Review of a data-driven adaptive disruption predictor for mitigation based on a nearest centroid approach

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

EUROfusion Consortium, JET, Culham Science Centre, Abingdon, OX14 3DB, UK

[1]Laboratorio Nacional de Fusión, CIEMAT. Av. Complutense 40. 28040 Madrid. Spain
[2]Consorzio RFX (CNR, ENEA, INFN, Università di Padova, Acciaierie Venete SpA), Padova, Italy
[3]United Kingdom Atomic Energy Authority, Culham Science Centre, Abingdon, Oxfordshire, UK
[4]Depto. De Informática y Automática. UNED. 28040 Madrid, Spain

* 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)

Abstract

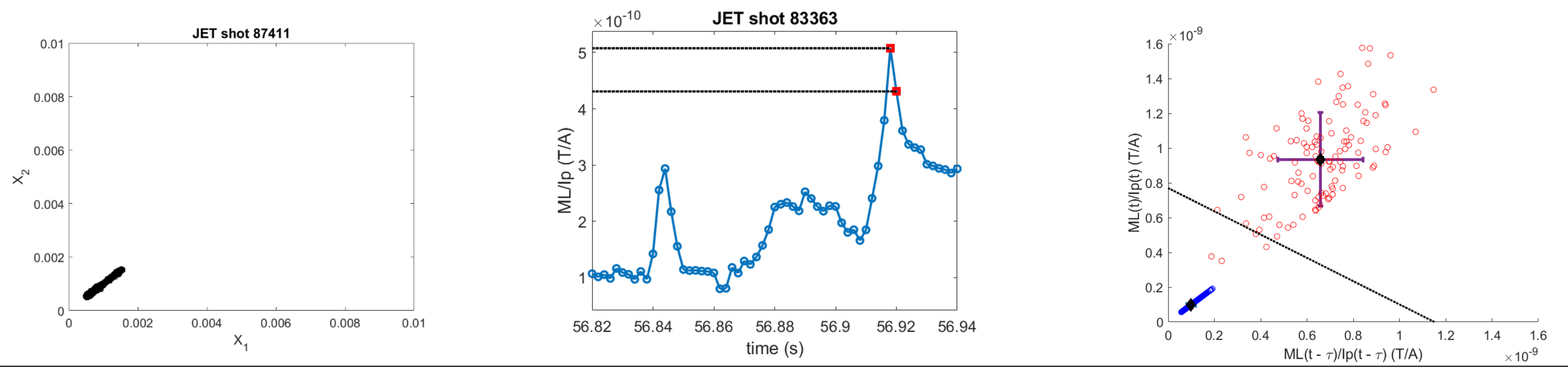
Any disruption mitigation system requires a trigger to trigger the corresponding remedial actions. Such trigger is the final step of a chain of events. This chain starts with an alarm that recognises an incoming disruption followed by interlocks protecting particular systems (for example, plasma heating systems). This contribution is a review of a specific disruption predictor that is installed in JET. The predictor uses only one signal, the mode lock normalised to the plasma current (NML), and its feature space, in which the separation frontier between disruptive and non-disruptive behaviour is linear, is two-dimensional. The linear frontier is defined based on two centroids, where each one summarises the disruptive and non-disruptive behaviours of past discharges, respectively. From a conceptual point of view, the predictor recognises a disruptive behaviour when large differences between consecutive samples of the NML appear. The predictor is installed in the JET real-time network from June 2019 (in open loop). The real-time predictions analysed so far confirm the following positive characteristics: fully deterministic (the running time of the algorithm for each prediction is less than 10 μs), not based on a simple threshold but on differences of amplitudes, easy physics interpretation (not a black-box), success rates above 96%, false alarm rates about 4%, most of the alarms very close to the disruption (26% of alarms within 10 ms) and average warning times of about 100 ms (can be smaller if assertion times are set-up). Off-line analyses with several databases (JET with C-wall, JET with ILW and JT-60U) have shown full compatibility with an adaptive development from scratch with about 10 re-trainings when tested in thousands of discharges. Re-trainings are performed after missed alarms. These properties make the predictor a potential candidate to be used as disruption predictor in ITER for mitigation purposes

