
Advanced techniques of disruption prediction, application to JET and extrapolation to ITER



PhD Thesis

Raúl Moreno Salinas

•

Supervised by

Dr. Jesús A. Vega Sánchez

Dr. Sebastián Dormido Canto

•

Grupo de Adquisición de Datos del TJ-II

Laboratorio Nacional de Fusión

Centro de Investigaciones Energéticas, Medioambientales y
Tecnológicas (CIEMAT)

•

Programa de Doctorado en Ingeniería de Sistemas y Control

Departamento de Informática y Automática

Escuela Técnica Superior de Ingeniería Informática (ETSII)

Universidad Nacional de Educación a Distancia (UNED)

•

2015

Advanced techniques of disruption prediction, application to JET and extrapolation to ITER

Dissertation submitted to obtain the Ph. D. Degree by

Raúl Moreno Salinas

Supervised by

**Dr. Jesús A. Vega Sánchez
Dr. Sebastián Dormido Canto**

Grupo de Adquisición de Datos del TJ-II

Laboratorio Nacional de Fusión

**Centro de Investigaciones Energéticas, Medioambientales y
Tecnológicas (CIEMAT)**

•

Programa de Doctorado en Ingeniería de Sistemas y Control

Departamento de Informática y Automática

Escuela Técnica Superior de Ingeniería Informática (ETSII)

Universidad Nacional de Educación a Distancia (UNED)

•

2015

*Take the time, look around you
And the world will be better*

Acknowledgements

Es difícil agradecer sin extenderme a todos lo que han colaborado y me han acompañado durante estos 4 años. En primer lugar tengo que dar las gracias a mis directores, el Dr. Jesús A. Vega y el Dr. Sebastián Dormido Canto cuya experiencia, consejos y guía han hecho posible esta tesis. Agradecer al Dr. Andrea Murari su apoyo y colaboración en todo momento.

Quiero destacar en especial el apoyo, ayuda y consejos que me han dado en todo momento mis compañeros de despacho Augusto Pereira y Ana B. Portas; así como a Rodrigo Castro. A Sergio González, otro compañero y amigo que dedicó tiempo cuando entré en el grupo y me ha ayudado cada vez que lo he necesitado.

No puedo nombrar a todos los que habéis aportado algo estos 4 años, por poco que sea, pero en Ciemat he conocido grandes personas y amigos. Sin duda gracias a ellos he disfrutado de un lugar de trabajo rodeado de amigos: Edu, Andrés, Isaac, Xan, Marcos, Kike, Iván, Elena, Yupi, Jesús, Regi, Rapi, David Jiménez, Avelino, Iole, Cristina, Lalia, Dani Alegre, Dani Iglesias, Juan,... A todos os podría dedicar unas líneas aquí, pero se haría interminable y seguiréis viéndome aunque sea por los bares de Madrid.

Los amigos de toda la vida, que me apoyan en cada paso profesional y personal: Juanjo, Alcaide, Víctor, Boche, Tere y todo el gran grupo de Córdoba. Sin vosotros cada vuelta a casa no habría sido lo mismo. El grupo de Madrid que desde que llegué a esta ciudad han llenado mi vida de buenos momentos: Rubén, Gonzalo, Fely, Eduardo, Gelines,...

El grupo de buenos músicos y mejores personas con los que llevo compartiendo años una de mis grandes pasiones: Juanjo, Gon, Dani y Elena.

Finalmente a mis padres y mis hermanos, que como en cada momento de mi vida siempre están ahí animando a seguir con lo que se presente. Y a la persona que me aguanta cada día, la que ha escuchado y leído este trabajo tanto que lo sabe mejor que yo, y la que hace que no importe lo difícil que pueda ser el camino, gracias Carlota.

Introduction

The human being has been consuming the resources of the planet since thousand years ago. The more advance is technology and human needs, the higher consumption of planet resources. According to the IEA (International Energy Agency) on 6th June, 2013 (IEA, 2013)¹, fossil fuels currently meet 80% of global energy demand. Even if every countries change their policy commitments to fight against climate change and achieve other challenges, global energy demand in 2035 is expected to be 40% (with fossil fuels still contributing 75%). Taking into account the global reserves and current consumption levels:

- Oil is expected to last in the range of 40 to 45 years but it will grow as resources are successfully converted into reserves. Despite growth in consumption, global reserves have increased since 1990.
- Conventional gas is thought to remain in the range of 55 to 60 years. Nevertheless, unconventional gas reserves are very difficult to be estimated, because of rock formation heterogeneity.
- Reserves of coal are high and the ratio is in the range 110 to 120 years at current consumption levels.

With adequate resources of fossil fuels to meet the increasing energy demand, Resources to Reserves 2013 shows that the emphasis now is on the technology, prices and policies that can ensure resources are developed into proven reserves in financially viable and environmentally sustainable ways.

Fossil fuels had a strong impact on environment, so nuclear energy is a great alternative. Remarkable amounts of energy can be obtained by this energy. Currently, nuclear fission, which was discovered in 1939 by the German researchers Hahn and Strassmann, can be safely handle. Despite of this, extraction, enrichment, chemical reprocessing and use of high molecular weight materials (like uranium), produce radioactive waste which harm the environment for thousand years. In spite of the high security, the danger of possible

¹This information has been obtained on the web page: <http://www.iea.org/newsroomandevents/news/2013/June/name,38548,en.html>

accidents in atomic power plants generate the population rejection against nuclear fission.

At this point, renewable energies are bound to be a great alternative. Towards the middle of the 21th century, the use of renewable energies is expected to increase significantly, albeit it will barely cover the growth of the energy demand. By regions, the EU will be in the head of the renewable primary energy demand, followed by the USA and China. Studies point out that there will be a reduction in the price of the electricity generated by renewable energy (de Vries et al., 2007).

At the crossroads of renewable and fossil energies, on the one hand, fossil fuels ensure supply for a few decades endangering the future of the environment. On the other hand, renewable energies represent an unlimited and harmless resource of energy, but unable to cover the worldwide energy demand. Hence, the importance of achieve nuclear fusion.

The energy from the stars is called nuclear fusion. Sir Arthur Eddington, in 1926, proposed that there is a limit where the pressure of the stars light on the atmosphere is large enough to balance the gravitational weight of the stellar atmosphere entirely, known as the Eddington luminosity limit. If the star were any brighter, the light of the star would push away the outer layers of the atmosphere, thus causing the star to lose mass. Eddington described stars as gas spheres at high temperature, mainly composed by hydrogen (Eddington, 1926). The pressure of gravitational forces can reach a temperature which turns gases into plasma (the fourth state of the matter). In 1927, Lewi Tonks and Irving Langmuir designated as *plasma*, a nearly neutral ionised gas in which ions and electrons were moving independently (Tonks and Langmuir, 1929). The temperature and pressure in the stars fuse hydrogen nuclei in plasma producing helium and energy. Nuclear fusion as renewable energy attempt to generate electricity reproducing the reaction which happens in the stars. The fuel used in nuclear fusion reactors will be tritium (a hydrogen radioactive isotope, which can be obtained from lithium). It should be considered the low CO_2 emissions (6 – 12 Cg/kWh) of nuclear fusion, followed by hydroelectric energy (4.8 Cg/kWh) and light water reactors (5.7 Cg/kWh). Photo-voltaic (16 Cg/kWh) and Wind (33.7 Cg/kWh) energies also have low CO_2 emissions but the inconvenient is the manufacture processes and the unstable production of energy, it means, they are not able to produce reliable continuous power. Albeit, oil(200 Cg/kWh) and coal (270 Cg/kWh) are the most contaminant energies (Kikuchi, 2009).

There is a great advantage in fusion against fission. In contrast to fission, which employs a large quantity of uranium or plutonium fuel, nuclear fusion employs a very small amount of tritium and deuterium fuel. If fuel is not continually replaced, the fusion reaction goes out and stays out, so there is no danger of an accident, nuclear fusion is a reaction happened in stars, not in the Earth. A second advantage is the radioactive waste. Radioactive

substances from the fission fuel cycle, that need to be separated from the unused fuel and stored safely for thousands of years, are the most difficult to process and store. Nuclear fusion generates none radioactive waste, it produces helium gas and tritium. Tritium is radioactive but it decays quickly (12.32 years half-life) against uranium (238.45 million years half-life), even, the purpose is to recycle the tritium fuel used in the power-plant. In addition, mining exploitation is not necessary in nuclear fusion because deuterium and tritium are easy to obtain. The materials can be obtained with lithium and water all around the planet for any country. There is enough lithium to last for ten thousands of years and enough deuterium in the oceans, (McCracken and Stott, 2005).

Goals

The challenge in this thesis is to deal with the massive and noisy databases in nuclear fusion to extract useful information and develop artificial intelligence algorithms and techniques to sort out specific problems in nuclear fusion. This is a work focused in disruption prediction and its databases. The most dangerous instabilities in tokamak devices need to be predicted to carry out mitigation actions but the data provided to develop predictors is not trivial, it must be studied, analysed and conveniently processed. The roadmap to fusion is long, but every big goal is formed by small steps.

Contents

This thesis is structured in five chapters. Chapter 1 contains the most important concepts about nuclear fusion and summarizes the plasma physics theory to understand the framework of this thesis. This chapter introduces the main problem in tokamak devices, the disruptions. Then in chapter 2 the data mining techniques which have been used in this thesis are explained. Chapter 3 shows the generation of nuclear fusion databases. Then in this chapter APODIS is widely study: its assessment and performance during the ILW campaigns at JET, its robustness and a sliding window mechanism. The determination of the disruption time automatically is treated in chapter 4. Finally chapter 5 covers the studies on disruption prediction in view of the next generation of tokamaks.

Contents

| | |
|---|------------|
| Acknowledgements | vii |
| Introduction | ix |
| 1 Fusion, the energy of the future | 1 |
| 1.1 Nuclear fusion | 1 |
| 1.1.1 A brief approach to plasma physics | 3 |
| 1.1.2 Fusion Devices | 9 |
| 1.1.3 Magnetic confinement devices, Tokamaks | 13 |
| 1.1.3.1 Joint European Torus - JET | 15 |
| 1.1.3.2 ITER and DEMO | 19 |
| 1.2 Plasma phenomena | 21 |
| 1.2.1 Disruptions, the most danger instability | 24 |
| 1.2.2 Disruption predictors | 28 |
| 2 Data mining techniques | 31 |
| 2.1 Data signal processing | 31 |
| 2.1.1 Data normalisation | 32 |
| 2.1.2 Data interpolation | 33 |
| 2.1.3 Fourier transform | 36 |
| 2.2 Wavelet analysis | 39 |
| 2.3 Classification | 43 |
| 2.3.1 Support Vector Machines | 44 |
| 2.3.1.1 The optimal separating hyperplane | 45 |
| 2.3.1.2 Mapping the input space into a higher dimensionally feature space | 49 |
| 2.3.1.3 Constructing SVM | 50 |
| 2.3.1.4 Examples of SVM classification | 51 |
| 2.3.2 Conformal Predictors | 53 |
| 2.3.2.1 Inductive Conformal Prediction | 55 |
| 2.3.2.2 Examples of Conformal Prediction | 56 |

| | | |
|----------|---|------------|
| 2.4 | Regression | 61 |
| 2.4.1 | Support Vector Regression | 61 |
| 2.4.2 | UMEL | 63 |
| 2.5 | Martingales | 65 |
| 3 | Databases and APODIS during ILW | 71 |
| 3.1 | Advanced Predictor of Disruptions, APODIS | 71 |
| 3.2 | Databases and signal processing | 75 |
| 3.2.1 | Analysis of JET campaigns | 78 |
| 3.2.2 | Signal processing | 84 |
| 3.2.3 | Summary of ILW databases | 86 |
| 3.3 | Assessment of the APODIS performance in ILW campaigns | 88 |
| 3.4 | Robustness of APODIS | 92 |
| 3.5 | Increased time resolution of APODIS | 98 |
| 4 | Automatic location of disruption times in JET | 101 |
| 4.1 | Previous studies | 103 |
| 4.1.1 | Classifier | 103 |
| 4.1.2 | Martingales | 105 |
| 4.1.3 | UMEL | 111 |
| 4.2 | Wavelets | 115 |
| 5 | Disruption predictions for the next generation of tokamaks | 119 |
| 5.1 | ITER-like Wall (ILW) predictors | 119 |
| 5.1.1 | Analysis of plasma dynamics to detect the approach to the disruption boundaries | 119 |
| 5.1.2 | Comparison between predictors trained and test with ILW data | 124 |
| 5.1.3 | Deuterium-Hydrogen ILW campaign | 128 |
| 5.2 | Disruption predictors from scratch | 130 |
| 5.2.1 | Probabilistic Venn predictor | 133 |
| 5.2.2 | APODIS from scratch | 138 |
| 5.2.3 | Disruption prediction from scratch during ILW campaigns C31-C34 | 144 |
| 6 | Conclusions and Future Work | 151 |
| I | Appendices | 155 |
| A | List of Journal Papers & Conferences | 157 |
| B | Parkinson's disease database | 161 |

| | |
|------------------------------------|------------|
| C Breast cancer database | 163 |
| D Martingale Test Algorithm | 165 |
| References | 167 |

List of Figures

| | | |
|------|--|----|
| 1.1 | Cross-section for the most common reactions in plasma fusion | 3 |
| 1.2 | Range of plasmas | 4 |
| 1.3 | Trajectories of a charged particle in a magnetic field | 8 |
| 1.4 | Progress of the fusion “triple product” of plasma ion density, ion temperature and energy confinement time | 10 |
| 1.5 | Inertial confinement procedure | 11 |
| 1.6 | Stellarator configuration (TJ-II, Ciemat, Spain) | 12 |
| 1.7 | Current and fields in a tokamak plasma | 13 |
| 1.8 | Original JET design | 16 |
| 1.9 | JET coil systems and iron transformer (EUROFUSION [©]) | 17 |
| 1.10 | Plasma X-point | 18 |
| 1.11 | Layout of JET diagnostic systems | 19 |
| 1.12 | ITER design | 20 |
| 1.13 | Reconnection of magnetic fields producing a magnetic island | 22 |
| 1.14 | Example of ELMs types | 23 |
| 1.15 | Example of a disruption at JET | 25 |
| 1.16 | Hugill diagram showing operating regime limited by disruptions | 26 |
| | | |
| 2.1 | Linear interpolation | 35 |
| 2.2 | Example of interpolation methods | 35 |
| 2.3 | Gaussian function for short-time Fourier transform | 38 |
| 2.4 | Haar wavelet levels | 41 |
| 2.5 | Example of Haar transform | 42 |
| 2.6 | Inductive and transductive prediction diagram | 44 |
| 2.7 | Example of optimal and non-optimal separating hyperplane | 47 |
| 2.8 | Input space is mapped into a feature space using a kernel function | 50 |
| 2.9 | Breast cancer conformal classification with linear kernel | 58 |
| 2.10 | Breast cancer conformal classification with polynomial kernel | 59 |
| 2.11 | Breast cancer conformal classification with RBF kernel | 60 |
| 2.12 | ϵ -insensitive loss function | 62 |

| | | |
|------|---|-----|
| 2.13 | UMEL example with different types of kernel | 63 |
| 2.14 | UMEL example | 65 |
| 3.1 | Temporal evolution of a JET discharge | 73 |
| 3.2 | APODIS structure formed by two layers | 74 |
| 3.3 | APODIS analysis procedure during a discharge | 74 |
| 3.4 | JET raw data estimated in bytes (B) | 76 |
| 3.5 | Rate of non-intentional disruptions and non-disruptive discharges in JET | 77 |
| 3.6 | Example of a right discharge to be considered for the databases | 79 |
| 3.7 | Example of a discharge not considered in the database with errors in several signals | 80 |
| 3.8 | Examples of EFCC effect on mode lock amplitude | 81 |
| 3.9 | Examples of non-intentional and intentional disruptions at JET | 83 |
| 3.10 | Signal processing overview to generate databases | 86 |
| 3.11 | The logarithmic graphic shows the APODIS results during C28 – C30 campaigns | 88 |
| 3.12 | APODIS performance compare with mode lock trigger from JPS, during ILW campaigns C28 – C34 | 91 |
| 3.13 | Example of both simulations carried out in the robustness analysis | 94 |
| 3.14 | Results for CFC and ILW campaigns with mode lock and plasma inductance signals in failure | 95 |
| 3.15 | Results for CFC and ILW campaigns with diamagnetic energy time derivative signal failure | 96 |
| 3.16 | Example of missed alarm due to a lack of resolution | 98 |
| 3.17 | Example of temporal resolution of 32 ms (how APODIS works currently at JET) and temporal resolution of 4 ms | 99 |
| 3.18 | Sliding window results | 100 |
| 4.1 | Example of a disruption during campaigns with the JET-C wall | 101 |
| 4.2 | Disruption from JET-ILW experiments | 102 |
| 4.3 | Signals used on disruption time location as a classification approach | 104 |
| 4.4 | Diagram of martingale for change detection in data streams . | 105 |
| 4.5 | Density and Peaking Factor in discharge 81867 | 106 |
| 4.6 | Scatter plot of the Density and Peaking Factor in discharge 81867 | 107 |
| 4.7 | Martingale change detection to locate disruption time | 108 |
| 4.8 | Histogram of RPM-S1 results with wavelet Haar level 2 | 109 |
| 4.9 | Histogram of RPM-S3 results with wavelet Haar level 2 | 110 |
| 4.10 | Signals used by UMEL on disruption time location | 112 |

| | | |
|------|---|-----|
| 4.11 | First step: UMEL is applied to the signals and the first shared ESV is detected | 113 |
| 4.12 | Second step: UMEL is applied to the detail coefficients of the plasma current data selection | 113 |
| 4.13 | Example of different parameters for UMEL to locate disrup- tion time | 114 |
| 4.14 | Example of the data processing algorithm for disruption time location | 116 |
| | | |
| 5.1 | Operational space during a discharge | 120 |
| 5.2 | Examples of temporal evolution of the distance to the APODIS separating hyper-plane | 121 |
| 5.3 | Example and distribuyion of transit speed | 121 |
| 5.4 | Credibility diagram | 122 |
| 5.5 | Plasma dynamics examples with conformal prediction | 122 |
| 5.6 | Plasma dynamics at non-return point | 123 |
| 5.7 | No-return times histogram | 123 |
| 5.8 | Comparison between ILW predictors | 127 |
| 5.9 | Scatter plot: <i>C34</i> features | 129 |
| 5.10 | Scatter plot: <i>C28</i> – <i>C30</i> features | 130 |
| 5.11 | Data samples of Venn predictor example | 135 |
| 5.12 | Features with the best results using Venn predictors | 138 |
| 5.13 | Simplified example of prediction from scratch | 139 |
| 5.14 | Unbalanced approach results | 142 |
| 5.15 | Balanced approach results | 142 |
| 5.16 | Hybrid approach results | 143 |
| 5.17 | Hybrid approach results retraining every missed disruption | 144 |
| 5.18 | Unbalanced approach results with one layer predictor | 145 |
| 5.19 | Balanced approach results with one layer predictor | 146 |
| 5.20 | Unbalanced approach results from scratch for all ILW campaigns | 147 |
| 5.21 | Unbalanced approach results from scratch for ILW campaigns <i>C28</i> – <i>C33</i> | 148 |
| 5.22 | Unbalanced approach results from scratch for ILW campaign <i>C34</i> | 148 |

List of Tables

| | | |
|------|---|----|
| 1.1 | Main parameters for some typical weakly coupled plasmas . . . | 7 |
| 1.2 | Stellarators devices | 13 |
| 1.3 | Tokamaks devices and features | 15 |
| 1.4 | ITER features | 20 |
| 2.1 | SVM classification of Parkinson’s disease | 52 |
| 2.2 | First SVM classification of breast cancer | 53 |
| 2.3 | Second SVM classification of breast cancer | 53 |
| 2.4 | First CP classification of breast cancer | 57 |
| 2.5 | Second CP classification of breast cancer | 58 |
| 3.1 | Signals used in APODIS predictor | 72 |
| 3.2 | Discharges used on Apodis training | 72 |
| 3.3 | Table of alarms for the second layer | 75 |
| 3.4 | Signals used in disruption databases with their directory of JPF signals in JET database | 78 |
| 3.5 | Recommended signals for disruption analysis | 78 |
| 3.6 | Standardisation parameters for databases during ILW campaigns <i>C28 – C34</i> | 85 |
| 3.7 | Count of discharges from database if ILW campaigns <i>C28 – C34</i> . 0-Non-disruptive shot, 1-non-intentional disruption, 2-intentional disruption | 87 |
| 3.8 | Count of discharges from ILW campaigns <i>C28 – C34</i> that cannot be considered safe neither disruptive | 87 |
| 3.9 | Count of discharges from ILW experimental campaigns <i>C28 – C34</i> for APODIS assessment. 0-Non-disruptive shot, 1-non-intentional disruption, 2-intentional disruption, 3-anomalous shot | 89 |
| 3.10 | APODIS results during ILW campaigns <i>C28 – C34</i> | 90 |
| 3.11 | Relation between DMV triggers and APODIS alarms in non-intentional disruptions without APODIS in close loop | 91 |

| | | |
|------|--|-----|
| 3.12 | CFC: <i>C23 – C27b</i> robustness result. Both types of signal failure provide similar outcomes | 96 |
| 3.13 | ILW: <i>C28 – C30</i> robustness result | 97 |
| 3.14 | AWTs and standard deviation (STD) for CFC and ILW campaigns results | 97 |
| 3.15 | Sliding window rates | 100 |
| 4.1 | Features for disruption time location as a classification approach | 104 |
| 4.2 | RPM-S1 results | 109 |
| 4.3 | RPM-S2 results | 110 |
| 4.4 | Results for disruption time location using UMEL. Simulation 1: plasma current and loop voltage. Simulation 2: plasma current, loop voltage and peaking factor. $ \Delta t = t_{dis} - t_{detect} $ | 113 |
| 4.5 | Values of levels for non-decimated wavelet and threshold that provide the best results | 117 |
| 5.1 | Best models for each set of signals and type of predictor | 126 |
| 5.2 | Results of APODIS offline and <i>P2-6</i> features | 127 |
| 5.3 | Results of disruption predictor trained with the hydrogen campaign | 129 |
| 5.4 | List of signals | 141 |
| 5.5 | Results for unbalanced approach from model 24 onwards | 141 |
| 5.6 | Results for balanced approach from model 24 onwards | 143 |
| 5.7 | Results for hybrid approach from model 24 onwards | 143 |
| 5.8 | Results for one layer predictor from scratch from model 24 onwards | 145 |
| 5.9 | Results for all ILW campaigns from scratch from model 40 onwards | 146 |
| 5.10 | Results for ILW campaigns <i>C28 – C33</i> from scratch from model 24 onwards | 147 |
| 5.11 | Results for ILW campaign <i>C34</i> from scratch | 148 |
| B.1 | Parkinson’s disease database | 162 |
| C.1 | Breast cancer database | 163 |

Chapter 1

Fusion, the energy of the future

In this chapter, a brief introduction to nuclear fusion, plasma confinement and fusion devices is exposed. The tokamak device, JET will be briefly explained and the main instability, disruptions, which happen in plasma at magnetic confinement devices.

1.1 Nuclear fusion

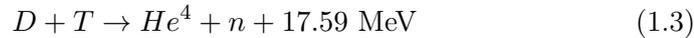
As it was mentioned in the introduction, nuclear fusion as renewable energy attempt to generate electricity reproducing the reaction which happens in the stars. Sir Arthur Eddington, in 1926, proposed that there is a limit where the pressure of the stars light on the atmosphere is large enough to balance the gravitational weight of the stellar atmosphere entirely, known as the Eddington luminosity limit. If the star were any brighter, the light of the star would push away the outer layers of the atmosphere, thus causing the star to lose mass. Eddington described stars as gas spheres at high temperature, mainly composed by hydrogen (Eddington, 1926). The pressure of gravitational forces can reach a temperature which turns gases into plasma (the fourth state of the matter). In 1927, Lewi Tonks and Irving Langmuir designated as *plasma*, a nearly neutral ionised gas in which ions and electrons were moving independently (Tonks and Langmuir, 1929). The temperature and pressure in the stars fuse hydrogen nuclei in plasma producing helium and energy.

There are many possible reactions to achieve fusion, but in all of them several light nuclei combine together into heavier and more stable nuclei. The difference in mass between the reactants and the products is converted to energy according to Einstein's formula, $\Delta E = \Delta mc^2$, where c is the speed of light. The intent is to use the kinetic energy from the reaction to produce electricity and to sustain new fusion reactions.

