolution

- These ***anomalies*** define the time limits of the segments

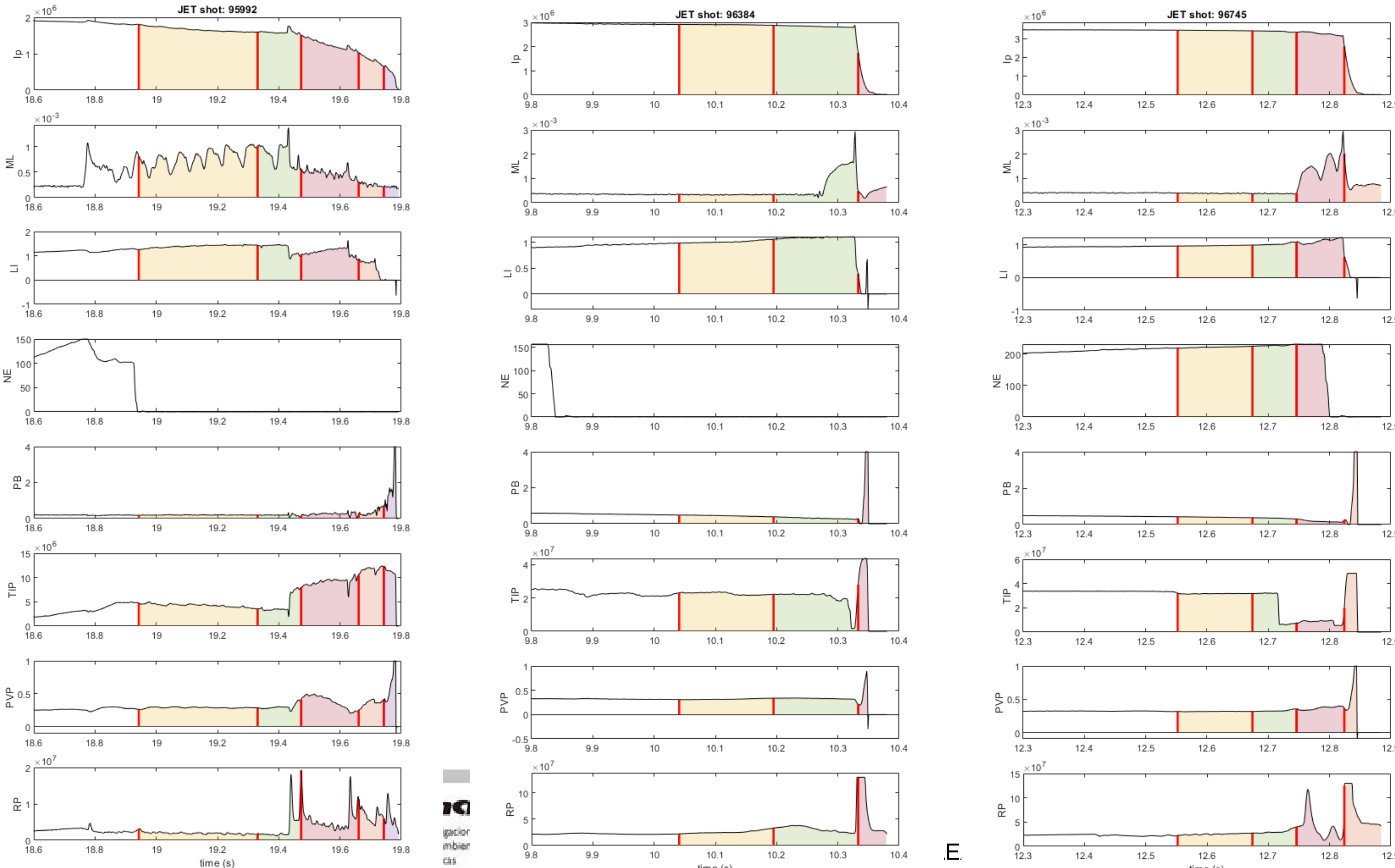
GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN
Ciemat
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Step 1: segment definition (discharge by discharge)