Disruption prediction and the Mode Lock (ML) or normalised ML signals

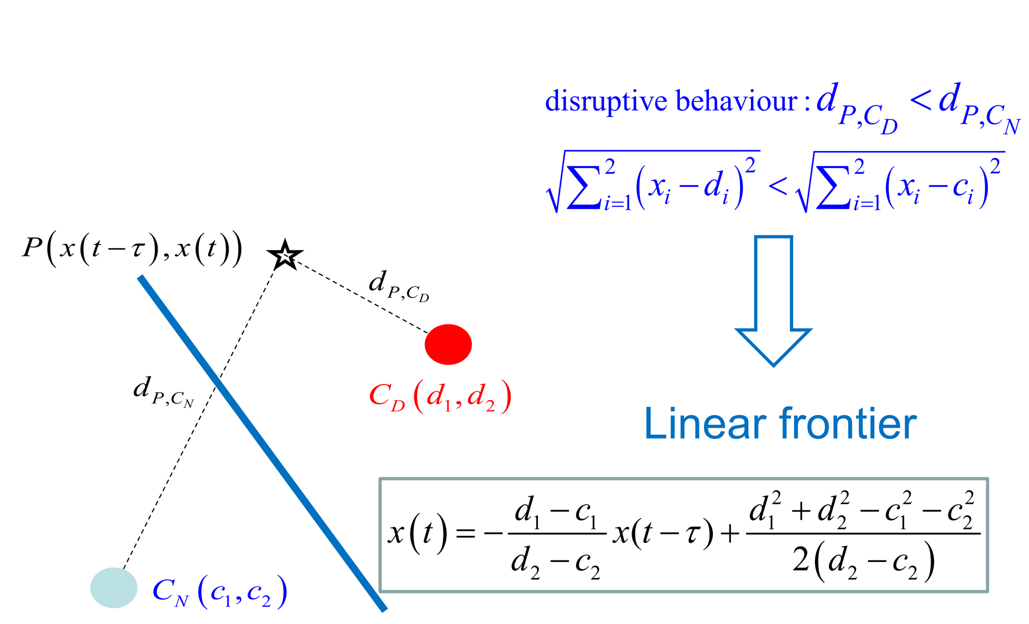
- X(t) = ML(t) or X(t) = ML(t)/Ip(t) thresholds are usually used to trigger alarms
  - Signal increases when
    - The rotation of an MHD mode slows down and can be locked
    - The MHD mode amplitude grows
  - Signal decreases when
    - The MHD mode amplitude drops
    - The MHD mode unlocks and the rotation speeds up
- An increasing signal is associated to a **disruptive behaviour**
- A decreasing signal is associated to a **non-disruptive behaviour**

Two-dimensional parameter space

- The **parameter space of consecutive amplitudes** (X(t - τ), X(t)) is chosen
- This means to put the focus on the differences of amplitudes (deltas) to recognize incoming disruptions
- Nearest centroid approach: disruptive/non-disruptive behaviours are summarised in two single points
  - Each discharge (disruptive or non-disruptive) contributes with a single centroid (disruptive or non-disruptive)