Unfortunately, nuclear fusion in practice is hard to attain. In random collision between two nuclei, they need extremely high energies and must

be close enough ($\approx 10^{-15}$) to overcome the repulsive electrostatic Coulomb force, known as the Coulomb barrier, and get over the strong nuclear force. To provide this energy to get close enough and fuse, their temperature must be raised until 10–100 keV¹. Under these conditions a thermonuclear plasma is obtained, it means that atoms are ionized and remain a gas formed by ions and electrons separated (Spitzer, 1962; Goldston and Rutherford, 1995). However, the closeness of the particles implies that the density of the fuel must also be very high. In a star, this density is maintained by the massive gravity, but on earth there's no conceivable material can confine the plasma and withstand such high temperatures. Hence, the confinement by walls is necessary.

The fusion of hydrogen atoms generates helium releasing 6×10^8 MJ per hydrogen Kg, this is above 60 millions more than the energy per Kg produced by chemical fuels (Wesson, 1999). The reactions used in fusion studies start with deuterium, either alone or mixed with the heavier hydrogen isotope, tritium. There are three possible reactions (Wesson, 1999):



where D is a deuteron, T a triton, H a proton, n a neutron and He^4 , the isotope of helium being an α - *particle*. The energies shown are the kinetic energy provided by the reaction products. It can be seen that tritium is produced in one reaction. Deuterium can be distilled from all forms of water. It is a widely available, harmless, and virtually inexhaustible resource. In every litre of seawater, for example, there are 33 mg of deuterium. Moreover deuterium is routinely produced for scientific and industrial applications. Nevertheless, tritium must be bred artificially: activated hydrogen in the water or during the fusion reaction through contact with lithium (when neutrons escaping the plasma interact with lithium contained in the blanket wall of the device).

In figure 1.1 (EUROFUSION[©]) is shown the cross section for reactions in equations 1.1, 1.2, 1.3. At temperatures of 100 keV, cross-section for $D - T$ reaction is much higher than the others. At this point, the real problem is to reproduce this reaction on Earth, where gravitational forces are 27.9 times smaller than the Sun. Furthermore, another problem is how to handle particles at temperatures tens of times larger than the Sun's temperature. For example, the temperature in the core of the Sun is 15 million K in

¹In nuclear physics, temperature is expressed in energy units, i.e. 1 eV = 1.16×10^4 K.

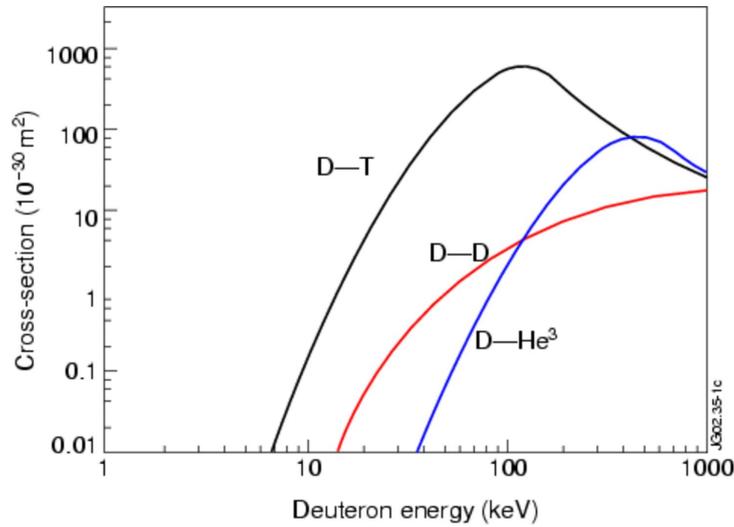


Figure 1.1: Cross-section for the most common reactions in plasma fusion

contrast to the temperature in the centre of the plasma in JET that can reach 200 million K. To overcome this problems, plasma properties are studied.

1.1.1 A brief approach to plasma physics

Plasma can thought of in two basic ways. The first one, how individuals electrons and ions behave whose motion is dominated by a magnetic field. The second one is to consider electrons and ions behaving as forming fluids. The electron and ion fluids then have their own temperatures and pressures, and also properties of gases as viscosity and thermal conductivity. It is thought sometimes ions and electrons as a single fluid to develop a simpler approach to analysing stability. The main properties of plasma are:

Basic parameters:

The temperature of the plasma is a way of describing the average kinetic energy of the particles. The kinetic temperature is invariably measured in electron-volts ($1 \text{ J} = 6.24 \cdot 10^{18} \text{ eV}$). Considering an idealized plasma with the same number of electrons, mass m_e and charge $-e$, and ions, with mass m_i and charge $+e$, the kinetic temperature would be,

$$T_s = \frac{1}{3} m_s v_s^2; \quad (1.4)$$

Assuming that both ions and electrons are characterized by the same T (which is, by no means, always the case in plasmas), we can estimate typical

particle speeds via the so-called thermal speed,

$$v_{ts} = \sqrt{2T/m_s}. \quad (1.5)$$

Thermonuclear plasmas have densities of a few times 10^{19} molecules per cubic metre, while the density of molecules in the atmosphere is about $3 \cdot 10^{25}$ molecules per cubic metre. Generally, density and temperature are functions of position in a plasma.

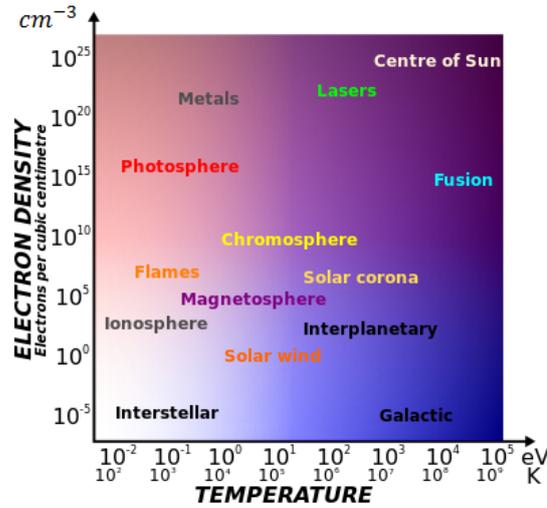


Figure 1.2: Range of plasmas

In figure 1.2 typical parameters of some plasmas can be seen and the enormous range of these magnitudes.

Plasma frequency and oscillations:

If we pull the electrons, with e charge, m_e mass and density η_e , in a region of the plasma ($x > 0$) slightly to the right a distance d , they oscillate with the electron plasma frequency:

$$\omega_p^2 = \frac{e^2 \eta_e}{\epsilon_0 m_e}. \quad (1.6)$$

A layer of the non-neutralized background with the charge density $\rho = -\eta_e e$ is left. The electric field, produced by this layer on the electrons on both edges is $E = 2\pi\rho d = -2\pi\eta_e e d$ (for the electrons at the right edge) and $E = 2\pi\rho d = 2\pi\eta_e e d$ (for the electrons at the left edge). The force $F = eE = -2\pi\eta_e e^2 d$ accelerates the electrons at the right edge to the left, while the electrons at the left edge experience similar acceleration to the

right. The relative acceleration of the electrons at the right and left edges would be $a = 2(eE/m_e) = -4\pi\eta_e e2d/m_e$. On the other hand, $a = \ddot{d}$, so that one has,

$$\ddot{d} = -w_p^2 d \rightarrow w_p^2 = \frac{4\pi\eta_e e^2}{m_e}. \quad (1.7)$$

The derived equation describes oscillations with the plasma frequency w_p . It should be emphasized that the motion is caused by the coordinated movement of many particles together and is thus a purely collective effect. In order to be able to observe these oscillations their period should be much smaller than the typical life time of the system.

Debye length and screening:

The Debye length is one of the most important and fundamental length scales in plasma physics. It describes a screening distance, beyond which charges are unaware of other charges inside of the Debye length. A charge in a plasma will attract opposite charges and repel like charges to the point that its electric field is screened by the charges it has attracted, so particles outside the screening charges are unaware of the presence of the interior charge. We are going to assume that ions and electrons have the same temperature T and density η . If we have a plasma with temperature T and density η , and a positive charge, e , is added, the particles will move around until they reach thermal equilibrium. So the probability of being in a state of energy U is proportional to Boltzmann factor $P(U) \approx e^{-U/T}$. At this point, given the potential of a single particle from the new charge $U = e_n V$, its distribution function would be,

$$f(u) = A e^{-(\frac{\mu^2}{2} + e_n V)/T} = \eta_0 \sqrt{\frac{m}{2\pi T}} e^{-(\frac{\mu^2}{2} + e_n V)/T}, \quad (1.8)$$

with η_0 as the equilibrium particle density, it means, before the arrival of the new charge. We can obtain the total particle number density by the integral,

$$\int_{-\text{inf}}^{\text{inf}} f(u) du = \eta \rightarrow \eta = \eta_0 e^{-e_n V/T}. \quad (1.9)$$

Once we have obtained the number density, the charge density is given by $\rho = \eta_i e - \eta_e (-e)$, the number density of ions minus the number density of electrons. With 1.9 using $-e$ rather than e_n ,

$$\rho = \eta_i e - \eta_e (-e) = e(\eta_i + \eta_e) = e\eta_0 (e^{-eV/T} - e^{eV/T}) = -e\eta_0 \sinh(eV/T). \quad (1.10)$$

Taking Posson's equation,

$$\nabla^2 V = -\frac{\rho}{\epsilon_0} = \frac{e\eta_0 \sinh(eV/T)}{\epsilon_0}. \quad (1.11)$$

A nonlinear second-order differential equation is obtained, albeit it is simplified by assuming the potential energy of the particles in the applied field is much smaller than their thermal energy, $eV \ll T$. Through a Taylor series it yields a second-order linear differential equation:

$$\nabla^2 V = \frac{e\eta_0}{\epsilon_0} \left(\frac{eV}{T} \right) = \frac{\eta_0 e^2}{T\epsilon_0} V. \quad (1.12)$$

At this point, we define the Debye length as:

$$\lambda_D = \sqrt{\epsilon_0 T \eta_0} e^2. \quad (1.13)$$

So equation 1.11 is now,

$$\nabla^2 V = \frac{V}{\lambda_D^2} \rightarrow V = V_0 e^{-x/\lambda_D}. \quad (1.14)$$

where the decaying solution follows the condition of $V \rightarrow 0$ and $x \rightarrow \text{inf}$. The meaning of λ_D is that charges feel the potential due to the central charge inside Debye's length; and the potential falls off exponentially without any presence of the central charge, outside the Debye's length. It can be easily explained, if you arrive late to a concert crowd of people, probably you will not be able to see anything because people will be screening you.

Plasma parameter:

Firstly, we are going to define the average distance between particles as

$$r_d = \eta^{-1/3}, \quad (1.15)$$

and the distance of closest approach as,

$$r_c = \frac{e^2}{4\pi\epsilon_0 T}. \quad (1.16)$$

The distance of closest approach, r_c , is the distance at which the Coulomb Energy, $U(r, v)$, of one charged particle in the electrostatic field of another vanishes, i.e., $U(r_c, v_t) = 0$. Attending to the ratio:

- $r_c/r_d \ll 1$: plasmas are termed strongly coupled. Charged particles are dominated by one another's electrostatic influence more or less continuously, and their kinetic energies are small compared to the interaction potential energies.
- $r_c/r_d \gg 1$: strong electrostatic interactions between particles are rare events. The common is that a particle is electrostatically influenced by all of the others within its Debye sphere. Albeit, this interaction does not cause usually a change in its motion and plasmas are termed weakly coupled.

A weakly coupled plasma can be described using the same type of equation which is usually used for a neutral gas, a standard Fokker-Planck equation. However, it is more difficult to understand strongly coupled plasma, which has more in common with a liquid. After that, the plasma parameter is defined as:

$$\Lambda = 4\pi\eta\lambda_D^3. \quad (1.17)$$

The plasma parameter is equal to the number of particles in a Debye sphere. This dimensionless parameter can be expressed combining 1.15, 1.16 and 1.17:

$$\Lambda = \frac{\lambda_D}{r_c} = \frac{1}{\sqrt{4\pi}} \left(\frac{r_d}{r_c} \right)^{3/2} = \frac{4\pi\epsilon_0^{3/2} T^{3/2}}{e^3 \eta^{1/2}}. \quad (1.18)$$

From equation 1.18, it can be deduced:

- If $\Lambda \ll 1$, we have strongly coupled plasma and Debye sphere would be slightly populated. These plasmas tend to be cold and dense. For example: solid-density laser ablation plasmas, the very 'cold' (with kinetic temperatures similar to the ionization energy) plasmas found in 'high pressure' arc discharges, and the plasmas which constitute the atmospheres of collapsed objects such as white dwarfs and neutron stars.
- If $\Lambda \gg 1$, we have a weakly coupled plasma and Debye sphere is highly populated. These plasmas tend to be diffuse and hot. For example: plasmas typically encountered in ionospheric physics, astrophysics, nuclear fusion, and space plasma physics

| Plasma \ Parameter | $\eta[\text{m}^{-3}]$ | $T[\text{eV}]$ | $w_p[\text{s}^{-1}]$ | $\lambda_D[\text{m}]$ | Λ |
|----------------------|-----------------------|----------------|----------------------|-----------------------|-------------------|
| Interstellar | 10^{-6} | 10^{-2} | $6x10^4$ | 0.7 | $4 \cdot 10^6$ |
| Solar chromosphere | 10^{18} | 2 | $6 \cdot 10^{10}$ | $5 \cdot 10^{-6}$ | $2 \cdot 10^3$ |
| Solar wind (1AU) | 10^7 | 10 | $2 \cdot 10^5$ | 7 | $5 \cdot 10^{10}$ |
| Ionosphere | 10^{12} | 0.1 | $6 \cdot 10^7$ | $2 \cdot 10^{-3}$ | $1 \cdot 10^5$ |
| Arc discharge | 10^{20} | 1 | $6 \cdot 10^{11}$ | $7 \cdot 10^{-7}$ | $5 \cdot 10^{12}$ |
| Tokamak | 10^{20} | 10^4 | $6 \cdot 10^{11}$ | $7 \cdot 10^{-5}$ | $4 \cdot 10^8$ |
| Inertial confinement | 10^{28} | 10^4 | $6 \cdot 10^{15}$ | $7 \cdot 10^{-9}$ | $5 \cdot 10^4$ |

Table 1.1: Main parameters for some typical weakly coupled plasmas

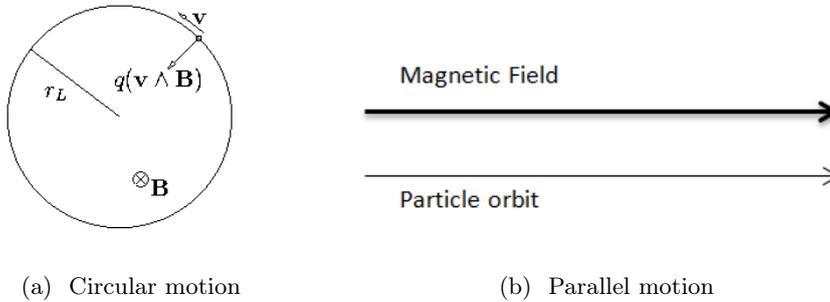
In table 1.1 are shown the key parameters for some typical weakly coupled plasmas.

Magnetic field effects:

The motion of a charged particle inside a magnetic field is very important in thermonuclear plasmas. In uniform magnetic fields, the motion of

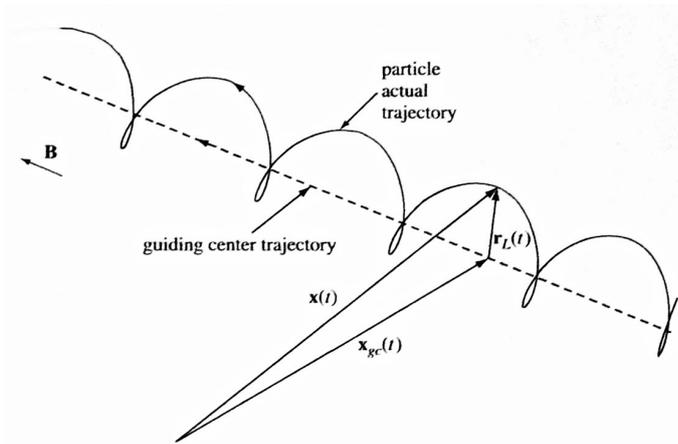
the charged particle has two parts. The first one is a circular motion perpendicular to the magnetic field, it can be seen in figure 1.3a. The radius is called Larmor radius, which increases with the energy of the particle and decreases with the strength of the magnetic field. For example, a typical plasma in JET has a Larmor radius of few millimetres. The electrons and ions go in opposite directions because of their charges. The second motion is along the magnetic field, the charged particle's motion parallel to the field is constant; figure 1.3b. Once the two motions are combined, an helical motion is obtained as it is shown in figure 1.3c.

If the magnetic field becomes stronger along its trajectory, the particle receives a force in the direction away from the stronger field. If the magnetic field increases enough, the particle can be reflected along its path, this is called “mirror force”. This force can be seen in figure 1.3c.



(a) Circular motion

(b) Parallel motion



(c) Compound motion: helical trajectory

Figure 1.3: Trajectories of a charged particle in a magnetic field

The fundamental measure of a magnetic field's effect on a plasma is the magnetization parameter δ . The fundamental measure of the inverse effect is called β , which is the ratio of the thermal energy density ηT to the magnetic energy density $B^2/2\mu_0$. Conventionally, plasma energy density is expressed

as the pressure, $p = \eta T$. So it yields,

$$\beta = \frac{2\mu_0 p}{B^2}. \quad (1.19)$$

Collisions:

In plasmas, collisions between charged particles are widely different than typically collisions between molecules in a normal gas. The reason is that in a normal gas the collision is produced by the contact between the molecules, but plasma particles collide with distant particles. Coulomb force leads the interaction between charged particles in the plasma; this force falls off with the inverse square of the distance between particles, hence this long range interaction makes that any particle collides simultaneously with many other particles. For example, at JET, each particle is colliding at the same time with millions of other particles.

Therefore, collisions are described by the global effect of all the separate collisions. There are two types of collisions. Although each distant collision causes only a small deflection of the particle trajectory, they are largely present that their total action upon the particle is greater than relatively rare close collisions. It means that the influence of close collisions on kinetic processes in plasma is, generally, negligibly small in comparison to the action of distant collisions.

The theory briefly exposed has been obtained from many sources (Wesson, 1999; Spitzer, 1962; Goldston and Rutherford, 1995; Miyamoto, 2000; Fitzpatrick, 2014).

1.1.2 Fusion Devices

First of all, it is necessary to define the conditions at fusion reactors can reach ignition, it means, the heating of the plasma by the products of the fusion reactions is sufficient to maintain the temperature of the plasma against all losses without external power input. In 1957, J. D. Lawson showed that the product of ion density and confinement time determined the minimum conditions for productive fusion, and that product is commonly called Lawson's criterion. It is typically used the triple product, it specifies that successful fusion will be achieved when the product of the three quantities, η , the particle density of a plasma, τ , the confinement time and T the temperature, reaches a certain value. Above this value of the triple product, the fusion energy released exceeds the energy required to produce and confine the plasma. For deuterium-tritium fusion this value is about :

$$\eta\tau T \geq 5 \cdot 10^{21} m^{-3} s K eV. \quad (1.20)$$

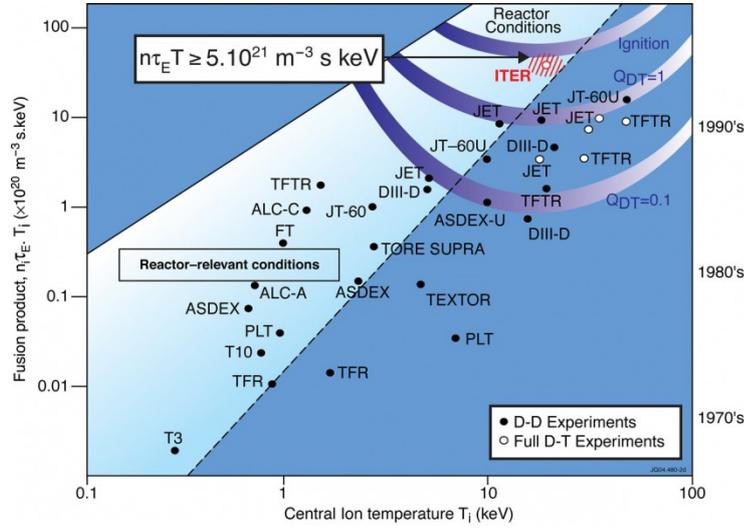


Figure 1.4: Progress of the fusion “triple product” of plasma ion density, ion temperature and energy confinement time

Figure 1.4 shows the fusion triple product achieved on different magnetic fusion facilities. Notice that the unit on the temperature scale is keV, and that at very high temperatures the difference between Kelvins ($^{\circ}K$) and degrees Celsius ($^{\circ}C$) is negligible. The graph shows clearly that new facilities performed better than previous ones. The present large machines, from the point of view of the fusion product, have now achieved their engineering limits so that only the next step facility, ITER, can bring about decisive progress.

Currently, fusion research is practically based on two approaches: inertial and magnetic confinement. In inertial confinement, an hydrogen sphere is compressed by a laser to such a high temperatures and densities that Lawson criterion is satisfied. The existing methods for inertial confinement, generally follow three steps, figure 1.5. Firstly, the hydrogen sphere is heated by an external power source (generally, lasers). This heating blows off the outer layers of the sphere and compresses the core to $1/1000^{th}$ of its original volume. The external pulse sends a shock wave to compress the centre of the sphere. The wave, without any hydrodynamic instabilities, will be able to produce high temperatures and pressures in the centre of the sphere, reaching densities about 10^{23} m^{-3} and temperatures about tens of KeV. These conditions are maintained during very short periods of time (nanoseconds).

The National Ignition Facility (NIF) probably has been the best approach to achieve inertial confinement fusion at the moment. Over a decade in the works, it began experiments at full power earlier 2010 (Glenzer et al., 2010). However, these devices requires high lasers energy whose development and

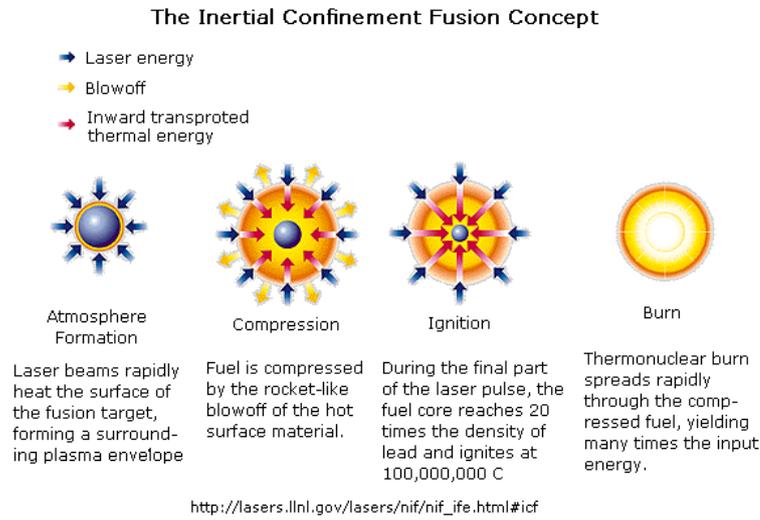


Figure 1.5: Inertial confinement procedure

application in a continual cycle are far to be solved (Dendy, 1995).

At magnetic confinement we can find two different devices, stellarators and tokamaks. In 1951, Lyman Spitzer designed a new configuration called stellarator (steady-state generator) (Spitzer, L. , 1951). This configuration is based on a magnetic field produced by solenoidal coils around the toroidal tube. The configuration is characterized by a rotational transform, it means, such a line magnetic field wrapping around the plasma surface, rather than simply go around as a set of individual disconnected lines (Spitzer, 1958). This configuration allows a stationary operation and current external control. The complexity of the geometry wraps the magnetic field describing an helical trajectory. This configuration is illustrated on figure 1.6. There are different types:

- Torsatron: configuration with continuous helical coils or discrete coils producing similar fields.
- Heliotron: configuration with a helical coil is used to confine the plasma, together with a pair of poloidal field (PF) coils to provide a vertical field. Toroidal field (TF) coils can also be used to control the magnetic surface features.
- Modular: configuration with a set of modular (separated) coils and a twisted toroidal coil (Johnson, 1999).
- Heliac: helical axis stellarator; configuration where the complexity of the geometry wraps the magnetic field describing an helical trajectory.

Typically can provide more twist than the Torsatron or Heliotron, especially near the centre of the plasma (magnetic axis). The original Heliac consists only of circular coils, and the flexible heliac (Harris et al., 1985) (H-1NF, TJ-II, TU-Heliac) adds a small helical coil to allow the twist to be varied by a factor of up to 2.

- Helias: helical advanced stellarator; configuration which use an optimized modular coil set designed to simultaneously achieve high plasma, low Pfirsch-Schluter currents and good confinement of energetic particles; i.e., alpha particles for reactor scenarios. The Helias has been proposed to be the most promising stellarator concept for a power plant, with a modular engineering design and optimised plasma, MHD and magnetic field properties. The Wendelstein VII-X device is based on a five field-period Helias configuration.