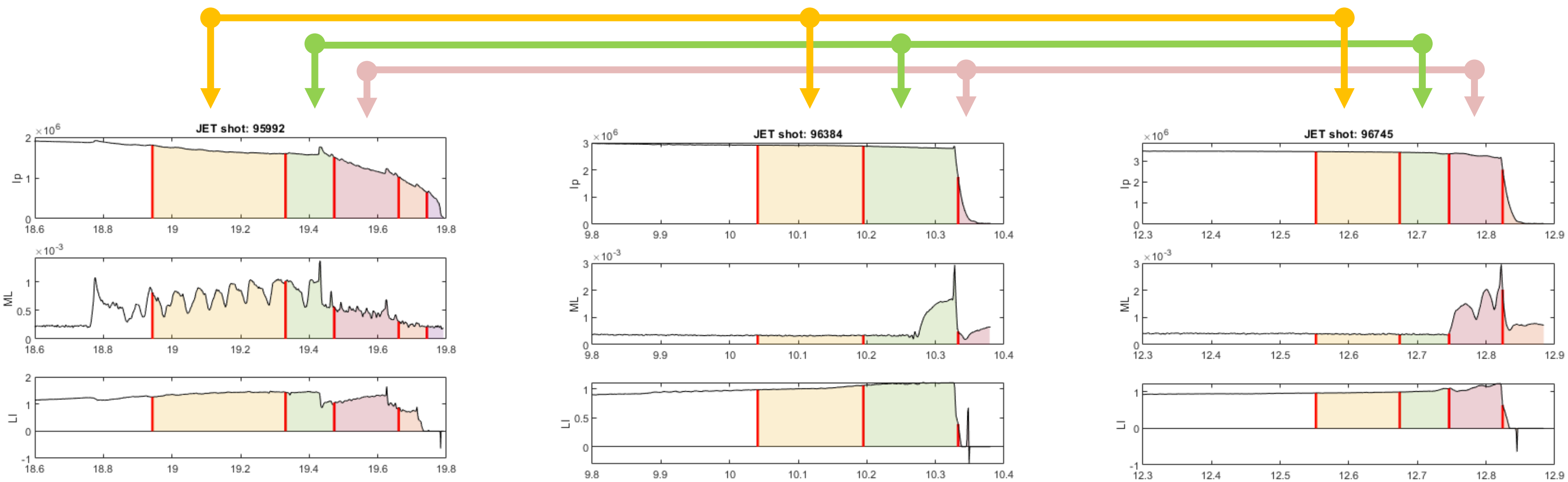
Step 2: Recognition of temporal segments related to disruptions

- A first filter is applied to only retain those segments that, according to the data evolution, are related to a disruptive behaviour
  - The other changes in the PDF (the other ***anomalies***) are not related to disruptions
- Each segment is represented by a single pattern to summarise the **plasma physics behaviour** along the segment
  - Different segments with different temporal length are present in different discharges



Step 3: Unsupervised classification of patterns with the dataset of discharges

- The plasma physics behaviour synthetized in the first (second, …) disruptive segment of one discharge is not necessarily related to the plasma physics behaviour summarised in the first (second, …) disruptive segment of other discharges
    - It is necessary to group the temporal segments that show a similar disruptive behaviour