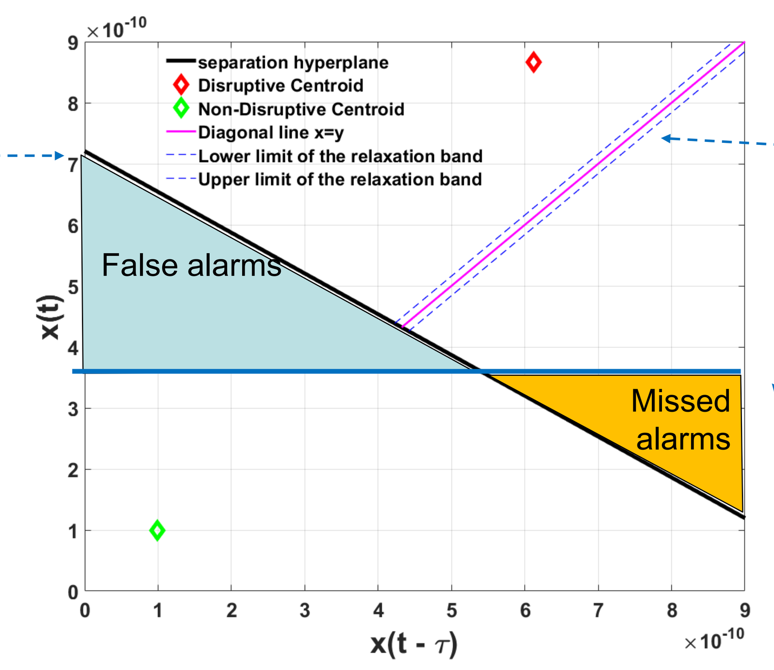
Rationale



J. Vega et al. Nuclear Fusion 60 (2020) 026001 (13pp)

Consequences

- The interception of the separation frontier defines a critical value above which the plasma is in a disruptive state regardless the amplitude of the previous sample
- Below the critical value, the disruptive behaviour depends on the previous amplitude