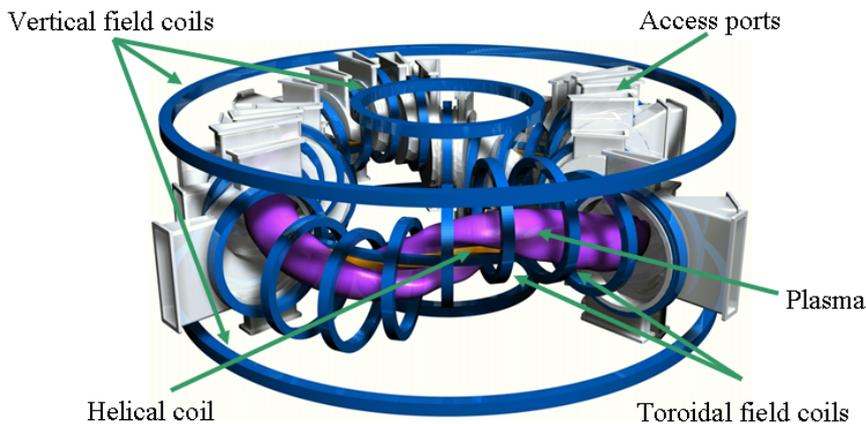


Figure 1.6: Stellarator configuration (TJ-II, Ciemat, Spain)

The absence of induced plasma current makes stellarator to be a great option for a commercial power-plant, because it is a configuration which works stationary and without disruptions. This last concept will be explained later. Moreover, stellarators have a wider range of possible configuration and a better control than tokamaks. However, it is difficult to calculate and engineer the correct shape for the helically twisted coils; only with the advent of powerful computers and sophisticated manufacturing techniques could stellarators become a commercial power-plant. Other disadvantages are the high electromagnetic forces between coils. In table 1.2 are shown some of the most relevant stellarators devices nowadays.

| Name | Location |
|-------|-----------|
| H1 | Australia |
| LHD | Japan |
| TJ-II | Spain |
| HSX | USA |
| NCSX | USA |
| W7-X | Germany |

Table 1.2: Stellarators devices

1.1.3 Magnetic confinement devices, Tokamaks

The other magnetic confinement devices, tokamaks, will be explain widely. In particular, the Joint European Torus (JET), because the topic of this thesis uses the data provided by the JET experiments. In this devices, a toroidal current is induced through the plasma itself by a central solenoid. This current produces a poloidal magnetic field, whose field lines pass around the plasma, being able to hold the plasma in place and provide an equilibrium force balance. In figure 1.7, these fields are illustrated and the helical field lines from the combination of the toroidal and poloidal magnetic fields.

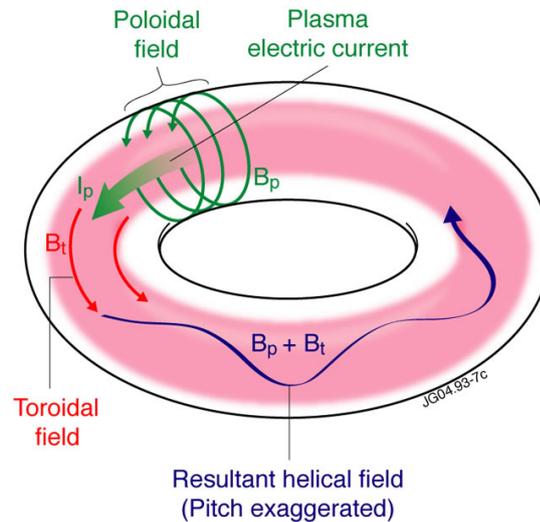


Figure 1.7: Current and fields in a tokamak plasma

By 1950, two scientifics, Sakharov and Tamn, from the Soviet Union designed the tokamak configuration. The meaning of tokamak comes from “toroidalnaya kamera ee magnitnaya katushka” in Russian, that could be translated as “toroidal chamber with toroidal coil”. In (Leontovich, 1961), Shakarov came up with the idea of confining plasma in a toroidal chamber

with a strong longitudinal magnetic field and Tamn proposed the general methods of solving the kinetic equation for toroidal plasmas in the presence of induced current and showed that thermal plasma conductivity in a torus may be higher than in a cylinder for equal magnitudes of a longitudinal magnetic field and induced current magnetic field. Although a tokamak is more robust and easier to build from an engineer point of view, there are two significant problems: the confinement of the plasma is not as good as the theory explains and the appearance of violent instabilities, called disruptions. The first tokamak was the TMF in 1955, whose parameters were: $R = 0.8$ m, $a = 0.13$ m, $B_t = 1.5$ T and $I = 0.26$ MA, where R is the major radius, a is the minor radius, B_t the magnetic field and I the induced current. The plasma volume was $V = 0.27$ m³. From 1955 until 1990 many devices have been developed (Azizov, 2012):

- During 1955-1969: eight facilities similar to TMF were built at the Kurchatov Institute, T-1, T-1, T-3, T-5. TM-1, TM-2, TM-3 and T-3A. Nearly every year, a facility was constructed.
- During 1970-1990: the first tokamak outside the USSR, LT-3 in Canberra (Australia). In all, over 200 tokamaks have been constructed in the world to date, including 31 in the USSR and Russia, 30 in USA, 32 in Europe and 27 in Japan.

The most important tokamak experimental devices currently in operation can be seen in table 1.3 and their principal features.

Furthermore, there are two future approaches, ITER and DEMO. DEMO is a proposed nuclear fusion power plant that is intended to build upon the expected success of the ITER experimental nuclear fusion reactor.

| Parameters \ Device | Alcator C-Mod | DIII-D | NSTX |
|------------------------------------|----------------|--------------|-------------------------|
| Location | USA | USA | USA |
| Year | 1991 | 1986 | 1999 |
| $R(m);R/a$ [range] | 0.67; 3.04 | 1.66; 2.47 | 1; 1.6 |
| $B(T)$ [range] | 8 | 2.2 | 0.3 |
| $I(MA)$ [max] | 2 | 3 | 1.4 |
| Plasma Volume(max,m ³) | 1 | 24 | 14 |
| Pulse Length (s) | 5 | 25 | 3.5 |
| Plasma Facing Materials | Solid Mo and W | C | CFC/Graphit, Li coating |
| Parameters \ Device | ASDEX-U | FTU | JET |
| Location | Germany | Italy | UK |
| Year | 1991 | 1990 | 1983 |
| $R(m);R/a$ [range] | 1.65; 3.2 | 0.94; 3.2 | 2.96; 3.1 |
| $B(T)$ [range] | 3.1 | 8 | 4 |
| $I(MA)$ [max] | 2 | 1.6 | 6 |
| Plasma Volume(max,m ³) | 13 | – | 100 |
| Pulse Length (s) | 10 | 1.7 | 60 |
| Plasma Facing Materials | C, W-coated | Mo limiters | C, Be, W-coated |
| Parameters \ Device | JT-60U | MAST | TCV |
| Location | Japan | UK | Switzerland |
| Year | 1985 | 1999 | 1992 |
| $R(m);R/a$ [range] | 3.2; 4 | 0.8; 1.3 | 0.86; 3.6 |
| $B(T)$ [range] | 4.8 | 0.52 | 1.43 |
| $I(MA)$ [max] | 5 | 1.4 | 0.8 |
| Plasma Volume(max,m ³) | 90 | 8 | – |
| Pulse Length (s) | 10 | 1 | 2 |
| Plasma Facing Materials | C | C | C |
| Parameters \ Device | KSTAR | EAST | |
| Location | Korea | China | |
| Year | 2008 | 2006 | |
| $R(m);R/a$ [range] | 1.8; 3.6 | 1.9; 4.2 | |
| $B(T)$ [range] | 3.5 | 3.5 | |
| $I(MA)$ [max] | 2 | 1 | |
| Plasma Volume(max,m ³) | 17.8 | 38 | |
| Pulse Length (s) | 300 | 1000 | |
| Plasma Facing Materials | CFC/Graphite | CFC/Graphite | |

Table 1.3: Tokamaks devices and features

1.1.3.1 Joint European Torus - JET

Currently, Joint European Torus (JET) is the largest nuclear fusion experimental device in the world, located in Culham, Oxfordshire, United Kingdom. The design was started in 1973 by P.H.Rebut and was completed with the publication “The Jet Design Proposal” (Team, 1976), in 1975. This document presented an impressively detailed design of the components. The structure in the design was almost identical to the device which would ultimately be built; the original design can be seen in figure 1.8. The construction of JET lasted 4 years, 1979 – 1983. The first pulse of JET was on 25th June, 1983.

The JET proposal was described in the design proposal: *the essential objective of JET is to obtain and study a plasma in conditions and dimensions approaching those needed in a thermonuclear reactor. These studies will be aimed at defining the parameters, the size and the working conditions of a Tokamak reactor. The realisation of this objective involves four main areas of work:*

1. *the scaling of plasma behaviour as parameters approach the reactor range,*
2. *the plasma-wall interaction in these conditions,*
3. *the study of plasma heating and*
4. *the study of alpha – particle production, confinement and consequent plasma heating.*

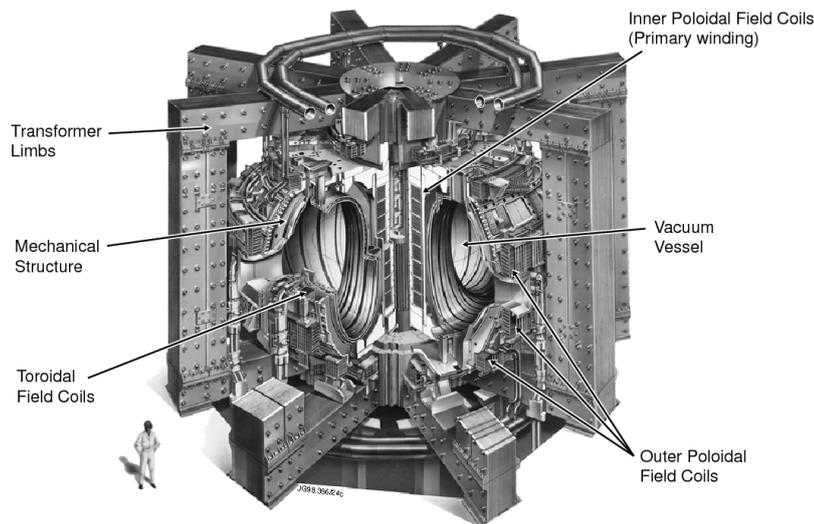


Figure 1.8: Original JET design

The vacuum vessel can hold a plasma volume in the range of $83 - 100 \text{ m}^3$. At first, it was designed to reach a plasma current of 3.8 MA , sufficient to confine α – particles, but then it was designed to reach 4.8 MA . Initially the toroidal magnetic field was 2.8 T and then it was extended to 3.5 T . In 2009 the carbon components (known as CFC-wall, Carbon Fiber Composite) were replaced with tungsten and beryllium ones (ITER-like wall, ILW), to bring JET's components more in line with those planned for ITER.

The toroidal magnetic field is produced by 32 D-shaped coils enclosing the vacuum vessel, whose layout is illustrated in figure 1.9a. Each coil is wound with 24 turns of copper bar and weigh 12 tonnes. The combined

current carrying capacity of all the coils is 51 MA. Each coil can receive a total force approximately about 2000 tonnes, directed toward the major axis of the torus. The D-shaped design is due to the magnetic field inside the toroidal field coils falls off with the major radius, $1/R$, so there is a larger magnetic force on the inboard side of the coils than the outboard side.

The poloidal field coils are 7 horizontal circular coils placed outside the toroidal field coils, as it can be seen in figure 1.9b. The main poloidal field coil is the inner coil wound round the central column of an iron transformer core, to act as the primary of the transformer. The other 6 coils are optimally placed to provide control of the plasma shape and position. The largest coil is 11 metres diameter.

The iron transformer core is formed by 8 limbs which envelope the other components. This structure weighs about 2600 tonnes, figure 1.9c.

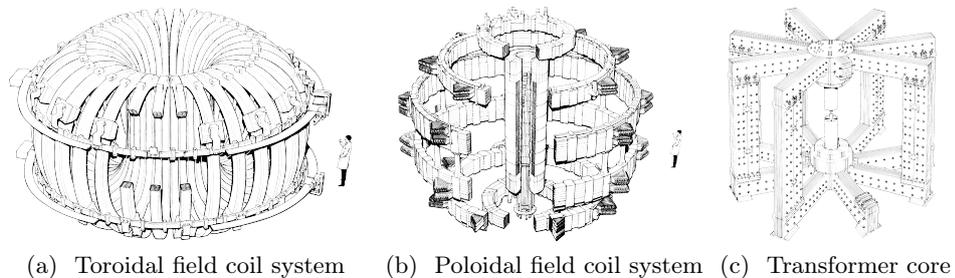


Figure 1.9: JET coil systems and iron transformer (EUROFUSION[©])

The plasma shape can vary from circular to elliptical shape. This shape is calculated from the magnetic field measured by poloidal coils. This shows the ability to produce both circular and elongated plasmas. With plasma elongation a separatrix is introduced into the magnetic geometry. This separates the internal closed flux surfaces from external open surfaces. This change of geometry implies that there are *X-points* on the separatrix flux surface, figure 1.10. These open surfaces are used to lead the escaping plasma to a receiving material surface remote from the plasma. The system is called, divertor. This type of operation with higher confinement is called H-mode, in contrast to the normal lower confinement known as L-mode.

Jet was designed to generate a pulse every 15 minutes. Each pulse needs a total power of up to 800 MW. The power supplies used at JET combine heavy flywheels driven up to speed by motor generators and several hundred MW taken directly from the electricity grid.

In JET, three plasma heating systems are used:

- Neutral beam injection (NBI): it supplies 23 MW power in total. High energy hydrogen beams are injected in plasma increasing its temperature. The particles injected in the plasma must be neutral particles,

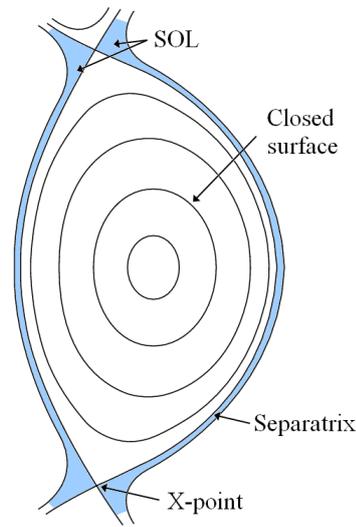


Figure 1.10: Plasma X-point

otherwise they will not be able to pass through magnetic field.

- ICRH antennae: it supplies a total power of 32 MW. High frequency electromagnetic waves are generated by an antennae in the vacuum vessel. Ion-Cyclotron Resonant Frequency (ICRH) system uses a frequency between 23 MHz (wavelength $\lambda = 13$ m) and 57 MHz ($\lambda = 5$ m) resonant to the second harmonic of deuterium or first harmonic of tritium and helium.
- LHCD antennae: power supply of 12 MW. Lower-Hybrid Current Drive (LHCD) heating uses a frequency of approximately 3.7 GHz ($\lambda = 0.1$ m) in between ion and electron cyclotron frequencies.

JET has the most complete set of diagnostics, about 100 different methods operating at the moment and new more diagnostics are been developing or upgrading for future experiments. JET also serves as a unique test bed for the development of diagnostics for the future fusion experimental device, ITER. A layout of the JET diagnostic systems is illustrated in figure 1.11.

At JET, signals from all diagnostic systems are digitised and stored in a central database. Every JET pulse produces almost 60 GBytes of raw diagnostics data. Most of the data need further processing, this is done automatically where possible by dedicated computer codes, but in many cases human intervention and/or data validation is required. All data are accessible to all scientists on the JET site and, moreover, any scientist from any EUROfusion Research Unit has remote access to the data from their home institute. These information has been obtained from (Wesson and Campbell,

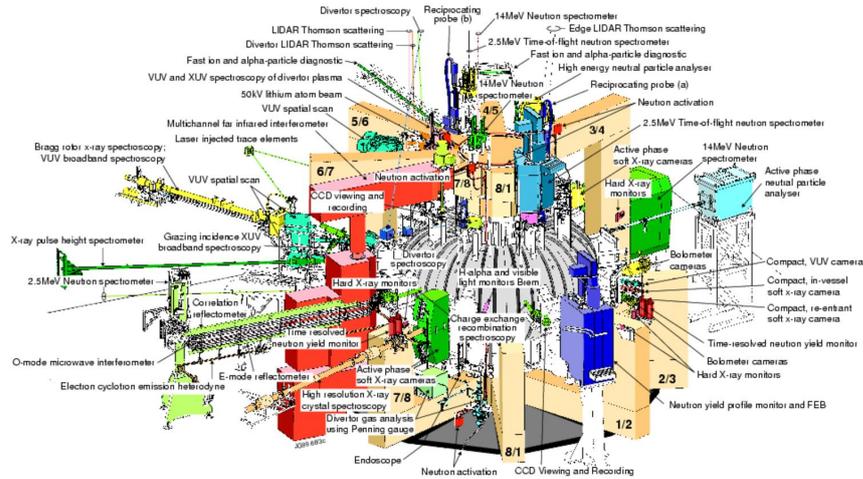


Figure 1.11: Layout of JET diagnostic systems

1987), (Wesson, 1999) and (Murari, 2014), where more information can be found.

1.1.3.2 ITER and DEMO

ITER is the largest and most advanced fusion experiment in the world. It is designed to generate 500 MW power, which is equivalent to the capacity of a medium size power plant, albeit, it will be an experimental fusion device. The fusion gain would be $Q = 10$, according to 50 MW injected power. ITER is also expected to be a world reference for the main technologies needed in a fusion power plant. Currently, ITER is being built in Cadarache, southern France, in the framework of a collaboration between China, Europe, India, Japan, Korea, Russia and USA. The complete tokamak assembly is expected by 2019, the first plasma would be produced in 2020, and the first Deuterium-Tritium operation would be in 2027.

The design of ITER is illustrated in figure 1.12(ITER[©]). Its main features are described in table 1.4.

ITER is an important step in the roadmap toward a first fusion reactor. The next device, DEMO(Demonstration Power Plant) will demonstrate the large-scale production of electrical power and tritium fuel self-sufficiency. A conceptual design of the device could be finished by 2017. The purpose, without any problem in the schedule, is to begin DEMO operations in the early 2030s, and to put power from nuclear fusion into the grid by 2040. Meanwhile, several devices will be operating around the world giving complementary research and development in support of ITER.

| | |
|---|------|
| Major radius, R (m) | 6.2 |
| Minor radius, a (m) | 2 |
| Toroidal field at R , B_T (T) | 5.3 |
| Plasma current, I_p (MA) | 15 |
| Safety factor, q_{95} | 3 |
| Average ion temperature, $\langle T_i \rangle$ (keV) | 8 |
| Average electron temperature, $\langle T_e \rangle$ (keV) | 8.8 |
| Average electron density, $\langle \eta_e \rangle$ (10^{19} m^{-3}) | 10.1 |
| Fusion power, P_{fusion} (MW) | 400 |
| Fusion gain, Q | 10 |
| Pulse length (s) | 400 |

Table 1.4: ITER features

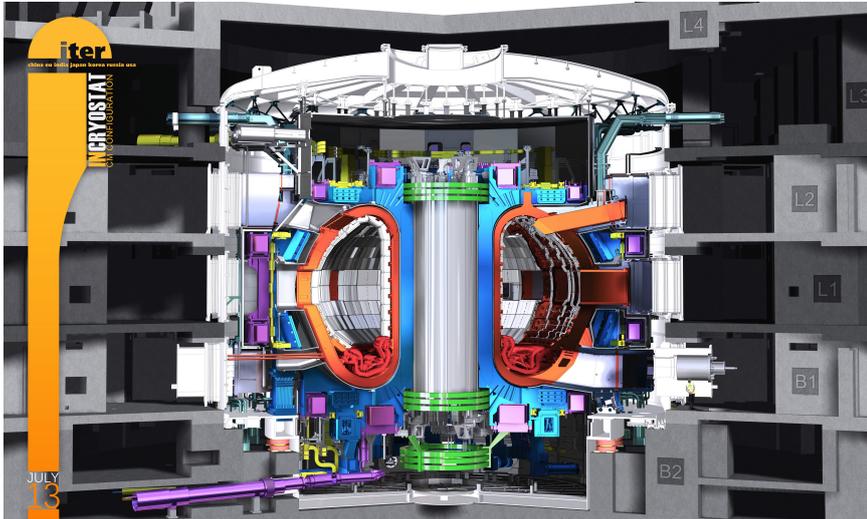


Figure 1.12: ITER design

ITER is not an end in itself: it is the bridge toward a first plant that will demonstrate the large-scale production of electrical power and tritium fuel self-sufficiency. This is the next step after ITER: the Demonstration Power Plant, or DEMO for short. A conceptual design for such a machine could be complete by 2017. If all goes well, DEMO will lead fusion into its industrial era, beginning operations in the early 2030s, and putting fusion power into the grid as early as 2040.

In the web page of ITER more detailed information can be found ².

²<http://www.iter.org/>

1.2 Plasma phenomena

The study of plasma operation is necessary to reach stable conditions for thermonuclear fusion, however plasma operation is not smooth and several events can appear. High plasma transport losses appear when high power auxiliary heating is used. This phenomena was showed the first time in 1980, and it was called low/L-mode confinement. A new confinement regime was found in 1982 (Wagner et al., 1982), it was called high/H-mode confinement. This mode is 2–3 times longer than L-mode and contrary to L-mode there is no confinement degradation with heating (Keilhacker, 1987). The variation of the parameters from the L-mode to the H-mode is called L/H transition. The creation of an edge transport barrier (ETB) characterizes the transition, and the plasma returns to the L-mode if the ETB is lost. For example this phenomena is helpful to achieve better plasma performance, but there are several instabilities which are harmful and potentially dangerous for the integrity of the device, therefore their mitigation is necessary. The most important and dangerous instabilities are disruptions, which are explained in section 1.2.1. The are other important events, although they are not fully understood they can mainly be attributed to identifiable magnetohydrodynamics (MHD) modes:

-Mirnov oscillations: during the start-up phase of the discharge, when the plasma current is rising, magnetic fluctuations occur. Mirnov discovered these oscillations using magnetic coils around the plasma surface. A decreasing sequence of m numbers identifies this event, and each m -number component arises from a distortion of the magnetic structure within the plasma due to a MHD instability with the same m -number (Wesson and Campbell, 1987).

-Sawtooth oscillations: the intensity of the soft X-rays emitted by plasma depends on the electron temperature and density, and these soft X-rays show a periodic relaxation in the core of the plasma (Wesson and Campbell, 1987; Wesson, 1999). The fluctuations were first observed in Princeton, in 1974 (von Goeler et al., 1974). A typical sawtooth cycle shows three phases:

1. The ramp phase where plasma density and temperature increase linearly with respect to time.
2. Precursor phase, a helical magnetic perturbation grows until the last phase.
3. The fast collapse phase, when the temperature and density drop fast (Chapman, 2011).

-Magnetic islands and tearing modes: MHD instabilities usually produce changes in the topology of the magnetic fields. This occurs at surfaces with rational values of q , the lines of the magnetic field are broken and reconnected generating magnetic islands, as it is illustrated in figure 1.13. Magnetic islands are generally associated with resistive instabilities and particularly tearing modes, firstly described in (Rutherford, 1973). Nevertheless, plasma is not perfectly conducting and magnetic islands can appear in the nonlinear development of all MHD instabilities with a resonant surface in the plasma. Two types of tearing modes can be discussed, the internally driven tearing, analysed in the context of disruptions in (Rebut and Hugon, 1985), and the neoclassical tearing mode, analysed in (Carrera et al., 1986) and firstly reported in (Chang et al., 1995).

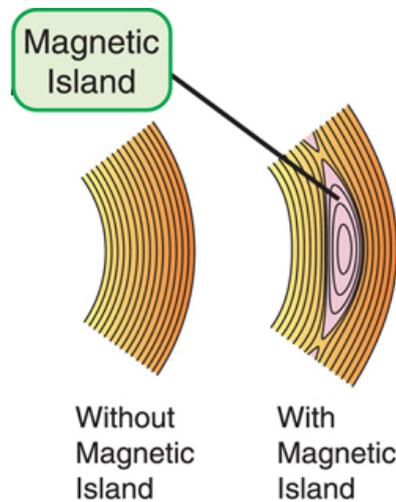


Figure 1.13: Reconnection of magnetic fields producing a magnetic island

-Edge Localized Modes, ELMs: In 1982, bursts or spikes were seen in H_α signal during the operation, in H-mode, in ASDEX (Wagner et al., 1982). These spikes are associated with and MHD instability in the edge of plasma, hence their name, edge localized modes (ELMs) (Keilhacker et al., 1984). These instabilities, associated with H-mode in tokamaks, are recognized as bursts in H_α or D_α signal which cause a drop in the density and temperature on the plasma edge, therefore it leads a deterioration of the plasma confinement. Firstly, in 1991 three different types of ELMs were identified at DIII-D tokamak (Doyle et al., 1991), an example of each type is illustrated in figure 1.14:

- Type I, giant ELMs: the plasma edge is close to the stability limit

$\alpha \approx \alpha_{limit}$. The repetition of these ELMs is proportional to the power and target density, and drops with increasing current. They appear as large isolated spikes on H_α or D_α signal. These ELMs are the most dangerous due to the heat loss and the divertor high heat (Wesson and Campbell, 1987).