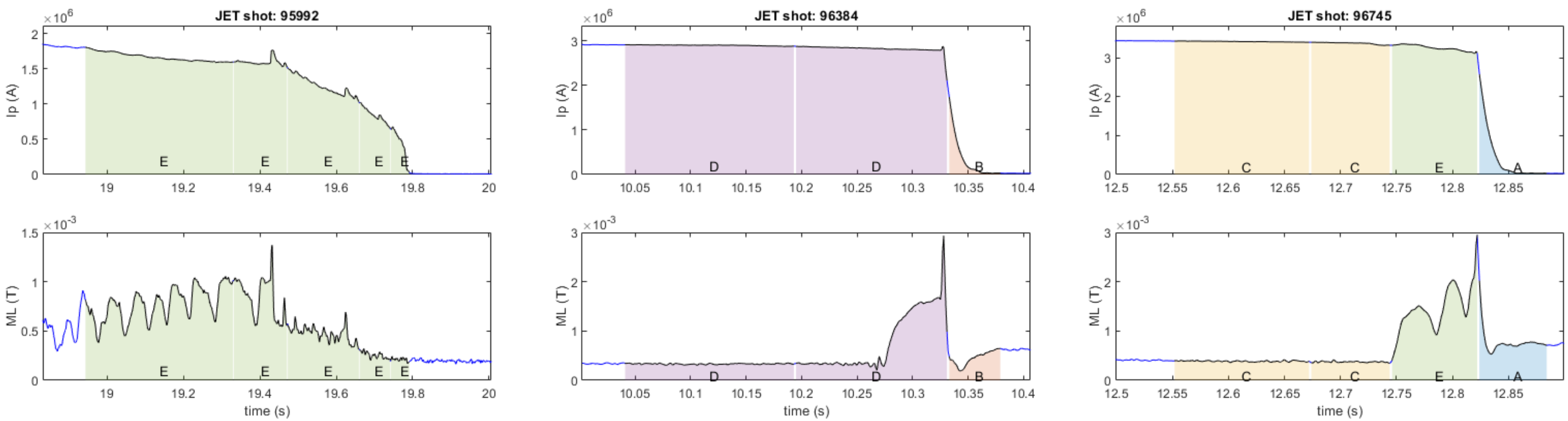


- An unsupervised classification allows associating a label to each segment to recognise similar physics behaviours
    - The unsupervised classification process is carried out with **all disruptive segments** of **all discharges** contained in the dataset of discharges that is analysed

Step 3: Unsupervised classification of patterns with the dataset of discharges

- Each physics behaviour (each segment) is represented by a label after the unsupervised classification process implemented with all segments of all discharges in the dataset

- Due to the fact that the classifier is unsupervised, the specific physics behaviour of each label is not identified by the classification process
  - The search for relationships between labels and physics is an additional step to be carried out by experts



- Segment fusion: sequence of labels
  - 95992: EEEEE ➔ E
  - 96384: DDB ➔ DB
  - 96745: CCEA ➔ CEA

Step 4: Grouping of similar temporal sequences to identify different behaviours leading to disruptions

- This step in intended for the automatic classification of the temporal sequences into a number of clusters

- The temporal sequences representing the end of disruptive discharges could have a large number of symbols (labels) and discharges with the same behaviour could have similar temporal sequences but not identical

Cluster 1
DJT

Cluster 2
JK
JK

Cluster 3
NH
NH
NG
NG
NG
NH
NG

Cluster 4
RP
RA
RB
RS
RK
RS
RO
RA
RB
RB
RB
DRUB
RB
RUB
RA
RA
RK
RQK
RB
RB
RB
RP
RS
RA
RO

Cluster 5
LEX
LEX
LFX
LEX
LEG
LEG
LEG

Cluster 6
CDIK
CDIK
VIK
IUB

Cluster 7
MW

Cluster 8
FX
FX
VX
FX
FNX

GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN
Ciemat
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Approaches

- Segment definition (discharge by discharge)
  - Anomaly detections by exchangeability martingales
- Recognition of temporal segments related to disruptions
  - Outlier detection (see later)

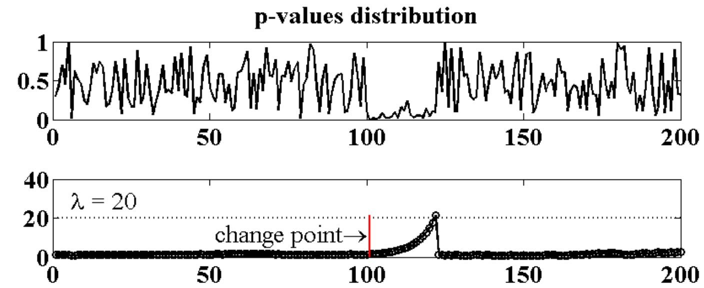
- Segment definition (discharge by discharge)
  - By dividing the discharges in time windows and applying a sliding window mechanism, changes in the data PDF are detected by means of unsupervised clustering
- Recognition of temporal segments related to disruptions
  - First segment after the longest one (see later)