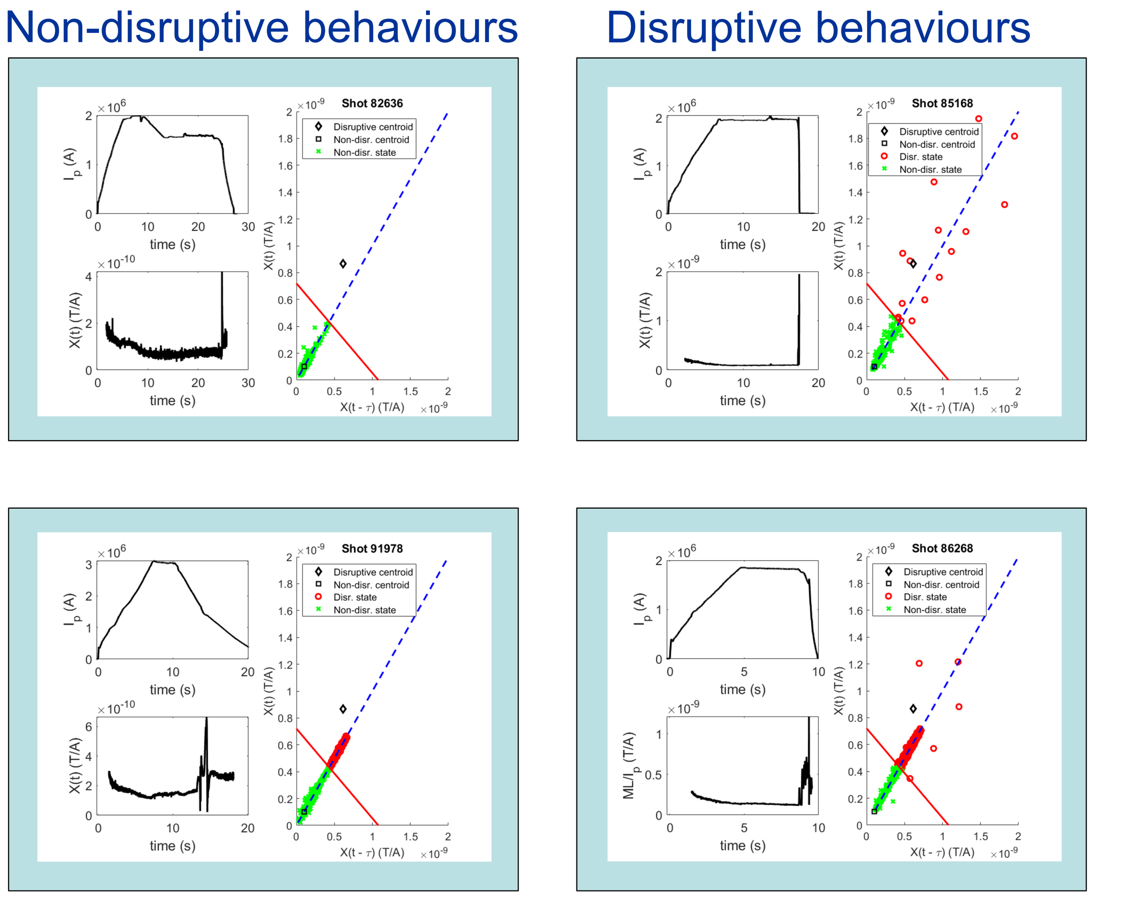
$$X(t) > A \cdot X(t-0.002) + B$$

Interpretation

- The extra 2nd dimension allows for the following physics interpretations
  - Smooth variations of the signal (deltas ≈ 0) mean non-disruptive behaviour
    - There is a band around the diagonal in the disruptive zone that means non-disruptive behaviour
  - An increasing X(t) signal does not necessarily mean a disruptive behaviour
    - Small deltas do not represent disruptive behaviours
  - A decreasing X(t) signal does not necessarily mean a non-disruptive behaviour
    - Large deltas represent disruptive behaviours
  - A simple threshold in the signal is not optimal to recognise disruptive behaviours

Predictions in the two-dimensional parameter space

Non-disruptive behaviours | Disruptive behaviours



Training/test

Type/use	Number of shots	Range
D/training	113	80181-82504 SEP 2011-MAR 2012
ND/training	1397	80176-82550 SEP 2011-MAR 2012
D/test	281	82569-92410 MAR 2012-NOV 2016
ND/test	3027	82552-92504 MAR 2012-NOV 2016

$$X(t) = -0.7441 \cdot X(t-0.002) + 6.1243e-10$$

dB,k (mT/MA)	SR (%)	WT>0 (%)	FA (%)
0	99.29	97.51	1.35
0.0021	99.29	97.51	1.32
0.0041	99.29	97.51	1.32
0.0062	99.29	97.51	1.32
0.0083	99.29	97.51	1.29
0.0103	99.29	97.51	1.29
0.0124	99.29	97.51	1.16
0.0145	99.29	97.51	1.16
0.0165	**99.29**	**97.51**	**1.12**
0.0186	98.58	97.15	1.12
0.0207	98.58	97.15	1.02

Success rate with positive warning time: 98%. False alarms: 4%. Average warning time: O(400 ms)

Results in the JET RT network in open loop

- Success rate: 96.2%
- Success rate with positive warning time: 84.6%
- Average warning time: 117 ms
- False alarms: 4%

D. Gadariya et al. 4th IAEA TM on Fusion Data Processing, Validation and Analysis

- Fully deterministic predictor: algorithm execution time for each prediction is less than 10 μs

C. I. Stuart et al. Fus. Eng. Des. 168 (2021) 112412

Adaptive prediction from scratch

J. Vega et al. Nature Physics (2022)

- Re-trainings are performed after missed alarms
- Off-line analyses with several databases have shown full compatibility with an adaptive development from scratch
  - JET case: 10 re-trainings when tested in more than 1200 discharges
    - CW: 2738 discharges (175 disruptive, 2563 non-disruptive). Success rate +98%, false alarms -3%, average warning time O(200 ms)
    - ILW: 4806 discharges (388 disruptive, 4418 non-disruptive). Success rate +98%, false alarms -3%, average warning time O(300 ms)
  - JT-60U: 2 re-trainings when tested with 154 discharges J. Vega et al. Fus. Eng. Des. 146 (2019) 1291
    - 154 discharges (62 disruptive, 92 non-disruptive). Success rate +97%, false alarms 20% (not enough training discharges), average warning times 60 ms

Summary

- ITER will require several disruptive event detectors whose combination will be used for avoidance, prevention and mitigation
- Both off-line simulations and real-time results in JET (open loop) show the centroid method as one of the potential candidates among others for mitigation purposes in ITER
- The centroid method shows an extremely simple real-time implementation to recognize disruptive behaviours: X(t) > A*X(t – τ) + B
- Off-line analyses with several databases have shown full compatibility with an adaptive development from scratch in a data-driven approach

This work was partially funded by the Spanish Ministry of Science and Innovation under the Project MCIN/AEI/ 10.13039/501100011033

This work has been carried out within the framework of the EUROfusion Consortium, funded by the European Union via the Euratom Research and Training Programme (Grant Agreement No 101052200 — EUROfusion). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.