- Type II, grassy ELMs: they appear when the plasma edge is in the connection regime between the first and the second stable ballooning regimes. The energy lost and amplitude is lower and the frequency is higher.
- Type III: these ELMs have similar amplitude to grassy ELMs, the frequency decreases as the power increases. They are associated to a deterioration of confinement.

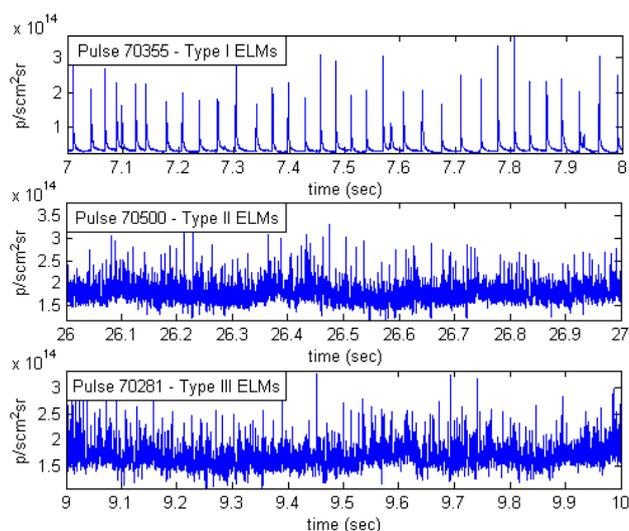


Figure 1.14: Example of ELMs types

The fact that ELMs of type II appear only under certain conditions, a new classification is proposed in (Zohm, 1996):

- Type I: the ELM repetition frequency ν_{ELM} increases with the energy flux through the separatrix³ (P_{sep}):

$$\frac{\nu_{ELM}}{P_{sep}} > 0. \quad (1.21)$$

³Separatrix: last closed flux surface. Surface that divides the plasma confined by the closed magnetic lines from the Scrape Off Layer (See figure 1.10).

- Type III: the ELM repetition frequency decreases with the energy flux through the separatrix:

$$\frac{\nu_{ELM}}{P_{sep}} < 0. \quad (1.22)$$

- Dithering cycles: for $P_{sep} \approx P_T$ where P_T is the power threshold for L/H transitions. L-H-L transitions may occur and the repetition frequency shows a slight decrease with increasing P_{sep} .

-Multifaceted Asymmetric Radiation From the Edge, MARFEs:

the MARFE is a radiation instability which appears in tokamaks as toroidally symmetric ring of enhanced radiation. This instability usually appears in the inner side of the torus, albeit it also appears around X-point of divertor configurations. The cause is generally the plasma cooling by radiation. The observation of a cold, dense region on the inboard side of the plasma was first reported in (Baker et al., 1982). Then, more observations, analysis and the name as MARFE appeared in (Lipschultz et al., 1984).

1.2.1 Disruptions, the most danger instability

Disruptions, plasma phenomena observed in (Gorbunov and Razumova, 1964), are the most dangerous instabilities in tokamaks, which lead the loss of the confinement and a abruptly drop of the plasma current (Wesson et al., 1989; Hender et al., 2007). The device is exposed to intense heat loads when the confinement is lost, hence, the different components and wall can be seriously damaged by the high temperature and current. This is a serious problem for tokamaks development, for the damage and the limit in the range of operation in current and density. Disruptions can overcome suddenly without precursors in a few milliseconds or several instabilities in hundreds milliseconds can finish disrupting. It is generally accepted that a disruption evolves through four phases (Schuller, 1995; Iter Physics Expert Group on Disruptions et al., 1999):

1. **Initial phase:** a change in plasma conditions leads to a less stable configuration, where plasma current and density slightly increase. This change is caused by external factors (operational or mechanical failures) or internal plasma factors (instabilities), however the variation on plasma conditions are almost imperceptible and practically undetectable.
2. **Precursor phase:** the onset changes in plasma configuration reach a critical point due to new increasing and detectable MHD instabilities. The importance of detecting early enough this phase is crucial. If the phenomena is detected with time enough, mitigation actions are able to be taken in order to prevent or reduce the effects. Theoretically

this instabilities start as magnetic islands rotating around the magnetic plasma axis, and they grow exponentially in tens of milliseconds leading into disruption. A model of this phase was developed at Culham and summarized in (Wesson et al., 1985).

3. **Fast phase (thermal quench)**: a second critical point happens, the radial current profile is flattened and the central temperature drops drastically in a few milliseconds.
4. **Quench phase (current quench)**: the magnetic energy of the poloidal magnetic field is dissipated as thermal energy. The deposition of this energy can be harmful for the integrity of the vessel. The decay time depends on the particular conditions and plasma position.

An example of a typical disruption at JET and the phases described previously is illustrated in figure 1.15.

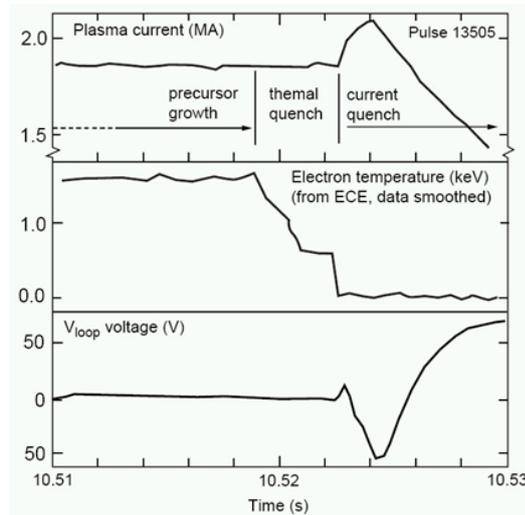


Figure 1.15: Example of a disruption at JET

Unless the physics of disruptions is not understood in detail, the general pattern of the behaviour can be described in some basics events:

1. The current profile becomes unstable leading to the growth of a tearing mode, $m = 2$ is particularly important.
2. Non-linear growth of the tearing mode.
3. Sudden relaxation of the equilibrium, the current profile is flattened and there is a dramatic loss of confinement with a collapse of the plasma temperature.

4. Plasma current decays.
5. Under some circumstances, the increased plasma resistance and toroidal electric field produce runaway electrons. These electrons can carry large currents which can persist after the current quench phase.
6. The loss of plasma energy and the current decay induce currents in the vacuum vessel which can produce very large forces on the vessel.

These events can be found in detailed in (Wesson and Campbell, 1987).

Despite of the lack of theoretically knowledge on disruptions, there are several studies and explanations for disruption causes. The operational limits for steady operation are imposed by low- q and density limit disruptions. These limits are combine in a Hugill diagram, see figure 1.16, to give a pattern to the experimental behaviour. This diagram shows the boundary of operation limited by disruptions plotting $1/q$ against $n_e R/B$, where the last parameter is introduced in (Murakami et al., 1976), q is the safety factor, n_e is plasma electron density, R is the major radius and B is the toroidal magnetic field.

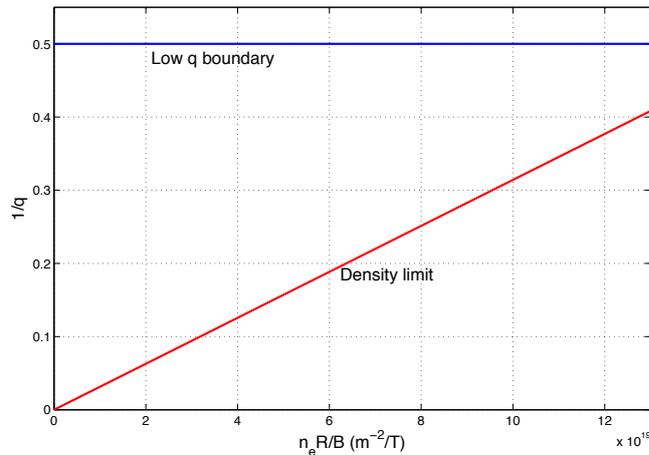


Figure 1.16: Hugill diagram showing operating regime limited by disruptions

There are differences in plasma parameters and evolution of the disruption phases, which allow establish a classification (Savtchikov, 2004; Murari et al., 2008a):

- **Low- q :** the safety factor q is the ratio of the times a particular magnetic field line travels around a toroidal confinement from toroidal direction to poloidal direction. In order to avoid instabilities, the safety factor should be $q > 1$. It has been demonstrated that low- q values and

particularly $q = 2$ are values which play a fundamental role in plasma stability. Magnetic surfaces with $q = 2$ usually became unstable and lead into disruption (Morris, 1992). The cause of these disruptions is probably the incompatibility of stability between $m = 1$ and $m = 2$ modes as q is lowered.

- **Density limit:** this instability begins when density exceeds an upper threshold or due to impurities in plasma. In a tokamak there are low and high density limit (Kleva and Drake, 1991). The purpose of fusion device is working at the highest density without leading into disruption, in order to maximize the energy output. This is because fusion reaction rate scales with the square of density. This upper threshold has not a fixed value, it depends on plasma configuration. Usually increasing density causes an increase in the radiated power, if the radiated power exceeds the heating power, plasma current and temperature contract while plasma detaches from the wall. This scenario usually leads into disruption.
- **Mode Lock:** due to temperature gradients, external perturbation of the magnetic field, pressure gradients or unstable currents, MHD instabilities can appear creating magnetic islands. The instability can reach a situation in which, it stops to rotate and “locks to the wall”, then it can grow faster enough to occlude a sector and modify the plasma current. All of this can induce a disruption.
- **High radiated power:** similar to density limit disruptions but the increase in the radiated power, in the presence of impurities, happens before the one in plasma density. MARFEs are particularly dangerous radiating instabilities which can lead into disruption.
- **H/L transition:** if a H/L transition happens at high density, usually after a reduction of the additional heating, it can induce instabilities and ends in a disruption, which it usually is a density limit disruption.
- **Internal transport barrier:** this type usually appears in plasmas configurations called “advanced scenarios”, due to the presence of internal transport barriers (ITB). The ITBs reduce the transport of particles and energy from the inner to the outer side of the plasma. The pressure gradient required by a ITB can generate instabilities finishing in a disruption. This type is very difficult to predict because of the short time scale of the precursor phase. Moreover, the time between the precursor phase and the current and thermal quench is very short, a few tens of milliseconds.
- **Vertical displacement:** they usually happen in plasma configurations with elongated shapes. Unless these configurations have better

performance than circular shapes, they are vertically unstable, hence the presence of a vertical position and velocity controller are necessary. If the controller fails, the plasma moves vertically towards the vacuum vessel disrupting. These disruption are very fast and hard to predict, albeit they rarely happen unless control system fails.

- β *limit*: this type is related to plasma pressure which can produce an MHD instability. β parameter is a MHD parameter which measure the confinement pressure, therefore it is directly linked with stability limits. These limits are not common in JET so this type of disruption does not happen frequently.

From an operational point of view, in (de Vries et al., 2011), the MHD instabilities and several events which can lead into a disruption are analysed. A statistic of the technical problems which cause disruptions and a classification of the different types are made in this work.

1.2.2 Disruption predictors

The prediction of the incoming disruption with enough time to carry out mitigation actions plays an increasing important role in the current and future devices. The closest approach to ITER is JET with the new ITER-like wall (ILW), unless the operational range is not the same. Although the development of more robust operational scenarios has reduced the JET disruption rate over the last decade from about 15 – 10% to below 4% (de Vries et al., 2009); disruptions probably will not be completely avoidable. The principal cause of disruptions comes from human errors on the operation, hence, even with better understanding of physics and being more careful, ITER aims to operate with a disruption rate of 1% or less. This rate is bigger enough to cause large damage on the device. To characterize plasma physics of disruption for their prediction is extremely hard. The complexity of developing a physic driven system is due to the lack of theoretically knowledge on disruption phenomena, the large number of parameters involved in this stability and the non linear relation between them. Therefore, several data driven systems have been developed in the last 15 years, mainly based on neural networks and support vector machines (SVM), and they have been employed as an alternative approximation to detect the phenomena. Machine learning techniques are highly adaptable to disruptions because the instability can be considered a classification problem: disruptive and non disruptive. Not every study approach disruptions making an adequate feature extraction or treatment of the problem. Despite of this, several studies have been carried out:

- Many studies by Cannas. In (Cannas et al., 2004a), it is shown that the classifier is able to classify between disruptive or non disruptive

discharges through a neural network training. Self Organising Maps (SOM) are used in (Cannas et al., 2004b), where a database formed by 274 discharges is divided in groups of 86 disruptive discharges to determine the training samples. This predictor based in neural networks reaches a success rate of 68%. Unless high success rates are obtained with the training set, in (Cannas et al., 2007) the rates notably decrease with the test set. Moreover the databases used in these studies are formed by a few hundreds of discharges. Other works as (Cannas et al., 2013) attempt determine the type of disruption, using the non-linear generative topographic map manifold learning technique.

- A neural network was developed to predict the occurrence of disruptions caused by edge cooling mechanisms in ASDEX Upgrade, obtaining good results in (Pautasso et al., 2001), around 85% of success rate in a database of 500 shots with 65 disruptive discharges. Another study in ASDEX Upgrade (Morabito et al., 2001), based on a fuzzy-neural network approach, uses a training database formed by 62 discharges and a test database of 46 discharges, obtaining 95% of success rate. In the same device it should be pointed out the study (Pautasso et al., 2002), following the work in (Pautasso et al., 2001), the system is tested in a simulated real time network obtaining 1% of false alarm rate and 79% of success rate. Other attempt to develop a real time disruption predictor in ASDEX Upgrade is described in (Cannas et al., 2010), where an alternative to the protection system in ASDEX Upgrade is developed, without making use of the LM signal. The results shows good rates on the training set but predictive performances deteriorate when the predictor is tested on pulses belonging to successive campaigns.
- In DII-D neural networks were used to train a predictor for β -limit disruptions (Wroblewski et al., 1997). This study attempt to stablish thresholds to trigger alarms. Using 33 input magnetic features, the predictor was trained with 56 disruptive discharges and it reached 90% of success rate in a test set with 28 disruptive discharges. The false alarm was high, about 20%.
- The study (Sengupta and Ranjan, 2001) in ADITYA tokamak was intended to detect density limit disruptions. A database with only disruptive discharges formed by 23 shots was used. This work was not suitable to be implemented in a real time network.
- Other works with neural networks were developed in JT-60U by Yoshino. In (Yoshino, 2003), the system is trained with 12 disruptive discharges and 6 non-disruptive discharges. The output of the neural network trained is validated through an analysis with a set formed by 12 dis-

ruptive discharges. The final predictor was tested on 300 disruptive discharges and 1008 non-disruptive discharges, reaching over 85% of success rate 50 ms before the disruption time. The work (Yoshino, 2005) uses a database of 525 discharges and shows that a system trained with only non-disruptive discharges and fitted with the previous work, can obtained good results on β -limit disruption prediction, above 76%.

- An interesting study links the devices JET and ASDEX Upgrade (Windsor et al., 2005). The database on this study is formed by 185 discharges, where 89 come from ASDEX Upgrade and 96 from JET. A neural network predictor is trained with data from JET and tested on data from ASDEX Upgrade, obtaining a success rate of 67%. The inverse procedure reaches 69% of success rate.
- In JET, Murari develops a technique with fuzzy logic (Murari et al., 2008a). This classifier based on rules obtains low success rates, however it shows an original analysis, using CART, of the signals significance for disruption prediction.

The majority of the previous works are used on small databases formed by a few hundreds discharges in the best cases. It is not possible to extrapolate to future devices due to the lacking of robustness of the predictors. Disruption prediction depends on the mitigation systems, which are based on the fast end of plasma heating, control of the plasma current and shape, gas and pellet injection. The purpose of these systems is to mitigate the effect of disruptions. In general, fast disruptions are been excluded from the previous works and this can be a serious problem cause these disruptions are uncommon but as dangerous as the others.

Alternative predictors have been developed with higher databases and more robust. The Advanced Predictor of Disruptions (APODIS)(Rattá et al., 2010), explained in section 3.1, obtains high success rates. This predictor was implemented in the real time network at JET (López et al., 2012), obtaining successful results on the first ILW campaigns (Vega et al., 2013a). In the next chapters of this thesis, the main contributions to the current disruption prediction knowledge on JET and future devices as ITER will be exposed.

Chapter 2

Data mining techniques

Large volumes of raw data are collected from nuclear fusion devices. During an experimental campaign, plasma properties are measured by several diagnostics which store the data in time series, (t, x) . The purpose of collecting all these data is to study the plasma physics in nuclear fusion devices, however raw data is not usually analysed before being processed. Data mining is the field which includes the statistical and artificial intelligence techniques that aim to obtain knowledge from data. These techniques analyse data from different perspectives, extract the useful information, seek hidden patterns and create models. So that, these techniques have become an important tool in several fields: science, statistics, economy, medicine, business, games, social studies, politics, etc.

This chapter covers all the techniques and theory which have been applied to nuclear fusion databases in the following chapters. Therefore this chapter is divided in some sections: section 2.1 describes all the techniques and methods that have been used to analyse and process the nuclear fusion data. Then, wavelet analysis is briefly exposed in section 2.2. Classification methods are explained in section 2.3, which is divided in the Support Vector Machines (SVM) tool in section 2.3.1, and Conformal Predictors in section 2.3.2. In section 2.4, Support Vector Regression (SVR) are shown followed by the UMEL tool. Finally, martingales used for change detection in data stream are explained in section 2.5.

2.1 Data signal processing

To extract knowledge from data requires applying data mining techniques, but in general the data are incomplete, it means lacking attribute values or certain attributes of interest. Unfortunately, data are also noisy, containing errors or outlier values that deviate from the expected, and inconsistent, containing discrepancies. Dealing with large databases and several attributes requires data preprocessing. Noisy and dirty data can introduce

false information into the mining procedure, resulting in unreliable output. Therefore it is necessary cleaning the data, it means filling in missing values, smoothing noisy data, identifying and removing outliers, etc. Other important step is data integration and transformation, sometimes it is necessary to deal with attributes and information from different sources, and attributes cannot be used if the measurement units or representation is different. Sometimes this problem is overcome normalising the different variables in order to scale the data, this allows comparing different variables or attributes. Other times, missing values need to be estimated by interpolation or extrapolation. Data transformation also includes more complex techniques as PCA (Principal Component Analysis), to study the importance of different variables, or Fourier and wavelet transforms, which decompose the input data and provide information about the time and frequency domains. These techniques allow to extract relevant information from data.

2.1.1 Data normalisation

Data normalisation consists of adjusting values measured on different scales to a common scale, which allows comparing attributes that originally are in a different domain or scale. If the data are not normalised, the results can be seriously influenced by extreme values due to range of values of raw data varies widely. Generally machine learning studies require previously data normalisation in order to each feature contributes approximately proportionately to the final result.

In this thesis it has been used decimal scaling and normalisation by scaling. *Decimal scaling* allows changing the order of magnitude, increasing or reducing the order in powers of 10. This transformation is particularly useful when some attributes from the data are measured in a high or low range compare to the remaining data or attributes. The transformation consist of:

$$x_i^* = 10^j \cdot x_i \quad (2.1)$$

where j is the power of 10 and $j \in \mathbb{N}$

Normalisation by scaling is a method used to standardize the range of independent variables or features of data. This method attempt to scale the range of the data between a maximum and minimum. The rescaling can be done in the range $[0, 1]$:

$$x_i^* = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (2.2)$$

Or between a specific value for the maximum and minimum, $[x'_{max}, x'_{min}]$:

$$x_i^* = \frac{x_i - x_{min}}{x_{max} - x_{min}}(x'_{max} - x'_{min}) + x'_{min} \quad (2.3)$$

where x is the original value and x_i^* is the normalised value.

Several software libraries on machine learning, as *libsvm*¹ that has been used in this thesis, make use of data normalised in ranges like $[0, 1]$ or $[-1, 1]$. The works on disruption prediction, like (Dormido-Canto et al., 2013; Moreno et al., 2014b; Vega et al., 2013a, 2014b, 2013b,c, 2014a), use this normalisation between $[0, 1]$ and previously extreme values in some features are changed by selected maximum and minimum values for this features. This is carried out in order to eliminate outliers from measurement or acquisition errors. Then, this maximum and minimum values are used to normalise in the range $[0, 1]$.

An important and extended normalisation is *standardisation*, which centres the values of each feature in the data around a mean value or zero and unit variance, in order to gather all the values. This normalisation is useful for methods which work with distances. However if the problem requires relations between different features measured in different units, these values will need to be scaled. Focusing on the data distribution, two normalisations can be employed:

- Scaling to unit deviation (the general method of standardisation, $(\bar{x}, \sigma) = (0, 1)$): determining the distribution mean and standard deviation for each feature. Next we subtract the mean from each feature. Then we divide the values (mean is already subtracted) of each feature by its standard deviation.

$$x_i^* = \frac{x_i - \bar{x}}{\sigma}, \text{ where } \sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \quad (2.4)$$

- Scaling to unit length, $(\bar{x}, \sigma) = (0, 1)$: another option, that is widely used, is to scale the data of a feature such that the complete vector has length one. This usually means dividing each component by the Euclidean length of the vector.

$$x_i^* = \frac{x_i - \bar{x}}{L}, \text{ where } L = \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (2.5)$$

Data standardisation is really useful in regression problems, in order to extract scale laws. Several studies uses this normalisation like (González et al., 2012; Vega et al., 2012).

2.1.2 Data interpolation

Sometimes some features present missed or unknown values, and this could be a problem in the analysis. For example, in time series cannot be correctly

¹<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

analysed if there are missed values at certain times. To sort out this problem, there are different methods to estimate missing or unknown values. Interpolation is probably the most widely method used to overcome this task. It is defined as the process which allows calculating unknown values from the known data. This operation is useful when signals are acquired with different sample rate. Therefore, interpolating the number of samples are changed but the sample frequency is equalised for every signal. It is important to note that interpolation is different from “smoothing”, using interpolation the estimated line passes through all the known data points. In this section the most common interpolation methods are explained, furthermore they are easy to compute:

- **Linear:** it is the simplest interpolation method. Given two known data points (x_a, y_a) and (x_b, y_b) , the points are joined by a straight line segment. The middle of this segment is chosen as the new point, as it can be seen in figure 2.1. Therefore, the value y at the point x , between x_a and x_b , is determined by:

$$\frac{y - y_a}{y_b - y_a} = \frac{x - x_a}{x_b - x_a} \rightarrow y = y_a + (y_b - y_a) \frac{x - x_a}{x_b - x_a} \quad (2.6)$$

This formula can be obtained geometrically from figure 2.1.

- **Polynomial:** it is the interpolation of a given data set by a polynomial. The idea is find a polynomial which goes exactly through some given points. Let’s assume a polynomial of degree n which defines the interpolation and the unknown value y at x :

$$p(x) = c_n x^n + \dots + c_1 x + c_0 \quad (2.7)$$

Therefore polynomial p interpolates the data points (x_i, y_i) :

$$p(x_i) = y_i \quad (2.8)$$

- **Spline:** while linear interpolation uses a linear function for each of intervals (x_a, x_b) ; spline interpolation uses low-degree polynomials in each of the intervals, and chooses the polynomial pieces such that they fit smoothly together. The resulting function is called a spline.