- Unsupervised classification of patterns with the dataset of discharges and segment fusion
  - Determination of the optimal number of clusters by means of hierarchical clustering and the S_DbW validity index[1]
  - Segment fusion
  - Each discharge is represented by a string of labels
- Grouping of similar temporal sequences to identify different behaviours leading to disruptions
  - Unsupervised clustering with the above string of labels is performed
    - Variable number of symbols per string
    - The strings represent sequential data
  - To perform agglomerative cluster from linkages, the distance between strings is computed by means of the S^3M similarity function[2]
  - Determination of the optimal number of clusters by means of the Silhouette validity index[3]

[1] M. Halkidi et al. Proc. of IEEE international conference on data mining, ICDM 2001 (pp. 187–194). http://doi.org/10.1109/ICDM. 2001.989517
[2] P. Kumar et al. International Journal of Data Warehousing and Mining, 6 (4) (2010), 16–32
[3] P. J. Rousseeuw. Journal of Computational and Applied Mathematics, 20 (1987) 53–65 http://doi.org/10.1016/0377-0427(87)90125-7

Detection of anomalies: test of exchangeability[1]

- The examples $x(t) \in \mathbb{R}^n$ arrive one by one and a valid measure of the degree to which the assumption of exchangeability has been falsified is obtained
  - Such measures are provided by exchangeability martingales
  - A martingale is a sequence of random variables that remains stable in value with some fluctuation as long as the process is random, i.e. without external inference
  - The Randomised Power Martingale is used



$$M_n^{(\varepsilon)} = \prod_{i=1}^{n} \varepsilon p_i^{\varepsilon-1},$$

$\varepsilon \in [0,1]$ is a parameter to optimise,

$M_0^{(\varepsilon)} = 1$ and the $p_i s$ are computed from a p-value function

  - An anomaly is detected when $M_n^{(\varepsilon)} > \lambda$
  - *1/λ* determines the false alarm rate that one is willing to accept

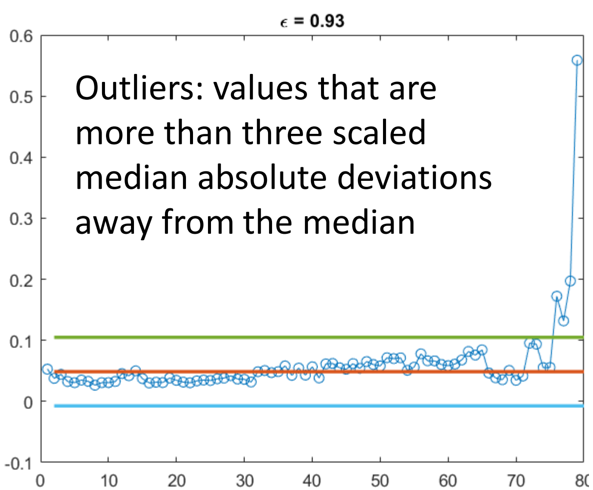
[1]S. Ho et al. IEEE Transactions on Pattern Analysis and Machine Intelligence. 32, 12 (2010) 2113-2127

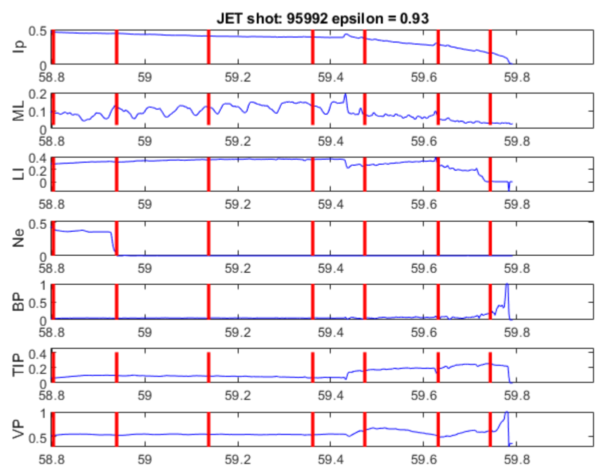
GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA E INNOVACIÓN