In figure 2.2 some interpolation methods are shown over the same raw data. It can be seen that linear interpolation, although it is a fast operation, results in discontinuities at each point. Often it is necessary a smoother interpolation, therefore, in these cases it would be better to use a different method. For example cubic or spline provide a smoother transition between adjacent segments. However polynomial interpolation is computationally expensive compared to linear interpolation. In the case of spline interpolation,

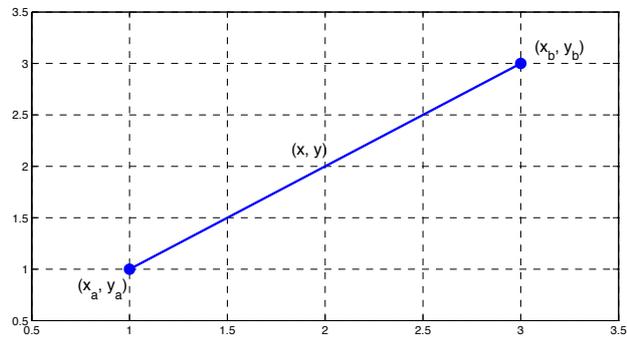


Figure 2.1: Linear interpolation

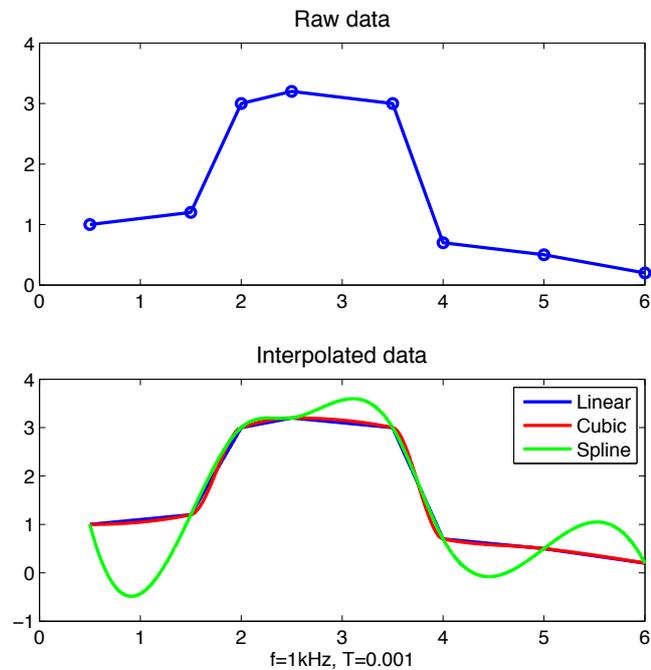


Figure 2.2: Example of interpolation methods

like polynomial, it incurs a smaller error than linear and provides a smoother result; and it is easier to evaluate than the high degree polynomials. The method is chosen depending on the problem to be solved.

The interpolation range is chosen by the maximum value of the set formed by the minimum temporal values of each signal, as first time; and the minimum temporal value of the set compound of the maximum temporal val-

ues of each signal, as the last time. This is extremely important in fusion databases; during a discharge many diagnoses acquired signals with different sample rates. Therefore, this way every signal will have the same sample rate and there will not be unknown values. This is necessary before carrying out any study with these data. This process has been applied to generate the databases explained in section 3.2, which have been used in many studies as (Vega et al., 2014b; Pereira et al., 2015; Murari et al., 2013; Dormido-Canto et al., 2013; Moreno et al., 2014b,a, 2015; Talebzadeh et al., 2015; Rattá et al., 2015).

Sometimes these methods cannot be used due to real time applicability. Sometimes in a real time network, we cannot interpolate between two known points because resolution time can be lost. Therefore, due to we do not have the two samples (x_a, y_a) and (x_b, y_b) to calculate the sample in between, it is applied the *interpolation by the nearest neighbour on the left*. The last sample provided at x is x_a, y_a , so the last known value is assigned to y , it means (x, y_a) . This has been used in works as (López et al., 2012; Moreno et al., 2014b)

2.1.3 Fourier transform

In 1807, Jean-Baptiste Joseph Fourier began the research in Fourier series by demonstrating that time series can be expressed in terms of trigonometric series. Obviously, trigonometric functions are periodic and consequently the time series development. By definition, a function $f(x)$ has a period P if $\forall x, f(x + P) = f(x)$, where $P \in \mathbb{N}$. The minimum P value is the period of $f(x)$.

Given an orthonormal basis of functions $\phi(x)$, the function $f(x)$ can be written:

$$f(x) = c_1\phi_1(x) + \dots + c_n\phi_n(x) + \dots, \quad (-c < x < c), n = 1, 2, \dots \quad (2.9)$$

The $f(x)$ series is expected to converge in the function $f(x)$. Then the coefficients c_n can be expressed:

$$c_n = \int_{-c}^c f(x)\phi_n^*(x)dx, \quad (2.10)$$

where ϕ^* is the complex conjugate of ϕ . These coefficients 2.10 with the equation 2.9 define the generic Fourier Series for $f(x)$:

$$f(x) = \sum_{n=1}^{\infty} c_n\phi_n(x). \quad (2.11)$$

If $f(x)$ has period 2π , the Fourier series of $f(x)$ can be written:

$$f(x) = \frac{a_0}{2} \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)), \quad n = 1, 2, \dots \quad (2.12)$$

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(nx) dx \quad (2.13)$$

$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx) dx \quad (2.14)$$

Using Euler's identity $e^{inx} = \cos(nx) + i\sin(nx)$ and the orthonormal basis of functions $e_k(x) = \frac{e^{ikx}}{\sqrt{2\pi}}$, Fourier series of $f(x)$ yields:

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}, \quad (2.15)$$

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx. \quad (2.16)$$

It can be seen that Fourier series 2.11 is not able to represent all values of x , if the function is not periodic. Therefore, if $f(x)$ is not periodic, in order to obtain a representation for all x values the previous representation is extended to $c \rightarrow \infty$. The result is the Fourier transform:

$$F(n) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-inx} dx. \quad (2.17)$$

where $n \in \mathbb{R}$. The function can be reconstructed from its Fourier terms using the inverse Fourier transform:

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(n) e^{inx} dn. \quad (2.18)$$

Fourier transform in $\mathbb{L}^2[-\infty, \infty]$ preserves the norm 2.19 and the inner product 2.20.

$$\int_{-\infty}^{\infty} |f(x)|^2 dx = \int_{-\infty}^{\infty} |F(n)|^2 dn. \quad (2.19)$$

$$\int_{-\infty}^{\infty} f(x) g^*(x) dx = \int_{-\infty}^{\infty} F(n) G^*(n) dn. \quad (2.20)$$

Further details are provided in (Brown and Churchill, 2001). Fourier transform is widely used in signal processing and analysis. If the signals are periodic and regular enough, successful results are obtained. On the other hand, it does not work properly with non stationary signals. Therefore, if the function $f(x)$ is non stationary, the Fourier basis functions are infinite along time but local in frequency. Fourier transform identifies a local frequency but it is not able to provide information about the whole spectrum along the temporal evolution. For instance, it cannot be correctly analysed by Fourier

transform the beginning and end of a discrete signal, or the time where a singularity happens.

To sort out this problem, the *short-time Fourier transform* is used, referred to as the Gabor transform, the Wigner distribution and the ambiguity function. The non stationary signal is divided in temporal segments where the signal can be considered quasi-stationary, and Fourier transform is applied to each local segment. In 1940, Gabor introduced the short-time Fourier transform:

$$S_f(w, \tau) = \int f(t)g^*(t - \tau)e^{-iwt} dt, \quad (2.21)$$

where $g(t)$ is a short-time window with a fixed width, which slides along the time axis with a factor τ (Poularikas, 2010, c. 10.1.3). The Gaussian function was proposed as the function $g(t)$, and it was demonstrated that Fourier transform of a short-time window remains Gaussian. The function is defined as (see figure 2.3):

$$g(t) = \frac{1}{s} e^{-\pi t^2/s^2}. \quad (2.22)$$

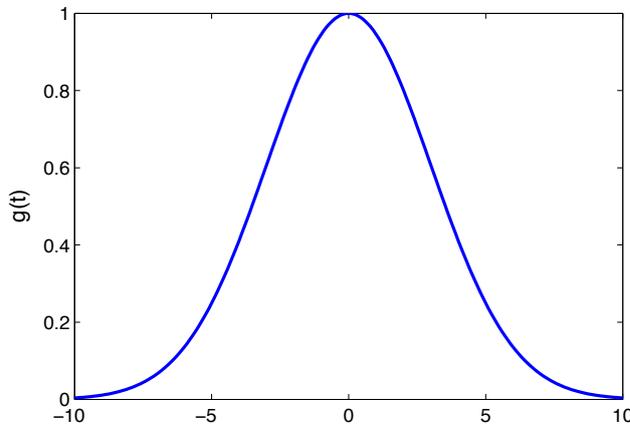


Figure 2.3: Gaussian function for short-time Fourier transform

Fourier transform applied to discrete signals is called discrete Fourier transform (DFT). In most typical applications the STFT, previously explained, is performed on a computer using the Fast Fourier Transform (FFT). The FFT can be applied when the total number of samples is a power of two. For example, a process using FFT takes approximately $O(n \log(n))$ operations, while DFT takes about $O(n^2)$, therefore FFT is significantly faster. The output from FFT provides negative frequencies, which come from the

FT properties. The first half of FFT output contains frequencies sorted in ascending order, where the first frequency corresponds to $0Hz$. The second half of the output provides the negative frequencies. In (Vega et al., 2014b; Pereira et al., 2015; Murari et al., 2013; Dormido-Canto et al., 2013; Moreno et al., 2014b, 2015; Talebzadeh et al., 2015; Rattá et al., 2015), it has been used the FFT considering only the positive frequencies and deleting the continuous component. The FFT has been used as it is explained in section 3.2.

2.2 Wavelet analysis

Wavelet analysis is an useful technique which allows analysing different types of data as waveforms and images. Wavelet transform decomposes the input data into time-frequency components with many applications: de-noising audio data, signal compression, object detection, speech recognition, etc. Wavelets are an extension of Fourier analysis (Burke, 1996, p. 5). In Fourier analysis information about the time domain of the data are not provided, however wavelets have good location in time and frequency domains. A wavelet can be defined as a function $\varphi \in \mathbf{L}^2(\mathbb{R})$ with a zero average:

$$\int_{-\infty}^{\infty} \varphi(t) dt = 0. \quad (2.23)$$

This function is normalized $\|\varphi\| = 1$, and centered around $t = 0$. After that, function φ is scaled by s and translated by u (remaining normalised $\|\varphi_{u,s}\| = 1$):

$$\varphi_{u,s}(t) = \frac{1}{\sqrt{s}} \varphi\left(\frac{t-u}{s}\right). \quad (2.24)$$

Given a function $f \in \mathbf{L}^2(\mathbb{R})$ at time u and scale s , the wavelet transform would be:

$$Wf(u, s) = \langle f, \varphi_{u,s} \rangle = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{s}} \varphi^*\left(\frac{t-u}{s}\right) dt. \quad (2.25)$$

Wavelets $\varphi_{u,s}(t)$ generated from the same function $\varphi(t)$, present the same form but different scale s and translation u . Scales $s > 0$ are always used. Wavelets are dilated for $s > 1$, and contracted for $s < 1$. Therefore, different ranges of frequency are covered by changing the value of s . Lower frequencies correspond to small values of s (Daubechies, 1992).

This wavelet transform 2.25 can be expressed as the convolution product:

$$Wf(u, s) = f * \bar{\varphi}_s(u), \text{ with } \bar{\varphi}_s(t) = \frac{1}{\sqrt{s}} \varphi^* \frac{-t}{s}. \quad (2.26)$$

And the Fourier transform of $\varphi_s(t)$ is

$$\hat{\varphi}_s(w) = \sqrt{s} \hat{\varphi}^*(sw). \quad (2.27)$$

Due to $\hat{\varphi}(0) = \int_{-\infty}^{\infty} \varphi(t) dt = 0$, $\hat{\varphi}$ can be defined as the transfer function of a band-pass filter. This information can be found widely explained in (Mallat, 1999, c. 4). At this point, it is necessary to construct a translation-invariant wavelet representation. For this purpose the scale s is made discrete along a dyadic sequence $2^j, j \in \mathbb{Z}$, but not the translation parameter u . The dyadic wavelet transform of $f \in \mathbf{L}^2(\mathbb{R})$ is written as:

$$Wf(u, 2^j) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{2^j}} \varphi\left(\frac{t-u}{2^j}\right) dt = f * \varphi_{2^j}(u), \quad (2.28)$$

with

$$\varphi_{2^j}(t) = \varphi_{2^j}(-t) = \frac{1}{\sqrt{2^j}} \varphi\left(\frac{-t}{2^j}\right). \quad (2.29)$$

A discrete dyadic wavelet transform can be computed with a fast filter bank algorithm if the wavelet is appropriately designed. This theory is widely explained in (Mallat, 1989, 1999). Wavelet analysis allows obtaining wavelet coefficients of a discrete series and a multi-resolution representation of the data. This multi-resolution representation provides a scale-invariant hierarchical representation of data. The information is provided in two successive detail levels 2^j and 2^{j+1} (Mallat, 1989), where each one of these levels provides different information of the data. Levels with a low resolution contain information about the larger patterns in the data (low frequency behaviours), while levels with a high resolution represent the smaller details of the data (high frequency behaviours).

In these thesis it has been used the *Haar* wavelet in discrete form, the *Haar transform* (Walker, 1999, c. 1). This wavelet is the simplest and most robust wavelet function, and it serves as a prototype for all other wavelet transforms. Given a discrete signal $f = (f_1, f_2, \dots, f_n)$, where $n \in \mathbb{N}$, and the values of f are $f_1, f_2, \dots, f_n \in \mathbb{R}$. These values are usually the measurements from an analogue signal at the time values t_1, \dots, t_n . It is assumed that time vector is equally spaced. Haar transform decomposes the discrete signal f in two sub-signals with length $n/2$. One sub-signal contains the *approximation coefficients*, which correspond to an average or trend; and the other the *detail coefficients*, which correspond to a difference or fluctuation. Haar transform can be applied at different levels. The first level corresponds to the first Haar transform H_1 given by:

$$f \xrightarrow{H_1} (a^1 | d^1), \quad (2.30)$$

where a^1 and d^1 correspond to the approximation and detail coefficients respectively. The approximation coefficients $a^1 = (a_1, \dots, a_{n/2})$ are computed by taking consecutive pair of values of f and calculating the average. Then this value is multiplied by $\sqrt{2}$. That is:

$$a_m = \frac{f_{2m-1} + f_{2m}}{\sqrt{2}}, \text{ with } m = 1, 2, \dots, n/2. \quad (2.31)$$

The detail coefficients $d^1 = (d_1, \dots, d_{n/2})$ are computed by taking consecutive pair of values of f and calculating half the difference. Then this value is multiplied by $\sqrt{2}$. That is:

$$d_m = \frac{f_{2m-1} - f_{2m}}{\sqrt{2}}, \text{ with } m = 1, 2, \dots, n/2. \quad (2.32)$$

Let's show an example, given a function $f = (4, 6, 10, 12, 8, 6, 5, 5)$, then approximation coefficients, following 2.31, are $a^1 = (5\sqrt{2}, 11\sqrt{2}, 7\sqrt{2}, 5\sqrt{2})$. The detail coefficients, following equation 2.32, are $d^1 = (-\sqrt{2}, -\sqrt{2}, \sqrt{2}, 0)$. Successive levels can be obtained by applying Haar transform to approximation coefficients at each level, halving the number of samples in every step. This can be seen in figure 2.4.

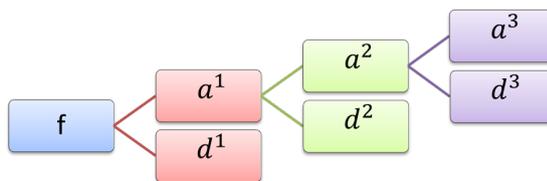


Figure 2.4: Haar wavelet levels

The mapping H_1 in 2.30 has an inverse. From coefficients a^1 and d^1 , it can be obtained the function f by applying:

$$f = \left(\underbrace{\frac{a_1 + d_1}{\sqrt{2}}}_{f_1}, \underbrace{\frac{a_1 - d_1}{\sqrt{2}}}_{f_2}, \dots, \underbrace{\frac{a_{n/2} + d_{n/2}}{\sqrt{2}}}_{f_{n/2}}, \underbrace{\frac{a_{n/2} - d_{n/2}}{\sqrt{2}}}_{f_{n/2}} \right). \quad (2.33)$$

Therefore, the diagram in figure 2.4 can be followed in the opposite direction.

It is important to note why the factor $\sqrt{2}$ is used. This factor allows ensuring that the Haar transform preserves the energy of a signal. The energy of a signal f is understood as the sum of the square of its values, $\epsilon_f = \epsilon_{(a^1|d^1)}$:

$$\epsilon_f = f_1^2 + f_2^2 + \dots + f_n^2. \quad (2.34)$$

Suppose the function from the previous example, $f = (4, 6, 10, 12, 8, 6, 5, 5)$, and its at level 1 coefficients $a^1 = (5\sqrt{2}, 11\sqrt{2}, 7\sqrt{2}, 5\sqrt{2})$ and $d^1 = (-\sqrt{2}, -\sqrt{2}, \sqrt{2}, 0)$.

Then the energy would be:

$$\epsilon_f = \sum_{i=1}^n f_i^2 = \sum_{i=1}^{n/2} (a_i^1)^2 + \sum_{i=1}^{n/2} (d_i^1)^2 = 440 + 6 = 446. \quad (2.35)$$

The first level Haar transform conserves energy. It can be seen the most part of the energy is compacted into the approximation coefficients, it means that only approximation coefficients can represent the morphology of a signal without losing much visual information, and reducing the dimensionality of the data. Let see an example in figure 2.5, given a signal (it corresponds to plasma current of discharge 81852 at JET) formed by 12180 samples, the signal is reduced and correctly plotted with 191 samples. It has been done by applying a Haar transform at level 6. The level depends on the resolution needed to analyse the data, higher resolutions requires lower levels of decomposition.

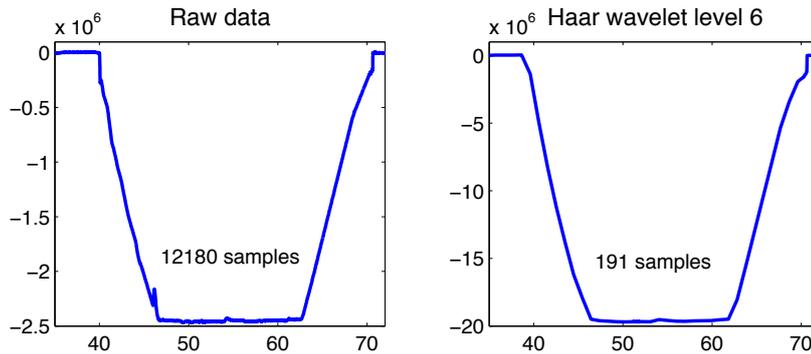


Figure 2.5: Example of Haar transform

In this thesis wavelet analysis is mainly used in sections 4.1.3 and 4.2, to reduce the dimensionality of the data and extract the most possible information with the less amount of data; and to de-noise and smooth signals in order to detect high frequency pikes. However, in section 4.2, it is a problem reducing the dimensionality because it is necessary a high resolution in order to detect an specific behaviour. The wavelet transform previously explained is also called decimated wavelet transform (DWT), where the signal is convolved and decimated. Therefore, a modified version has been used to overcome the resolution problem, the Non-Decimated Wavelet Transform (NDWT) or stationary wavelet transform. NDWT has no sub-sampling step so it keeps the same number of coefficients of each level.

2.3 Classification

Classification problems aim to identify the label y_{n+1} of a new sample x_{n+1} , given a set of training samples with their labels $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$. Any classification method which incorporates information from training samples employs learning. Learning process is the hardest step in a classification problem, where automatic learning algorithms and pattern recognition techniques are employed to extract the information from the training samples and develop a model. Learning comes in different forms (Duda et al., 2000, p. 16-17):

- *Supervised learning*: the training samples and their respective labels, (x_n, y_n) , are previously known. The goal is to find patterns and build mathematical models based on the training samples. These models are validated with test samples and their predictive capacity is evaluated. In this thesis this learning is broadly used by Support Vector Machines, see section 2.3.1.
- *Unsupervised learning*: the training samples x_1, \dots, x_n are known but their respective labels y_1, \dots, y_n are unknown. So that, the task is to develop classification labels automatically. Unsupervised learning seeks out the similarities between the data samples and attempts to create clusters or groups of similar samples. There are several clustering techniques. The application of martingales (see section 2.5) in section 4.1.2 could be understood as an unsupervised learning approach.
- *Reinforcement learning*: given an initial set of samples x_1, \dots, x_n (with their respective labels y_1, \dots, y_n in the case of supervised learning, or without the previous knowledge of the labels in the case of unsupervised learning) the algorithm evaluates the new incoming sample x_{n+1} and it is included in the initial set to retrain or reinforce the model for the next test sample x_{n+2} . The purpose is to improve the model and the predictive capacity with every new incoming sample. There are several techniques and applications, however in these thesis it could be consider as reinforcement learning Venn predictors (section 5.2.1), and prediction from scratch (section 5.2.2).

In addition, looking at how is the process between the training set and the prediction it is possible to distinguish two classification paradigms (Gammerman and Vovk, 2007):

- **Induction**: induction prediction firstly infers a general rule, model or predictor from the initial training samples, this is called inductive step. Then, in the deductive step, the model or general rule predicts the label of the incoming sample. The bulk of the computations happens

during the inductive step when the model is calculated. Examples of this paradigm are the Support Vector Machines (see section 2.3.1) and Inductive Conformal Predictors (see section 2.3.2.1).

- **Transduction:** the prediction is obtained in a single step. It is taken a shortcut, the training samples x_1, \dots, x_n and the incoming sample x_{n+1} are used to predict the label y_{n+1} . This process must be done for each new incoming sample, it usually takes a longer time than deductive step. Therefore this process becomes unfeasible for large training sets. An example of transductive prediction is Conformal Predictors, which is explained in section 2.3.2.

Vapnik's (Vapnik, 1998, 2000) distinction between induction and transduction, as applied to the problem of prediction, is illustrated in figure 2.6.

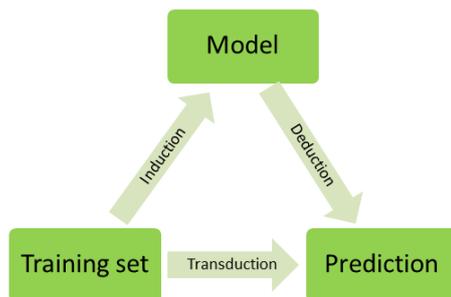


Figure 2.6: Inductive and transductive prediction diagram

2.3.1 Support Vector Machines

The Support Vector Machines (SVM) is a set of statistical supervised learning algorithms which can be used to obtain either classification or regression. SVM theory was developed by Vladimir N. Vapnik between 1963 and 1995. The first algorithm was created in 1963 by Vladimir N. Vapnik and Alexey Ya. Chervonenkis. In 1992, B.E. Boser, I.M. Guyon and V.N. Vapnik presented a training algorithm that maximized the margin hyperplanes by kernel methods (Boser et al., 1992). The current version was presented in (Cortes and Vapnik, 1995), where the underlying idea is that input vectors are non-linearly mapped into a higher dimensionally space. The original space of the input vectors, the training samples, is called *input space*; while the higher dimensionally space is called *feature space*. In the feature space the data are linearly separable by a hyperplane, called *optimal separating hyperplane*. This hyperplane is understood as a generalisation of a plane in a high dimensionally space which divides the feature space into two regions. Unlike conventional statistical and neural network methods, SVM

does not attempt to deal with a small number of features. The mapping into a higher dimension allows SVM to control the complexity of large number of features. The reason is that linear decision boundaries constructed in the feature space, with a large number of dimensions, correspond to non-linear decision boundaries in the input space. SVM overcomes two main problems:

- *Conceptual statistical problem*: how to deal with the complexity of high dimensional spaces and provide good generalization ability. This is solved by using penalized linear estimators with a large number of basis functions or by penalising the norm of the weighting coefficients. The results are approached in a controlled quadratic optimisation problem formulation of the learning problem.
- *Computational problem*: how to perform the numerical optimisation, it means, how to solve the quadratic optimisation problem in a high dimensional space. This problem is overcome by using the dual kernel representation of linear functions.

SVM combines four different concepts, that are explained in greater depth on (Cherkassky and Mulier, 2007, p. 355-356):

1. New implementation of the SRM (Structural Risk Minimization) inductive principle.
2. Input samples mapped onto a very high dimensional space using a set of non-linear basis function defined a priori.
3. Linear functions with constraints on complexity used to approximate or discriminate the input samples in the high dimensional space.
4. Duality theory of optimisation used to make estimation of model parameters in a high dimensional feature space computationally tractable.

The SVM theory can be explained in three steps: optimal separating hyperplane (section 2.3.1.1), mapping the input space into a higher dimensionally feature space (section 2.3.1.2), constructing SVM (section 2.3.1.3).

2.3.1.1 The optimal separating hyperplane

In the high dimensional feature space the separating hyperplane is understood as a linear function that is able to separate the training data without error. In (Vapnik and Kotz, 2006, p. 430) it is mentioned that the algorithm for constructing an optimal separating hyperplane was firstly formulated in 1964 by Chervonenkis and Vapnik, and the detailed theory is contained in the book (Vapnik and Chervonenkis, 1974). Given a training data set of

n separable samples $(x_1, y_1), \dots, (x_n, y_n)$ with $x \in \mathfrak{R}$ and $y \in +1, -1$; the separating hyperplane can be defined as:

$$(w \cdot x) + w_0 \quad (2.36)$$

This hyperplane separates the data training samples and allows us to solve the classification problem directly. The hyperplane satisfies the following constraints, where a sample x_i has the same label as the sign of the distance from the separating hyperplane:

$$(w \cdot x_i) + w_0 \geq 1, \text{ if } y_i = +1 \quad (2.37)$$

$$(w \cdot x_i) + w_0 \leq -1, \text{ if } y_i = -1, i = 1, \dots, n \quad (2.38)$$

These equations can be expressed:

$$y_i[(w \cdot x_i + w_0)] \geq 1, i = 1, \dots, n \quad (2.39)$$

The *margin*, δ , is defined as the minimal distance from the separating hyperplane to the closest data point; and it is called *optimal* when the margin is maximum and the generalization capability is better, see figure 2.7. The distance between the hyperplane and a sample x_i is $|(w \cdot x_i) + w_0| / \|w\|$, so that every sample obeys:

$$\frac{y_k[(w \cdot x_k) + w_0]}{\|w\|} \geq \delta, k = 1, \dots, n \quad (2.40)$$

where $y_k \in -1, 1$. To find the optimal hyperplane is necessary to maximize the margin δ , and this is equivalent to minimizing the norm of w . Therefore an optimal hyperplane satisfies 2.39 and minimises:

$$\eta(w) = \|w\|^2 \quad (2.41)$$

In figure 2.7 it can be seen a separating hyperplane and the optimal separating hyperplane which satisfies 2.39 and minimises 2.41. The samples in red and blue are called *support vectors*, they define the decision surface. The number of support vectors is directly related to the generalisation capability, and they provide a bound on the expectation of the error rate for a sample (Vapnik, 2000, Chapter 6.8):

$$E_n[\text{Error rate}] \leq \frac{E_n[\text{Number of support vectors}]}{n} \quad (2.42)$$

where E_n is the expectation over the training set of n samples.

The determination of the optimal separating hyperplane is a quadratic optimisation problem with linear constraints, so it is not a practical problem to be solved for high dimensional spaces. An optimisation problem can be represented by a dual form if the cost and constraint functions are strictly

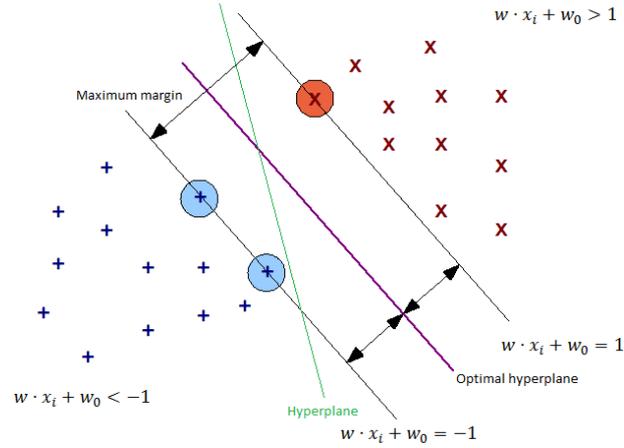


Figure 2.7: Example of optimal and non-optimal separating hyperplane

convex, so that solving the dual problem is similar to solve the original problem. In this case, 2.39 and 2.41 satisfies these criteria and it is possible to represent the problem in its dual form using Kuhn-Tucker theorem (Strang, 1986). Two steps have to be carried out to obtain the dual form from 2.39 and 2.41. Firstly the problem is constructed using Lagrange multipliers:

$$L(w, w_0, \alpha) = \frac{1}{2}(w \cdot w) - \sum_{i=1}^n \alpha_i y_i [(w \cdot x_i) + w_0] - 1 \quad (2.43)$$

where α_i are Lagrange multipliers. This equation has to be minimised with respect to w , w_0 and maximised with respect to $\alpha_i \geq 0$. After that, in the second step the parameters w , w_0 in 2.43 are expressed in terms of α_i using Kuhn-Tucker conditions. Therefore the solutions w^* , w_0^* and α^* should satisfy:

$$\frac{\partial L(w^*, w_0^*, \alpha^*)}{\partial w_0} = 0 \quad (2.44)$$

$$\frac{\partial L(w^*, w_0^*, \alpha^*)}{\partial w} = 0 \quad (2.45)$$

Solving 2.44 and 2.45, the following properties of the optimal hyperplane are obtained:

1. The lagrange multipliers α_i^* should satisfy the constraints:

$$\sum_{i=1}^n \alpha_i^* y_i = 0, \quad \alpha_i^* \geq 0 \quad (2.46)$$

2. Vector w^* is a linear combination of the vectors in the training set:

$$w^* = \sum_{i=1}^n \alpha_i^* y_i x_i, \quad \alpha_i^* \geq 0 \quad (2.47)$$

To obtain the equation of the optimal separating hyperplane, the dual problem is constructed using 2.46 and 2.47 in 2.43:

$$L(\alpha) = -\frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) + \sum_{i=1}^n \alpha_i \quad (2.48)$$

This is the dual form of the optimisation problem which should be maximised with respect to the parameters α_i . Using 2.47 in 2.36, the optimal separating hyperplane is obtained:

$$\sum_{i=1}^n \alpha_i^* y_i (x \cdot x_i) + w_0^* \quad (2.49)$$

The parameter w_0^* can be computed considering the conditions on support vectors. Given a support vector (x_s, y_s) , it satisfies:

$$y_s [(w^* \cdot x_s) + w_0^*] = 1 \quad (2.50)$$

Then, using 2.47 in 2.50, it yields:

$$w_0^* = y_s - \sum_{i=1}^n \alpha_i^* y_i (x_i \cdot x_s) \quad (2.51)$$

In a non-separable case (Vapnik, 2000, p. 136) (Cherkassky and Mulier, 2007, p. 364), it means data that cannot be separated without error. Previously it has been described that a sample is non-separable if it does not satisfy 2.39. It is important to note that misclassification differs from non-separable, the first one refers to the samples which fall within the margin in the wrong side, while the second one corresponds to samples which fall within the margin in the right side. The purpose is to minimise the classification error and maximise the margin of the separating hyperplane, therefore the slack variables $\xi_i \geq 0$ and the function 2.52 are introduced to quantify the non-separable data.

$$F(\xi) = \sum_{i=1}^n \xi_i^p \quad (2.52)$$

where $p \geq 0$. The function 2.52 under the constraints:

$$\begin{aligned} y_i [(w \cdot x_i) - w_0] &\geq 1 - \xi_i, \quad i = 1, \dots, n \\ (w \cdot w) &\leq \Delta^{-2} \end{aligned} \quad (2.53)$$

The function 2.52 is minimised to $p = 1$, which is computationally simple. The Δ -margin separating hyperplane is determined by the vector:

$$w = \frac{1}{C^* \sum_{i=1}^n \alpha_i y_i x_i} \quad (2.54)$$

where α_i and C^* are the solutions to the convex optimisation problem which maximises the functional:

$$W(\alpha, C^*) = \sum_{i=1}^n \alpha_i - \frac{1}{2C^*} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) - \frac{C^*}{2\Delta^2} \quad (2.55)$$

subject to constraints:

$$\sum_{i=1}^n \alpha_i y_i = 0, \quad C^* \geq 0 \quad (2.56)$$

$$0 \leq \alpha_i \leq 1, \quad i = 1, \dots, n \quad (2.57)$$

In order to simplify the computations it is employed the concept of *soft-margin optimal hyperplane*, firstly introduced in (Cortes and Vapnik, 1995), which is determined by the vector w that minimises the functional:

$$\phi(w, \xi) = \frac{1}{2}(w \cdot w) + C \sum_{i=1}^n \xi_i \quad (2.58)$$

where C is a given value. The solution of this optimisation problem is calculated following an equivalent computation as the one followed in the separable case.

2.3.1.2 Mapping the input space into a higher dimensionally feature space

After defining the optimal separating hyperplane, the next step is mapping the input space into a higher dimensionally feature space, which is described in (Vapnik and Kotz, 2006, p. 432). The task is to find a feature space where data can be linearly separable. In (Boser et al., 1992), it is presented a way to construct the optimal separating hyperplane in a Hilbert space without mapping the input vectors into vectors of Hilbert space. This is carried out using Mercer's theorem (Minh et al., 2006):

Let vectors $x \in X$ be mapped into vectors $z \in Z$ of some Hilbert space.

1. *There exists in X space a symmetric positive definite function $K(x_i, x_j)$ that defines the corresponding inner product in Z space:*

$$(z_i, z_j) = K(x_i, x_j)$$

2. Also, for any symmetric positive definite function $K(x_i, x_j)$ in X space there exists a mapping from X to Z such that this function defines an inner product in Z space.

According to this theorem, the separating hyperplane 2.49 in the feature space has the following form:

$$\sum_{i=1}^n \alpha_i^* y_i K(x, x_i) + w_0^* = 0 \quad (2.59)$$

The function $K(x, x_i)$ is called kernel function and depending on the kernel function chosen one can make specific mappings from input space into feature space. Kernel functions are widely explained in (Schölkopf and Smola, 2002, c. 2).

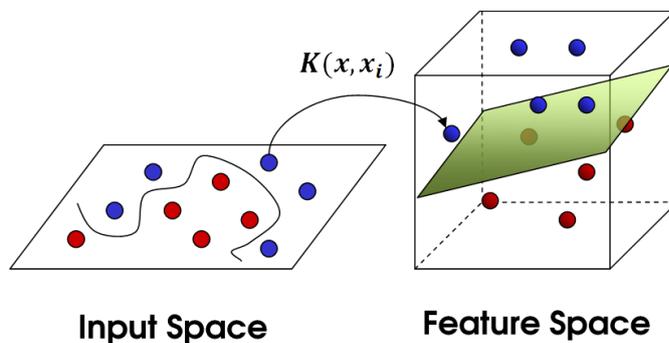


Figure 2.8: Input space is mapped into a feature space using a kernel function

2.3.1.3 Constructing SVM

Finally the SVM theory, as mentioned above, was presented in (Cortes and Vapnik, 1995). Input vectors are non-linearly mapped into a higher dimensionally space called feature space. The SVM construction can be followed in (Vapnik and Kotz, 2006, p. 433). The parameters α_i^* from equation 2.59 are calculated in the same way as the non-separable case:

$$W(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (2.60)$$

subject to constraints:

$$\sum_{i=1}^n \alpha_i y_i = 0 \quad (2.61)$$

$$\alpha_i \geq 0, \quad i = 1, \dots, n \quad (2.62)$$

The optimal separating hyperplane can be expressed in terms of the Support Vectors and the decision function yields:

$$D(x) = \sum^{SV} \alpha_i y_i K(x, x_i) + w_0 \quad (2.63)$$

where the value of this function represents the distance from the sample evaluated to the hyperplane in the feature space. There can be used several kernel functions, however the most used are the following:

- *Linear*: this is the simplest kernel function. It is given by the common inner product.

$$K(x, x_i) = x \cdot x_i$$

- *Polynomial*: polynomial functions of p degree. It is well suited for problems where all data is normalized.

$$K(x, x_i) = [(x \cdot x_i) + 1]^p$$

- *Radial basis function (RBF)*: is by far one of the most versatile Kernels. It is the preferred Kernel when we don't know much about the data we are trying to model.

$$K(x, x_i) = \exp \left\{ -\frac{|x - x_i|^2}{2\sigma^2} \right\}$$

The given parameter σ determine the influence area of SVs. If the value is large, the exponential will behave almost linearly and the hyperplane will be smoother. On the other hand if the value is low, the hyperplane will be more sensitive to noise in training data and the fitting is higher.

2.3.1.4 Examples of SVM classification

Parkinson's disease classification

This first example shows the classification of the Parkinson's disease data set shown in appendix B. The data from table B.1 are classified using *libsvm*² in Matlab. The data set has been divided in a training set of 20 samples, it means 10 samples of each class, and 175 samples for testing the model. There have been developed 4 classifiers using different kernels: two RBF kernels with different values of γ parameter, one polynomial and one linear kernel. The results are shown in table 2.1, where it can be seen that the best model

²<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

is obtained with RBF kernels and the lowest value of γ parameter fits better the data reaching a success rate of 80.57% but the generalisation capability decreases. The success rate is understood as $(TP/N * 100)$ the ratio of true positive classifications to the total number of test samples.

| Kernel | Success rate |
|---------------------------------------|--------------|
| Linear | 68.57% |
| Polynomial ($p = 5$) | 72.57% |
| RBF ($C = 100$ and $\gamma = 1.5$) | 78.29% |
| RBF ($C = 100$ and $\gamma = 0.01$) | 80.57% |

Table 2.1: SVM classification of Parkinson's disease

Breast cancer classification

This example presents the classification of the breast cancer database shown in appendix C. This database is classified using *libsvm*³ in Matlab. Two classifications have been carried out, and following the same process as the previous example, each classification analyse 4 models using RBF, polynomial and linear kernels.

First, the database is divided in 20 samples for training (10 benign and 10 malignant samples) and 663 for the test (434 benign and 229 malignant samples). The results are shown in table 2.2. RBF kernels generate a non linear hyperplane. The lowest value of γ fits better the data and is capable to classify both benign (96.31%) and malignant (81.22%) cases. Theoretically, when parameter γ is increased, the data fitness decreases but the generalisation capability is higher; however in this example it can be seen that a higher value of γ parameter classifies all the malignant cases but it is not able to classify benign cases, 59.68%. Polynomial kernel with dimension $p = 5$ obtains similar results than RBF kernel with $\gamma = 0.01$, however linear kernel reaches the best results with 96.77% of success rate in benign cases, and 94.76% in malignant cases.

The second classification trains with 100 samples for training (50 benign and 50 malignant samples) and 583 for the test (394 benign and 189 malignant samples). The results are shown in table 2.3. If the training set is increased, better results are obtained. The conclusion is similar in the case of RBF kernels, but now the best result is reached by polynomial kernel. Although it is similar than linear and RBF ($\gamma = 0.01$) kernels, the malignant cases are better classified reaching a 100% of success rate.

³<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

| | Benign cases | Malignant cases |
|------------------------------|---------------------|------------------------|
| Train samples | 10 | 10 |
| Test samples | 434 | 229 |
| RBF $\gamma = 0.01, C = 100$ | | |
| Success | 96.31% (418) | 81.22% (186) |
| Error | 3.69% (16) | 18.78% (43) |
| RBF $\gamma = 1.5 C = 100$ | | |
| Success | 59.68% (259) | 100% (229) |
| Error | 40.32% (175) | 0% (0) |
| Polynomial $p = 5$ | | |
| Success | 96.54% (419) | 83.41% (191) |
| Error | 3.46% (15) | 16.6% (38) |
| Linear | | |
| Success | 96.77% (420) | 94.76% (217) |
| Error | 3.23% (14) | 5.24% (12) |

Table 2.2: First SVM classification of breast cancer

| | Benign cases | Malignant cases |
|------------------------------|---------------------|------------------------|
| Train samples | 50 | 50 |
| Test samples | 394 | 189 |
| RBF $\gamma = 0.01, C = 100$ | | |
| Success | 95.43% (376) | 96.30% (182) |
| Error | 4.57% (18) | 3.7% (7) |
| RBF $\gamma = 1.5 C = 100$ | | |
| Success | 64.97% (256) | 100% (189) |
| Error | 35.03% (138) | 0% (0) |
| Polynomial $p = 5$ | | |
| Success | 95.69% (377) | 100% (189) |
| Error | 4.31% (17) | 0% (0) |
| Linear | | |
| Success | 96.70% (381) | 95.77% (181) |
| Error | 3.30% (13) | 4.23% (8) |

Table 2.3: Second SVM classification of breast cancer

2.3.2 Conformal Predictors

A. Gammerman, V. Vovk and V. Vapnik introduced conformal prediction (CP) in (Gammerman, A. and Vovk, V. and Vapnik, V., 1998), assigning values of confidence to predictions made by SVM. Then a complete theory on CP was developed where prediction algorithms (nearest-neighbour, SVM, ridge regression,...) can be transformed into randomness tests and, therefore,

be used for producing hedge predictions. Given a initial data set and an error probability ϵ , the new samples x_{n+1} are evaluated obtaining a prediction \hat{y} , which produces a set of labels y that also contains the label y_{n+1} with probability $1 - \epsilon$ (Shafer and Vovk, 2008).

Firstly, the concept of *nonconformity measure* is defined. It is a measure which represents how different is a new incoming sample from a bag of initial samples. If we consider the samples $z_i = (x_i, y_i)$ where $z_i \in Z$, the union of all bags (subsets) $Z^{(n)}$ is written as $Z^{(*)}$. Therefore, the nonconformity measure can be defined as a measurable mapping

$$A : Z^{(*)} \times Z \rightarrow \bar{R}; \quad (2.64)$$

to each possible set $Z^{(n)}$ and new sample, where A is the numerical value which indicates how different the new sample is from the initial set $Z^{(n)}$ (Vovk et al., 2005, p. 23). Given a nonconformity measure A and a subset $Z^n = z_1, \dots, z_n$, it can be calculated how different is a new sample z_{n+1} from Z^n . In (Vovk et al., 2005, p. 25), it can be seen that this nonconformity value is computed as

$$\alpha_{n+1} = A(Z^n, z_{n+1}). \quad (2.65)$$

However the value of α_{n+1} does not tell how different is the sample z_{n+1} , so that it is necessary to compare α_{n+1} with the nonconformity values $\alpha_1, \dots, \alpha_n$ from the samples of the subset Z^n . This comparison is called *p-value*:

$$\frac{\#i = 1, \dots, n : \alpha_i \geq \alpha_{n+1}}{n + 1} \quad (2.66)$$

This fraction is the p-value of the sample z_{n+1} . If this p-value is small, it means close to $1/(n+1)$, then z_{n+1} is nonconforming (an outlier); while if the p-value is large, it means close to 1, then z_{n+1} is conforming. Nonconformity measures can be computed in several different ways, and each one defines a conformal predictor. In a classification problem with k classes, the p-value for the sample z_{n+1} has to be computed k times. Therefore, it is considered that sample z_{n+1} belongs to each one of the k classes and there will be k p-values. According to (Saunders et al., 1999), the highest p-value, P_1 , determines the class predicted by the algorithm; and the second highest p-value, P_2 , defines the confidence in prediction:

$$\text{Credibility} \rightarrow P_1 \quad (2.67)$$

$$\text{Confidence} \rightarrow 1 - P_2 \quad (2.68)$$

The credibility serves as indicator of how suitable the training data are for classifying the example; while the confidence tells how likely each prediction is of being correct. In the ideal case, a confidence of 99% or higher for all

samples in a set, means that the percentage of erroneous predictions in that set will not exceed 1% (Papadopoulos, 2008, c. 18).

This is the transductive CP approach, where the label of sample z_{n+1} is directly predicted using the training set (see figure 2.6). The computation is carried out for each test sample, and unfortunately it means a high computational cost. On the other hand, the inductive approach, as it was explained previously, extracts from the initial training set a general rule. This general rule is called model or decision rule and it is used to carry out the prediction of the new incoming samples. The CP is a transductive algorithm which can be highly inefficient for large data sets, so that there is an inductive CP approach.

2.3.2.1 Inductive Conformal Prediction

There is also an inductive approach with CP, inductive conformal prediction (ICP). The idea is to carry out the bulk of computations only once. Given a training set $z_1, \dots, z_n \in \mathcal{Z}^n$, classifying a test set of l samples would involve implementing l independent computations. Unfortunately this entails a high computational cost, which can be overcome following the procedure suggested in (Papadopoulos et al., 2002a,b).

1. The initial training set is divided in two subsets: the proper training set z_1, \dots, z_k with $k < n$, and the calibration set z_{k+1}, \dots, z_n .
2. A decision rule or model, F is constructed from the proper training set. This model remains unchanged during the whole ICP process, it means, the computation is done only once.
3. The samples from the calibration set are used to calculate the nonconformity scores of the new samples. So that, for every test sample the nonconformity score and the p-value are computed for every possible label $y \in Y$:

$$\alpha_{n+1} = A(z_{k+1}, \dots, z_n, (x_{n+1}, y)), \quad (2.69)$$

$$p - value_y = \frac{\#\{i = k + 1, \dots, n : \alpha_i \geq \alpha_{n+1}\}}{n - k + 2}. \quad (2.70)$$

4. Predict the label of the test sample using the conformal values 2.67 and 2.68, where the label y_{n+1} of the sample x_{n+1} is given by the largest $p - value_y$.

The main advantage of ICP is their computational efficiency; the bulk of the computations is carried out only once. Then only remains to compute the conformal values for each test sample using the calibration set and the general rule found at the inductive step. On the other hand, the disadvantage

is a possible loss of the prediction efficiency because for CP it can be used the whole training set.

There two approaches of ICP:

- *Off-line ICP*: after classifying a new sample x_{n+1} , it is not added to the calibration set. The advantage is that the calibration set is not modified and it is irrelevant the order of test samples classified.
- *Semi-off-line ICP*: the new sample is added to the calibration set. Therefore, the calibration set is extended with the new sample classified. The advantage is that conformal measures are improved when a new sample is classified and initially it requires a smaller calibration set than off-line ICP.

The on-line version of ICP are explained in detail, with a stronger notion of validity, in (Vovk, 2002) (Vovk et al., 2005, c. 4.1).

2.3.2.2 Examples of Conformal Prediction

Breast cancer classification

This example shows a conformal predictor using the same breast cancer database (see appendix C) as example in section 2.3.1.4. This database is classified using *libsvm*⁴ in Matlab. Following the same steps in section 2.3.1.4, two classifications with 3 models (using RBF, polynomial and linear kernels) have been carried out.

The nonconformity measure chosen to develop the conformal predictor is:

$$\alpha(x) = \begin{cases} - |decisionvalue(H, x)| & \text{if } x \text{ is properly classified} \\ |decisionvalue(H, x)| & \text{if } x \text{ is misclassified} \end{cases} \quad (2.71)$$

where *decision value*(H, x) is a value calculated by SVM which can be understood as the distance from the hyperplane H to the sample x . The reason is that it cannot be used for RBF kernel the formula for linear to calculate w and consequently the distance. Note that for RBF w is infinite dimensional, so it is not possible to have w . Instead it is calculated the decision values. The sign of the decision value gives the predicted label of x . Therefore x is properly classified if $sign(decision\ value(H, x)) = y$, where y is the real label of x and misclassified otherwise.

First, the database is divided in 20 samples for training (10 benign and 10 malignant samples) and 663 for the test (434 benign and 229 malignant samples). The results are shown in table 2.4. The best results are obtained by linear kernel. This table shows the samples used for train and test. Moreover

⁴<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

for each kernel and case (benign and malignant), it is shown the rate of samples properly classified (success), the rate of samples wrongly classified (error), and the mean values of confidence and credibility. After that, a second classification has been made by dividing the database in 100 samples for training (50 benign and 50 malignant samples) and 583 for the test (394 benign and 189 malignant samples). The results are shown in table 2.5. The best results are reached again by linear kernel with an improvement of the success and error rate compare to the smaller training set. In addition the confidence is also higher, it means the classification is more likely to be correct.