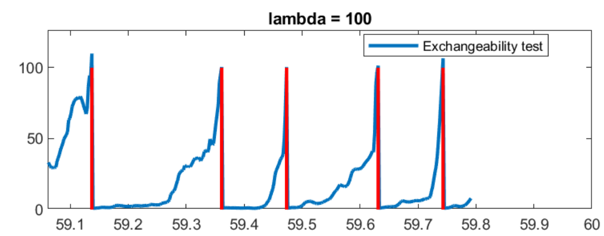
Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Detection of anomalies: test of exchangeability

- Parameters to optimise: ε and λ
- Optimization of ε and selection of disruptive segments
  - Each discharge is analysed with $0.9 \leq \varepsilon \leq 0.99$ and $\lambda = 100$
  - Each temporal segment of a shot (for each epsilon) is represented by the mean value of the distances between the feature vectors of the segment
  - The winner epsilon will be the one with the greatest number of outlier segments close to the disruption
    - If several epsilons have the same number of outlier segments, the winner is the one with the higher increasing rate of the martingale



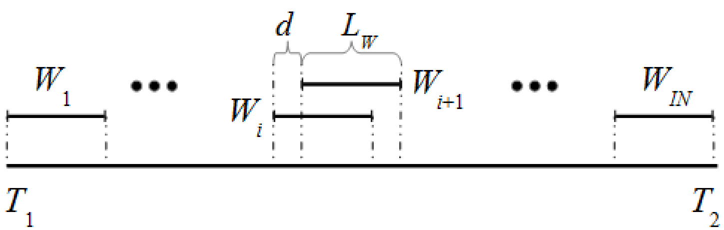




GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN
Ciemat
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Detection of anomalies: unsupervised detection method[1]

For each Disruptive discharge → **Dividing the discharge in temporal windows (sliding window mechanism)**



Feature vectors in each window

$$W(m) = \sum_{i=1}^{2} a_i W(m-i) + e(m)$$

(Two autoregressive coefficients)

Optimal unsupervised process by Hierarchical Clustering algorithm

Optimization of window size

Optimization of slide width

Optimization of number of clusters to classify the temporal windows

- Cophenetic Correlation Coefficient[2]

[2]R. Sokal et al.
Taxon, 11 (2) (1962) 33–40
http://doi.org/10.2307/1217208

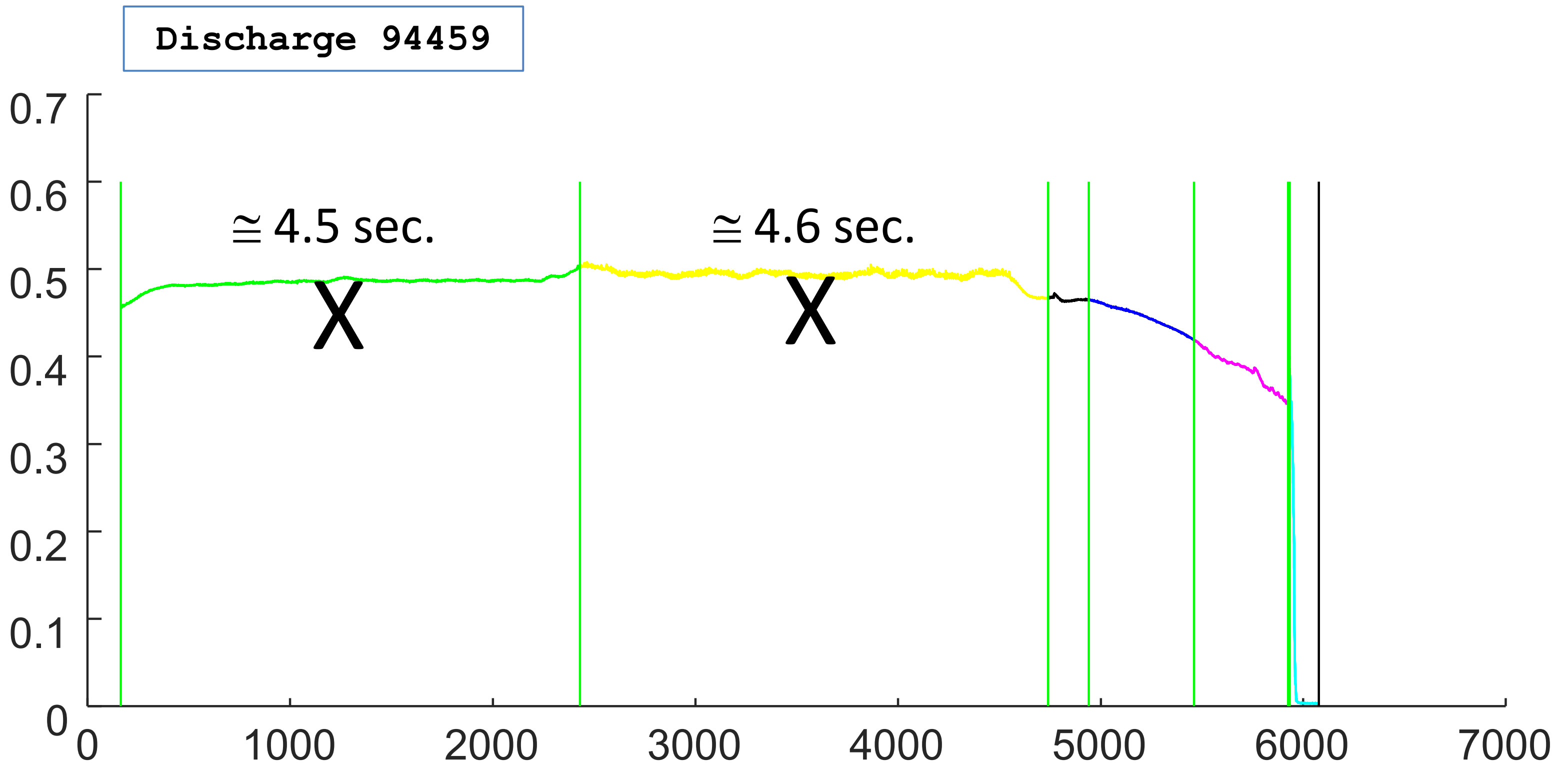
- S_Dbw Validity Index

Computational optimization time for each discharge around 3 hours

[1]A. Mur et al. Expert System with Applications, vol. 54, pp. 294-303, 2016

Event detection: consecutive segments of different types define anomalies

- Window size
- Slide width
- Number of clusters to classify the temporal windows



Results

- JET disruptive discharges in the range 94422 - 96996
  - Unintentional disruptions
  - 52 discharges corresponding to the baseline scenario

- With the 2 approaches that have been described, the ending part of disruptive discharges are grouped into 3 clusters

Anomaly detections with temporal windows and unsupervised clustering

Anomaly detections by exchangeability martingales

	Cluster 1	Cluster 2	Cluster 3	
Cluster 1	33	1	1	35
Cluster 2	2	7	0	9
Cluster 3	6	1	1	8
	41	9	2	

- Alarms fired by the JET control system
  - Cluster 1: high radiation emission
  - Cluster 2: MHD and locked mode
  - Cluster 3: mix of causes

- 79% of disruptions are classified in the same cluster by both approaches

GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA E INNOVACIÓN

Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Conclusions

- First results show that the detection of changes in the data PDF (multi-dimensional spaces) are useful to recognise several plasma behaviours
  - Both the initial PDF and the final PDF are unknown
- In particular, disruptive segments can be identified by recognising changes in the PDFs of data streams
- The disruptive part of discharges can be represented by strings of labels that represent the temporal evolution of the shots
- The string of labels can be grouped together to identify similar behaviours in the disruptive phase of the shots
- Tests with JET unintentional disruptions and two different methods show promising results
  - 79% of disruptions are grouped in the same clusters with the two methods
- A lot of additional work remains

Work in progress

- Parameter optimization
  - Selection of physics quantities and proper representations
  - Martingale parameters
  - Window width, sliding window mechanism
- Unsupervised clustering techniques
  - Hidden Markov Models, spectral clustering
- Application to other plasma scenarios
  - Hybrid scenario
- Physics interpretation of segments and label sequences
  - Relation with the root causes of disruptions

Thank you very much for your attention!

Questions?