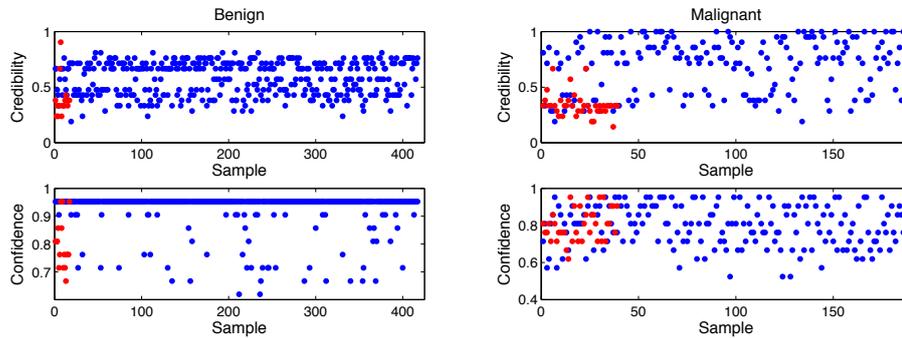
Figures 2.9, 2.10 and 2.11 illustrate the credibility and confidence values obtained for the test set. The values corresponding to the smallest training set (20 samples) are shown in figures 2.9a, 2.10a and 2.11a; and to the largest training set (100 samples) are shown in figures 2.9b, 2.10b and 2.11b. Each figure shows benign and malignant cases separately, and the errors are illustrated in red. It is important to note how the confidence values are placed higher for larger training sets; and errors, in the majority of the cases, are located in the low credibility areas.

| | Benign cases | Malignant cases |
|------------------------------|---------------------|------------------------|
| Train samples | 10 | 10 |
| Test samples | 434 | 229 |
| RBF $\gamma = 0.01, C = 100$ | | |
| Success | 87.33% (379) | 65.50% (150) |
| Error | 12.67% (55) | 34.50% (79) |
| Credibility | 0.8063 | 0.5598 |
| Confidence | 0.7738 | 0.7571 |
| Polynomial $p = 5$ | | |
| Success | 83.87% (364) | 93.01% (213) |
| Error | 16.13% (70) | 6.99% (16) |
| Credibility | 0.4003 | 0.7428 |
| Confidence | 0.7677 | 0.8424 |
| Linear | | |
| Success | 96.08% (417) | 82.97% (190) |
| Error | 3.92% (17) | 17.03% (39) |
| Credibility | 0.5734 | 0.6380 |
| Confidence | 0.9243 | 0.8008 |

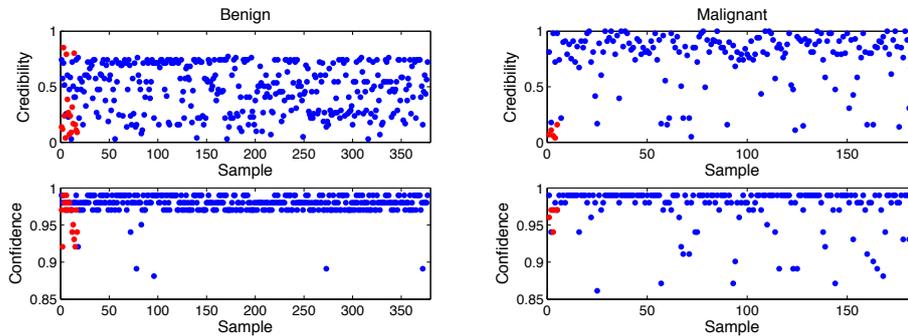
Table 2.4: First CP classification of breast cancer

| | Benign cases | Malignant cases |
|------------------------------|--------------|-----------------|
| Train samples | 50 | 50 |
| Test samples | 394 | 189 |
| RBF $\gamma = 0.01, C = 100$ | | |
| Success | 90.61% (357) | 80.42% (169) |
| Error | 9.39% (37) | 10.58% (20) |
| Credibility | 0.5754 | 0.5846 |
| Confidence | 0.9677 | 0.8862 |
| Polynomial $p = 5$ | | |
| Success | 96.45% (380) | 94.71% (179) |
| Error | 3.55% (14) | 5.29% (10) |
| Credibility | 0.3687 | 0.8199 |
| Confidence | 0.9427 | 0.9391 |
| Linear | | |
| Success | 95.69% (377) | 97.35% (184) |
| Error | 4.31% (17) | 2.65% (5) |
| Credibility | 0.4787 | 0.758 |
| Confidence | 0.978 | 0.9738 |

Table 2.5: Second CP classification of breast cancer



(a) Linear kernel - train 20 samples



(b) Linear kernel - train 100 samples

Figure 2.9: Breast cancer conformal classification with linear kernel

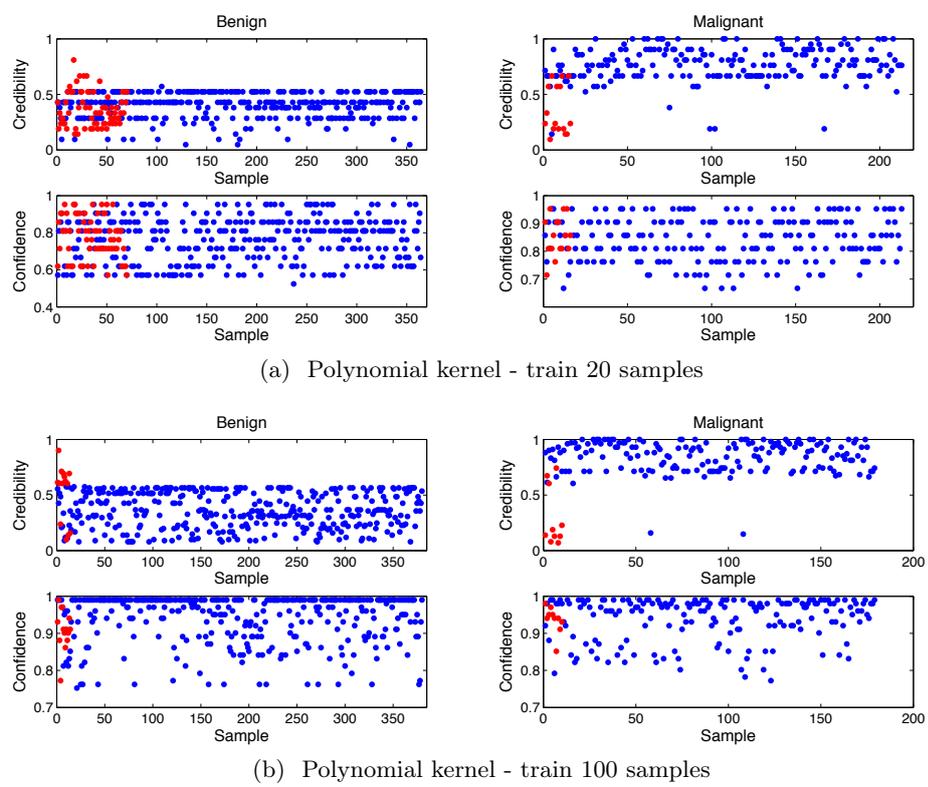
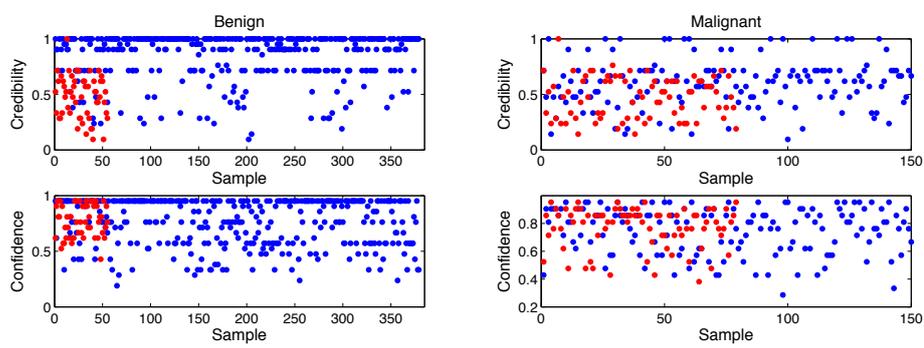
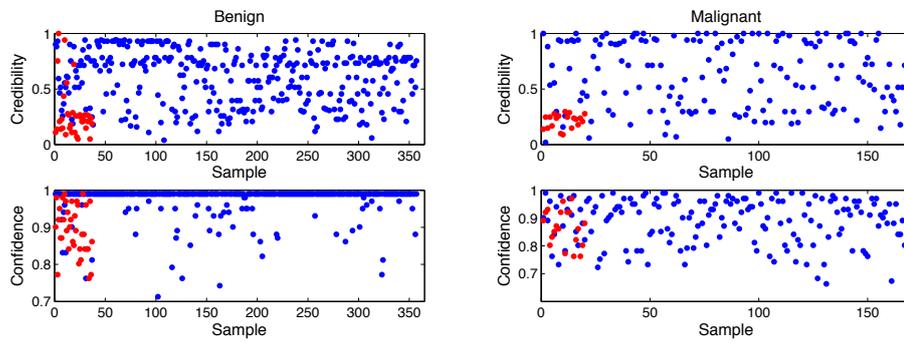


Figure 2.10: Breast cancer conformal classification with polynomial kernel



(a) RBF kernel - train 20 samples



(b) RBF kernel - train 100 samples

Figure 2.11: Breast cancer conformal classification with RBF kernel

2.4 Regression

Regression is a statistical technique, which allows to predict or calculate the behaviour of a dependant variable y as a function of independent variables (x_1, \dots, x_n) . The relation between independent and dependant variables can be expressed by the regression model (Chatterjee and Hadi, 2006, p.2):

$$y = f(x_1, \dots, x_n) + \epsilon, \quad (2.72)$$

where ϵ is defined as the error estimated between the function and the dependant variable. Depending on the dependant variable y , different regression techniques can be applied: Ordinary Least Squares, if y is linear; Ridge Regression, which is a linear regression that solves the problem of collinearity among variables; Conformal Regression, which can be linear or non-linear regression and provides values of confidence and credibility;...

There are several regression techniques, however we will focus on Support Vector Regression (SVR), which have been employed in these thesis in section 4.1.3.

2.4.1 Support Vector Regression

Support Vector Regression (SVR) is a technique that can obtain non-linear regression functions. SVR is a version of SVM, previously explained in section 2.3.1. SVR goal is finding a function $f(x)$ as flat as possible, whose prediction y_i has an error smaller or equal as ϵ for all the training data (Smola and Schölkopf, 2004). Only the points in the region $[-\epsilon, \epsilon]$ are considered, while the points outside the regions are minimised. At this point, the slack variables ξ and ξ^* (figure 2.12) are introduced, the formulation proposed in section 2.3.1.1 yields:

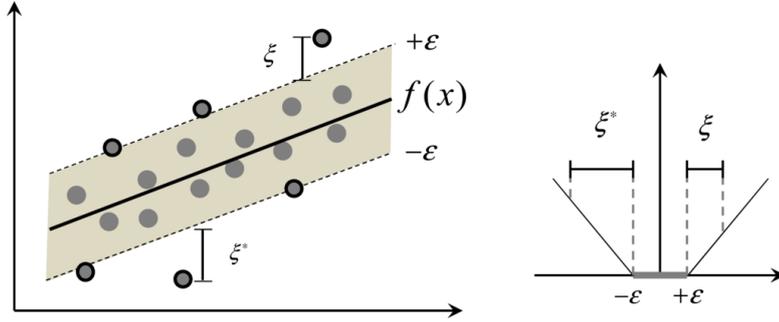
$$\text{Minimise: } \frac{1}{2} \|w\|^2 + \frac{C}{n} \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (2.73)$$

subject to the constraints:

$$\begin{cases} y_i - (w \cdot x_i) - b & \leq \epsilon + \xi_i \\ (w \cdot x_i) + b - y_i & \leq \epsilon + \xi_i^* \\ \xi, \xi_i^* & \geq 0 \end{cases} \quad (2.74)$$

where n is the number of samples and w are the coefficients of the estimated linear function. The parameter $C > 0$ determines the trade-off between the flatness of the function $f(x)$ and the tolerance of ϵ . SVR uses the formulation called ϵ -insensitive loss function $|\xi|_\epsilon$ (figure 2.12):

$$|\xi|_\epsilon = \begin{cases} 0 & \text{if } |\xi| \leq \epsilon \\ |\xi| - \epsilon & \text{otherwise} \end{cases} \quad (2.75)$$

Figure 2.12: ϵ -insensitive loss function

The problem on its dual form can be solved by maximising the Lagrangian:

$$\begin{aligned}
 L(\alpha_i, \beta_i) = & -\frac{1}{2} \sum_{i,j=1}^n (\alpha_i - \beta_i)(\alpha_j - \beta_j)(x_i \cdot x_j) + \\
 & -\epsilon \sum_{i=1}^n (\alpha_i + \beta_i) + \sum_{i=1}^n y_i(\alpha_i - \beta_i)
 \end{aligned} \tag{2.76}$$

subject to the constraints:

$$\begin{cases} \sum_{i=1}^n (\alpha_i - \beta_i) = 0 \\ \alpha_i, \beta_i \in [0, C] \end{cases} \tag{2.77}$$

Solving equation 2.76, SVs, which have non-zero coefficients (α^*, β^*), are provided. The regression function yields:

$$f(x) = \sum_{i=1}^n (\alpha_i^* - \beta_i^*)(x_i \cdot x) + b_0, \tag{2.78}$$

where b_0 is the bias given by:

$$b_0 = y_s - \sum_{i=1}^{SV} (\alpha_i^* - \beta_i^*)(x_i \cdot x_s). \tag{2.79}$$

Using non-linear kernels, the regression function 2.78 is expressed by:

$$f(x) = \sum_{i=1}^n (\alpha_i^* - \beta_i^*)K(x_i, x) + b_0. \tag{2.80}$$

2.4.2 UMEL

In this section, the Universal Multi-Event Locator (UMEL) is described (Vega et al., 2010). UMEL is an automatic tool which allows detecting local events within signals, or it can be used as a filter to recognise signal segments or image areas with relevant information.

UMEL is based on the specific regression technique SVR, explained in the previous section. SVR fits the data regardless of factors such as sampling rate or noise distribution. It computes a fitting functions and provides the SVs from the training set. In SVM, in this case SVR, the number of SVs is determined by the complexity of the model; it means simple data sets would require smaller numbers of SVs, while complex data sets would require large numbers of SVs. However, the number of SVs also depends on the smoothness of the regression function. A smoother function requires less SVs than a spiky function.

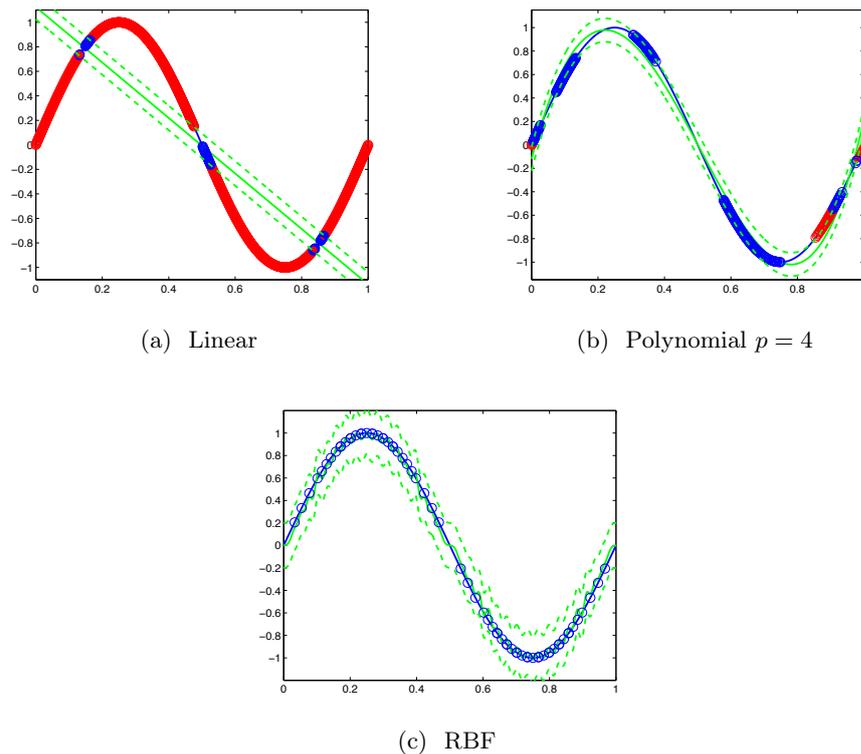


Figure 2.13: UMEL example with different types of kernel

The smoothness of a SVR fit is function of four parameters:

1. *Kernel type*: there are many kernel types, e.g. linear, polynomial or RBF. In figure 2.13 it can be seen an example of UMEL applied to a

sinusoidal signals with 3 different kernels. Figure 2.13a corresponds to a linear kernel, it does not fit correctly and follows practically a straight line. Figure 2.13b illustrates a fit computed using a polynomial kernel which follows much better the global tendency. Finally figure 2.13c shows a RBF fitness, which corresponds to the best fit.

2. *Parameter ϵ* : this parameter determines the margin size of the ϵ -insensitive zone, which is also called ϵ -tube. It can be calculated by using the equation (Cherkassky and Mulier, 2007, p. 449):

$$\epsilon = K_\epsilon \cdot \sigma_{noise} \cdot \sqrt{\frac{\ln(n)}{n}} \quad (2.81)$$

where σ_{noise} is the standard deviation of the noise of the function, n is the number of training samples, $\ln(n)$ is the natural logarithm of the number of training samples, and K_ϵ is a proportionality constant. Due to the value of the noise is not usually known, the variable σ_{noise} is replaced by σ_y , where y is the function to be fitted. Therefore, the parameter ϵ is settled by using K'_ϵ .

3. *Regularization parameter C* : it controls the model complexity (equation 2.58). Low values of C provides smooth regressions, and large values of C provides spiky regressions. Its values is given by (Cherkassky and Mulier, 2007, p. 449):

$$C = K_C \cdot \max(|\bar{y} + 3 \cdot \sigma_y|, |\bar{y} - 3 \cdot \sigma_y|) \quad (2.82)$$

where \bar{y} is the mean of the function to be fitted, σ_y is the standard deviation of the function to regress, and K_C is a constant that can vary for different kind of signals. This estimation can effectively handle outliers in the training data.

4. *Kernel parameters*: depending on the kernel used, it can include different parameters. For example, for polynomial kernel it must be chosen the polynomial degree p ; or σ parameter for RBF kernel. RBF is the most used kernel, so it exists an equation to estimate the parameter σ (Martinez and Martinez, 2008, p. 325):

$$\sigma_k = K_\sigma \cdot 1.06 \cdot \sigma_y \cdot n^{-1/5} \quad (2.83)$$

where σ_y is the standard deviation of the function to regress, n is the number of samples, and K_σ is a constant dependent on the type of signal to regress.

UMEL provides different SVs which have not the same relevance. The SVs that lie on or outside the ϵ -insensitive margin are called External Support

Vectors (ESVs). On the other hand, the SVs within the ϵ -insensitive margin are called Internal Support Vectors (ISVs):

$$ESV \subseteq SV \text{ text } \forall i \in ESV, |y_i - f(x_i)| \geq \epsilon \quad (2.84)$$

$$ISV \subseteq SV \text{ text } \forall i \in ISV, |y_i - f(x_i)| < \epsilon \quad (2.85)$$

ISVs are necessary samples for the regression estimation. UMEL considers the ESVs the most difficult samples to regress, due to they cannot be fitted inside the ϵ -insensitive margin. Therefore, ESVs provide essential information in the regression process, they show special patterns or behaviours in the signal: peaks, gradients, anomalous behaviours, etc.

This tool has been used in section 4.1.3 to detect the disruption time as a sudden fall of the plasma current, it means as a local event which supposes a different pattern on the normal evolution of the signal. In figure 2.14, it is shown an example of UMEL applied to a signal with multiple pikes. It can be seen how SVR fit the function in the green line, and the ϵ -tube in the discontinuous green line. In blue and red circles are marked the ISVs and ESVs respectively.

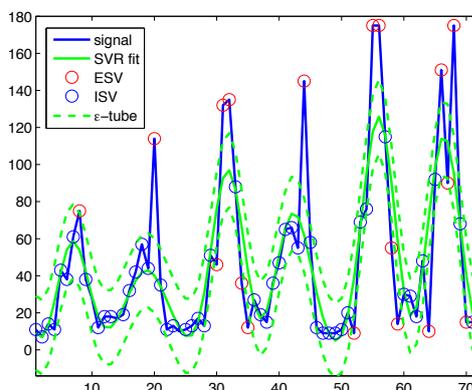


Figure 2.14: UMEL example

2.5 Martingales

Martingales are able to detect changes in data streams by testing exchangeability (Ho and Wechsler, 2010). In the martingale framework, the hypothesis testing decides whether a change has occurred when a new sample is observed, this method has been used and can be found widely explained in works like (Ho and Wechsler, 2010; Vega et al., 2015b). First of all, let

see a brief explanation of martingale and exchangeability concepts. Given a sequence of random variables $Z_i : 1 \leq i < \infty$, the sequence Z_1, \dots, Z_n is *ex-changeable* if for every finite subset of the random variable sequence (containing n random variables), the joint distribution $p(Z_1, \dots, Z_n)$ is invariant under any permutation of the random variables, i.e.,

$$p(Z_1, \dots, Z_n) = p(Z_{\Pi(1)}, \dots, Z_{\Pi(n)}), \quad (2.86)$$

for all permutations Π defined on the set $1, \dots, n$. Ex-changeable random variables are identically distributed and can be either independently or dependently distributed.

The concept of martingale, which comes from gambling, describes a fair game of chance. Considering a person gambling in a dice game, odd results are considered winnings and even are considered defeats. If Z_i is the money that the person earns on the *ith* game, this value will be negative when he loses. M_i is defined as the total number of winnings at the end of the *ith* game. In every game, Z_i is independent of all previous game earnings, so that the conditional expectation $E(Z_i | Z_1, \dots, Z_{i-1}) = E(Z_i)$. Therefore the conditional expectation of the winnings M_{n+1} is the value of the current random variable M_n . A martingale is a sequence of random variables that remains stable in value with some fluctuations as long as the process is random, i.e., without any external interference. Therefore a sequence of random variables $M_i : 0 \leq i < \infty$ is a martingale with respect to the sequence of random variables $Z_i : 0 \leq i < \infty$ if, for all $i \geq 0$, the following conditions hold:

- M_i is a measurable function of Z_0, Z_1, \dots, Z_i ,
- $E(|M_i|) < \infty$, and
- $E(M_{n+1} | Z_0, \dots, Z_n) = M_n$.

The idea of testing exchangeability on-line using a martingale was firstly introduced in (Vovk et al., 2003). An observer outputs a positive martingale value for each new incoming data point, reflecting the strength of evidence found against the null hypothesis of data exchangeability. Firstly, it is needed to explain the concept of *strangeness measure* before introducing a martingale for change detection. The *strangeness measure* scores how much a data point is different from the other data points. There are different strangeness measures related to classification, clustering and regression (Ho and Wechsler, 2010). For example, given an unlabelled training set $X = x_1, \dots, x_n$, the strangeness of a sample x_i with respect to a cluster model is defined as (Ho and Wechsler, 2010, 2007):

$$s(X, x_i) = \|x_i - c\|, \quad (2.87)$$

where c is the cluster centre of the training set X and $\|\cdot\|$ is a distance metric.

Randomized Power Martingale and \hat{p} -values

Using any strangeness measure, a family of martingales, indexed by $\epsilon \in [0, 1]$ and referred to as the *randomized power martingale* (RPM) (Vovk et al., 2003; Ho and Wechsler, 2010), is defined as:

$$M_n^{(\epsilon)} = \prod_{i=1}^n (\epsilon \hat{p}_i^{\epsilon-1}), \quad (2.88)$$

where \hat{p}_i values are computed from the \hat{p} -value function

$$\hat{p}_i((x_1, y_1), \dots, (x_i, y_i), \theta_i) = \frac{\#j : s_j > s_i + \theta_i \#j : s_j = s_i}{i}, \quad (2.89)$$

where s_j is the strangeness measure for (x_j, y_j) with $j = 1, \dots, i$, and θ_i is a random value in the range $[0, 1]$ at instance i . The initial martingale value is $M_0^{(\epsilon)} = 1$. These \hat{p} -values are a special case of the statistical notion of p -value in (Ho, 2008).

Simple Mixture Martingale

From RPM, equation 2.88, in order to eliminate the dependency on ϵ , a different martingale can be used (Vovk et al., 2003):

$$M_n = \int_0^1 M_n^{(\epsilon)} d\epsilon \quad (2.90)$$

This martingale is called the *simple mixture martingale* (SMM) of $M_n^{(\epsilon)}$.

Plug-In Martingale

A third family of martingales can be considered. It is called, *Plug-In martingale* (PIM) and it is fully described in (Fedorova et al., 2012). The procedure in PIM involves the computation of probability density functions using \hat{p} -values output by martingale test algorithm. This is achieved using kernel methods (Martinez and Martinez, 2008). It has been used a Gaussian kernel and the normal reference rule for bandwidth selection. In particular, the bandwidth range used in the probability density function estimations covers the interval $K_N \cdot N_{REF}$, where N_{REF} is the value corresponding to the normal reference rule.

Martingale Test

Let see an important property of the \hat{p} -values obtained from equation 2.89, using data generated from a source that satisfies the exchangeability assumption in (Ho and Wechsler, 2010):

- *Lemma: \hat{p} -values \hat{p}_i , $i = 1, 2, \dots$ from 2.89 are independent and uniformly distributed on $[0, 1]$. (See [32, Theorem 8.2 and Section 8.7].)*

The Lemma fails if the observed data samples are no longer exchangeable. At this point, the \hat{p} -values obtained from 2.89 have smaller values. In addition, they are no longer uniformly distributed on $[0, 1]$, because the newly observed data samples are likely to have higher strangeness values compared to the previously samples. As a result, $M_n^{(\epsilon)}$ increases and provides evidence against the null hypothesis H_0 of exchangeability in a data stream. This property of $M_n^{(\epsilon)}$ is employed in the change detection in a data stream. In the martingale framework for change detection, when a new data sample is observed, hypothesis testing takes place to decide whether a change in the data generating model occurs in the data stream. The decision is based on whether the exchangeability assumption is violated using a RPM, SMM or PIM.

If it is considered the null hypothesis H_0 , it means no change in the data stream against the alternative H_1 , a change occurs in the data stream. The martingale test continues to operate as long as

$$0 < M_n^{(\epsilon)} < \lambda, \quad (2.91)$$

where λ is a positive number. The null hypothesis H_0 is rejected when $M_n^{(\epsilon)} \geq \lambda$. Following the procedure in (Ho and Wechsler, 2010), suppose that $M_k : 0 \leq k < \infty$ is a non-negative martingale. If $E(M_n) = E(M_0) = 1$, then from the Doob's Maximal Inequality (Steele, 2001),

$$\lambda P \left(\underbrace{\max}_{k \leq n} (M_k) \geq \lambda \right), \quad (2.92)$$

it is reached the inequality

$$P \left(\underbrace{\max}_{k \leq n} (M_k) \geq \lambda \right) \leq 1/\lambda, \quad (2.93)$$

for $\lambda \geq 0$ and $n \in \mathbb{N}$. Equation 2.93 shows that it is unlikely for any M_k to have a high value. The null hypothesis is rejected when $M_k > \lambda$. However it is still possible to detect a change when there is none. The equation 2.93 is an upper bound for the false alarm rate for detecting a change when there is

none. The value of λ is determined by the false alarm rate that one is willing to accept. The martingale test algorithm is defined as (Ho and Wechsler, 2010), see appendix D.

This method for change detection in data stream has been used as an approach to locate the disruption time in section 4.1.2.

Chapter 3

Databases and APODIS during ILW

In this chapter it will be explained the Advanced Predictor of Disruptions, APODIS in section 3.1. Then the procedure for generating databases will be exposed in section 3.2. In section 3.3, it is carried out the assessment of APODIS during ILW experimental campaigns at JET. After this, the robustness analysis is illustrated in section 3.4. Finally it is shown the implementation of the sliding window mechanism following APODIS structure in section 3.5.

3.1 Advanced Predictor of Disruptions, APODIS

The Advanced Predictor of Disruptions (APODIS) is a multilayer predictor based on support vector machines (SVM) classifiers. This project started in 2008 (Rattá et al., 2008). The current version in the real time network at JET (López et al., 2012) differs in some aspects from the first version (Rattá et al., 2010): development in real time, number of discharges used for training, signals selected and an exhaustive pre-processing to delete outliers.

This predictor uses a set of models based on SVM, a machine learning technique explained in section 2.3.1. It classifies the samples from the incoming discharge between disruptive or non disruptive. Seven signals available in real time are used to characterize plasma, see table 3.1.

For the training of APODIS, a total of 10845 discharges from campaigns between 24th April 2006 and 23th October 2009 (C15-C27b) were taken. These initial discharges were analysed following a similar process used in section 3.2:

- A visual inspection of the discharges and the different signals is used to remove discharges with signals without data or wrong measures.

| Signal name | Units |
|------------------------------------|----------|
| Plasma current | A |
| Mode lock amplitude | T |
| Plasma internal inductance | – |
| Plasma density | m^{-3} |
| Diamagnetic energy time derivative | W |
| Radiated power | W |
| Total input power | W |

Table 3.1: Signals used in APODIS predictor

- A threshold analysis removed the discharges with plasma density $> 10^{22}$, mode lock amplitude $> 6 \cdot 10^{-3}$, plasma internal inductance > 10 and poloidal beta > 30 .

Table 3.2 shows 8407 discharges selected, it means 2438 removed. From campaigns $C19 - C22$, 100 non-disruptive of 2312 discharges randomly selected and 125 non intentional disruptive discharges (all possible) have been used for the training models. From campaigns $C23 - C27b$, 3578 non-disruptive discharges and 228 non intentional disruptive discharges have been employed for the test and validation.

| Campaigns | Non-disruptive | Non-int. discr. | Int. discr. | Total |
|---------------------|----------------|-----------------|-------------|-------------|
| $C19$ (02 – 04/07) | 585 | 47 | 41 | 673 |
| $C20$ (04 – 06/08) | 703 | 28 | 12 | 743 |
| $C21$ (06 – 07/08) | 573 | 16 | 3 | 592 |
| $C22$ (07 – 08/08) | 451 | 34 | 3 | 488 |
| $C23$ (09 – 10/08) | 490 | 24 | 8 | 522 |
| $C24$ (10 – 11/08) | 362 | 14 | 12 | 388 |
| $C25$ (11 – 12/08) | 570 | 19 | 22 | 611 |
| $C26$ (01 – 03/09) | 1323 | 58 | 49 | 1430 |
| $C27a$ (07 – 08/09) | 320 | 43 | 10 | 373 |
| $C27b$ (09 – 10/09) | 513 | 70 | 59 | 642 |
| Total | 7648 | 521 | 238 | 8407 |

Table 3.2: Discharges used on Apodis training

The features are selected using the tools and methods described in (Murari et al., 2008b; Rattá et al., 2008), hence, the feature vectors are formed by the mean value and the standard deviation of the Fourier transform module (the continuous components are erased) of each signal, every 32 ms. This resolution was chosen because the calculus of Fourier transform is faster for the power that is multiple of 2; and the mitigation valve in JET at this time needed at least 30 ms to carry out the action.

The temporal evolution of the discharge is analysed in temporal win-

dows of 32 ms, figure 3.1. The classification system analyses every window consecutively and decides to trigger or not an alarm to notice an incoming disruption. If any alarm is triggered, then the predictor has decided there is no disruptive behaviour and the discharge is safe (non-disruptive).

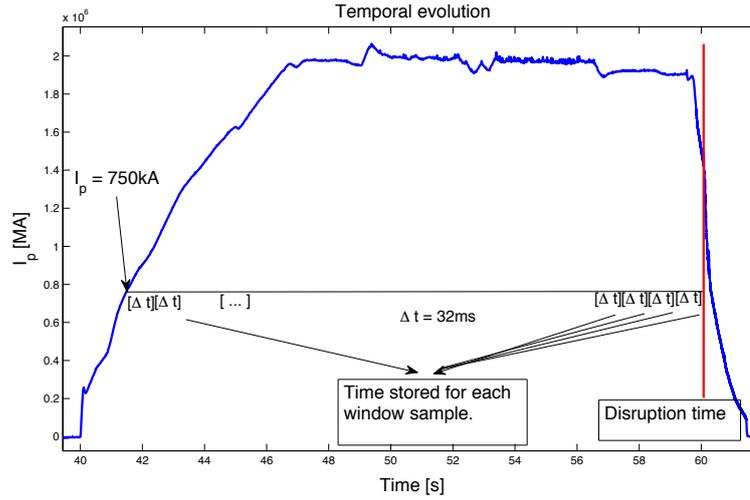


Figure 3.1: Temporal evolution of a JET discharge

The structure of the predictor, which is illustrated in figure 3.2, is compound by two layers. The first layer is formed by the combination of three consecutive models. Each model is a SVM classifier trained with a RBF kernel (section 2.3.1). The difference of these models is the disruptive features considered for the training process, where:

- Model $M1$ takes the window $(-64 \text{ ms}, -32 \text{ ms}]$ before the disruption time as disruptive sample.
- Model $M2$ takes the window $(-96 \text{ ms}, -64 \text{ ms}]$ before the disruption time as disruptive sample.
- Model $M3$ takes the window $(-128 \text{ ms}, -96 \text{ ms}]$ before the disruption time as disruptive sample.

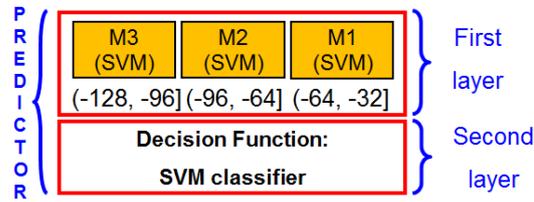


Figure 3.2: APODIS structure formed by two layers

The temporal window $(-32 \text{ ms}, 0 \text{ ms}]$ is not considered in this predictor as disruptive sample. Three output values are obtained after the three model analyse 3 consecutive windows. These 3 values are introduced as input for the second layer, which is a linear SVM classifier that works as decision function. This second layer obtained a final value which show if the behaviour is disruptive or not. The time assigned to the evaluation is the one which correspond to the model that analyses the last temporal window, $M1$. These procedure is illustrated clearly in figure 3.3, as a discharge is in execution, the most recent 32 ms temporal segments are classified as disruptive or non-disruptive, the three models may disagree about the discharge behaviour and the second layers decides.

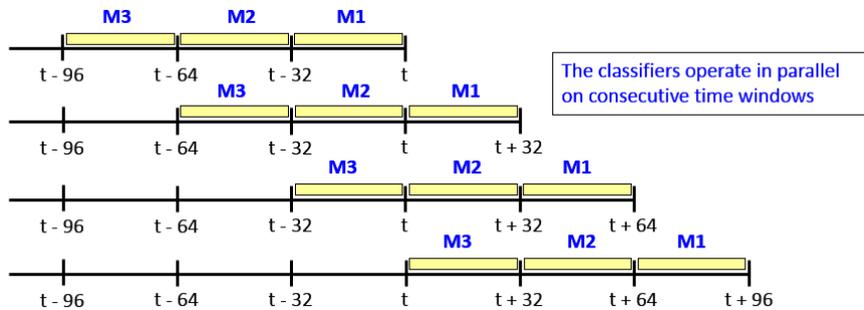


Figure 3.3: APODIS analysis procedure during a discharge

Following the training for both layers will be described. The first layer takes:

- For non-disruptive discharges, the first 22 window samples are not considered to avoid values out of range in the beginning of the discharge. From the window 23 until the end of the discharge, one sample in, one sample out, are considered as non-disruptive samples.
- For disruptive discharges, each model considers its temporal window as disruptive sample.

Then three models are trained with a kernel RBF in SVM.

For the second layer, the discharges are evaluated by the models obtaining three signed outputs (distances), which are called $V1$, $V2$ and $V3$ for $M3$, $M2$, and $M1$ respectively. The decision rule employed is:

- If ($V3 > -0.8$ and $V2 > -0.4$ and $V1 > 0$) then: an alarm is triggered and a row must be written in the table of alarms. This mean a disruptive sample for the second layer.
- Every 64 ms a row is written in the table of alarms as non-disruptive sample if any alarm is triggered.

There are 3 possible scenarios:

1. Right alarm: in a disruptive discharge, the warning time (time of the triggered alarm) is less than 1 s before the disruption time. The label of this sample is +1.
2. Premature alarm: in a disruptive discharge, the warning time is higher than 1 s before the disruption time. The label of this sample is -1.
3. False alarm: in a non-disruptive discharge an alarm is triggered. The label of this sample is -1.

The discharge is evaluated from the beginning until an alarm is triggered or until reach the last window. Once all discharges have been evaluated and the table of alarms is finished, the second layer is trained with a linear kernel. The table of alarms is shown in the table 3.3, the red cells are used for the training of the second layer, and examples of the different possible scenarios are shown. The model obtained in the second layer provides the coefficients of the hyperplane which classify a sample as disruptive or safe.

| Shot | t_{alarm} [s] | t_{dis} [s] | Δt [ms] | $M1_{output}$ | $M2_{output}$ | $M3_{output}$ | Label |
|-------|-----------------|---------------|-----------------|---------------|---------------|---------------|-------|
| 56658 | 63.911 | 64.021 | 110 | -0.910 | 0.448 | 0.0852 | +1 |
| 53740 | 45.881 | 50.22 | 4339 | -0.896 | -1 | 0.445 | -1 |
| 52461 | 40.916 | 0 | -40916 | -1.176 | -1.202 | -1.008 | -1 |

Table 3.3: Table of alarms for the second layer

3.2 Databases and signal processing

Currently, Joint European Torus is the largest nuclear fusion device in the world as it was mentioned and explained in section 1.1.3.1. From the onset JET was designed to use a computer based integrated control and data acquisition system (CODAS) (Krom, 1999). This system has evolved over the

years and huge databases are stored in JET. In (Farthing et al., 2007), some simulations for data volume are done where it is shown that ≈ 10 GB/pulse = 0.25 TB/day (figure 3.4 are produced reaching ≈ 60 GB/pulse in 2010. It is estimated to store ≈ 40 TB of data in 25 years and approximately $\leq 10\%$ would be processed. More recent is the information given in (Murari, 2014) where it is told that data acquired per pulse is ≥ 50 GB and the total is ≥ 250 TB.

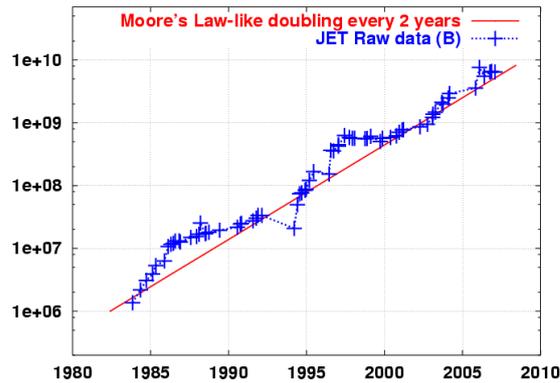


Figure 3.4: JET raw data estimated in bytes (B)

The preparation of the databases is one of the most important steps in data mining. The knowledge extracted from data is an iterative sequence of some steps (Jiawei et al., 2011):

- Data cleaning (to remove noise and inconsistent data).
- Data integration (where multiple data sources may be combined).
- Data selection (where data relevant to the analysis task are retrieved from the database).
- Data transformation (where data are transformed or consolidated into forms appropriate for mining by performing summary or aggregation operations, for instance).
- Data mining (an essential process where intelligent methods are applied in order to extract data patterns).
- Pattern evaluation (to identify the truly interesting patterns representing knowledge based on some interestingness measures).
- Knowledge presentation (where visualization and knowledge representation techniques are used to present the mined knowledge to the user).

Steps 1 to 4 are different forms of data preprocessing, where the data are prepared for mining. It could be considered these steps as the highest workload of the process.

The development of the databases for disruptions needs several steps, beginning with the analysis of the discharges from every experimental session at JET where many conditions have to be applied. Once the analysis is finished, the signal processing is realised to obtain the processed data. In this thesis the whole ILW campaigns $C28 - C34$ (September 2011 - October 2014) have been analysed and processed for several studies. These campaigns are extremely important because of the new ITER-like wall (ILW) and the high performance plasmas (plasma current 3.5 MA, input power 25 MW, electron temperature 1 keV and density $2 \cdot 10^{20} \text{ m}^{-3}$). In JET, the number of disruptions varies between campaigns (de Vries et al., 2012). Comparing JET ILW (ITER-like wall) campaigns $C28 - C34$ (September 2011-September 2014) with the C-wall campaigns $C15 - C27b$ (April 2006-October 2009), the number of non-intentional disruptions has notably increased, although it is expected to be reduced as experience is gained on how to operate JET with the new all (de Vries et al., 2012; Matthews, 2013). This is illustrated in figure 3.5, on the left of the black line C-wall campaigns are shown and on the right ILW campaigns. The percentage of non-intentional disruptions in every campaigns per date is shown in blue bars and the percentage of non-disruptive shots is illustrated in red bars. This count considers the databases used in disruption prediction studies. Globally it can be seen that C-wall campaigns contain 474 non-intentional disruptions and 7648 safe discharges, while during ILW campaigns 589 non-intentional disruptions and 2067 safe discharges happened.

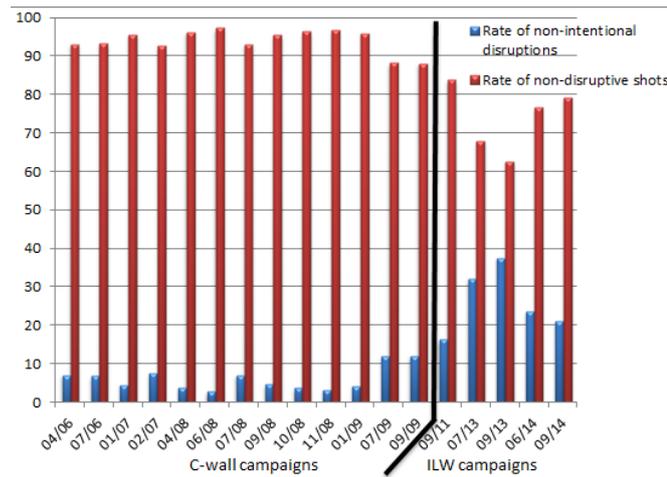


Figure 3.5: Rate of non-intentional disruptions and non-disruptive discharges in JET

3.2.1 Analysis of JET campaigns

First of all, starting from the first discharge until the last one, a visual inspection of the discharges and the different signals is used to remove discharges with signals without data or wrong measures. The main plasma parameters, shown in table 3.4, and other useful signals to analyse disruptions, shown in table 3.5, are observed. If every signal is right and there is no anomalous behaviour at first sight, as it is illustrated in figures 3.6a and 3.6b, the discharge is considered for the database. Albeit, if there is any failure in a diagnostic or anomalous behaviour in any signal, the discharge is not taken into account, as it is illustrated in figure 3.7. Data is stored in two ways at JET: the JET pulse file (JPF) server, and the processed pulse file (PPF) system (Layne and Wheatley, 2002). The data used for the databases is JPF because is the one available in real time. Moreover all the discharges which belong to restart or reconditioning sessions are not considered.

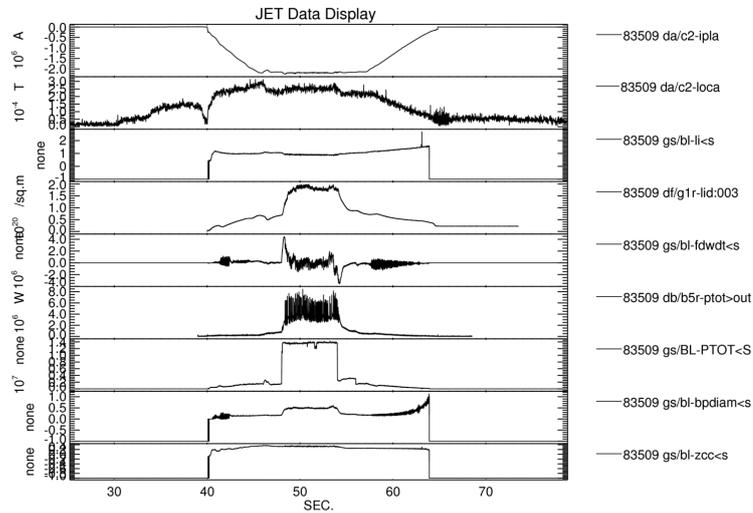
| Signal name | DDA/Datatype |
|------------------------------------|---------------------|
| Plasma current | jpf/da/c2-ipla |
| Mode lock amplitude | jpf/da/c2-loca |
| Plasma inductance | jpf/gs/bl-li<s |
| Plasma density | jpf/df/g1r-lid:003 |
| Diamagnetic energy time derivative | jpf/gs/bl-fdwdt<s |
| Radiated power | jpf/db/b5r-ptot>out |
| Total input power | jpf/gs/BL-PTOT<S |
| Poloidal beta | jpf/gs/bl-bpdiam<s |
| Plasma vertical centroid position | jpf/gs/bl-zcc<s |

Table 3.4: Signals used in disruption databases with their directory of JPF signals in JET database

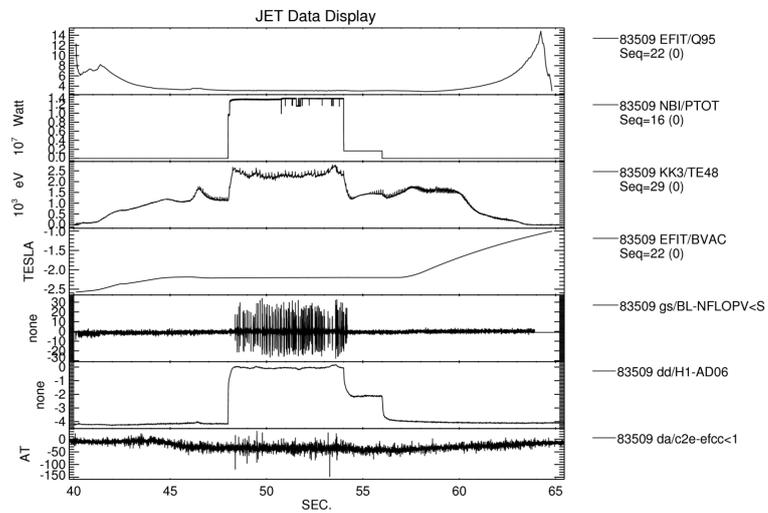
| Signal name | DDA/Datatype |
|------------------------------------|--------------------|
| Safety factor, Q95 | ppf/efit/q95 |
| Neutral beam injection, NBI | ppf/nbi/ptot |
| Plasma electron temperature | ppf/kk3/te48 |
| Vacuum BPHI at 2.96M | ppf/efit/bvac |
| Loop voltage | jpf/gs/bl-nflopv<s |
| Soft X-rays | jpf/dd/h1-ad06 |
| Error field correction coils, EFCC | jpf/da/c2e-efcc<1 |

Table 3.5: Recommended signals for disruption analysis

After the visual inspection, a deep analysis is carried out. It is observed if the plasma evolves without instabilities or control action which end in disruption or anomalous behaviour. For this purpose, the comments of the



(a) Main plasma parameters



(b) Recommended signals for disruptions

Figure 3.6: Example of a right discharge to be considered for the databases

session leader and the outputs of the JPS (JET Protection System) help to understand the experiment. The case any control action has been taken and

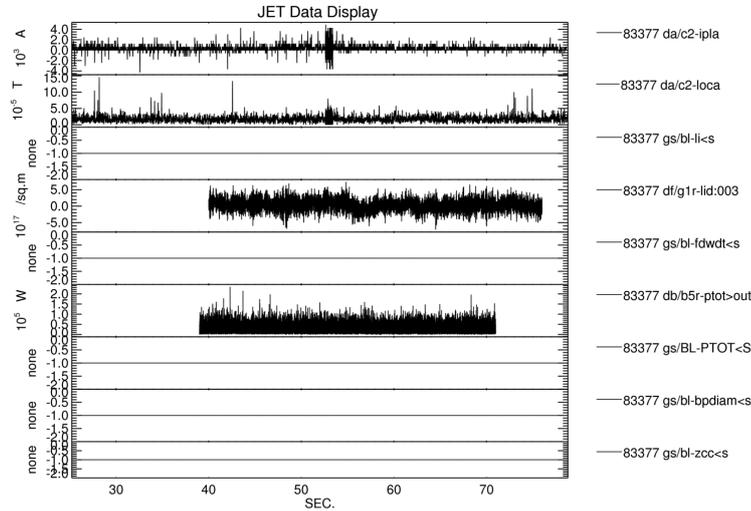
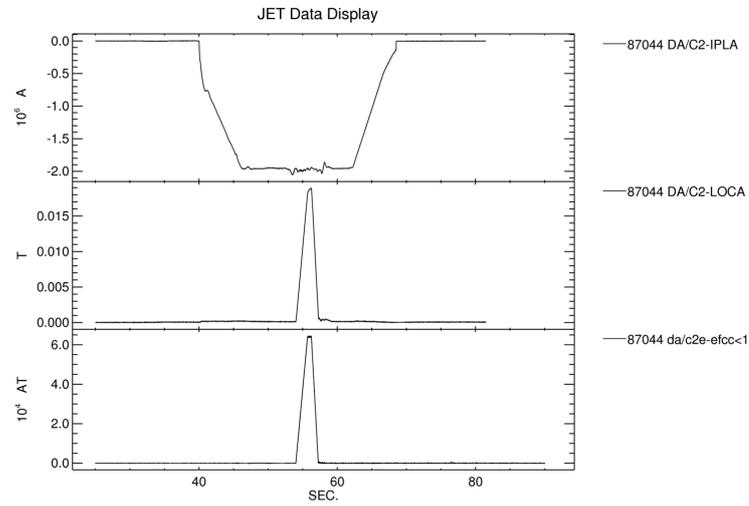


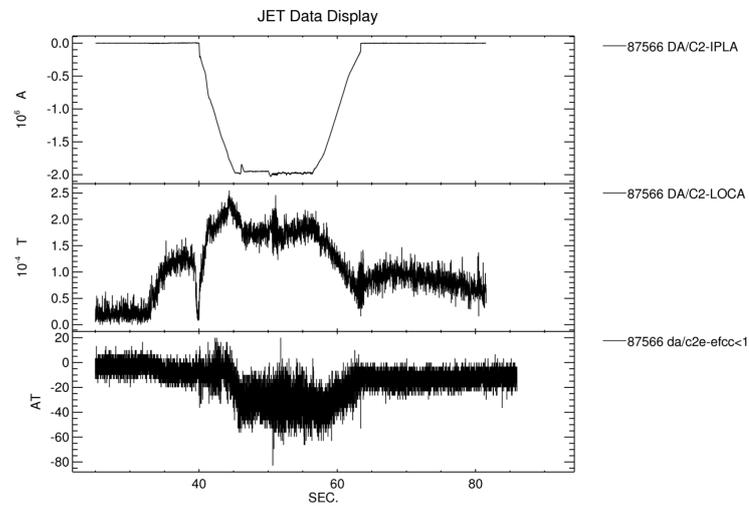
Figure 3.7: Example of a discharge not considered in the database with errors in several signals

changes the normal evolution of the plasma, the discharge is not considered disruptive neither non-disruptive. We catalogue these shots as “strange” discharges, because they cannot be taken into account in disruption studies or predictors. For example, during discharge 84634 the JPS triggers a stop which ends the plasma, so that this shot cannot be used as a non-disruptive or disruptive discharge for the database. There are several cases with different causes and consequences so it is not possible to enumerate all of them. In order to show other examples, the discharge 84670, 84693, 84912 are minor disruptions; discharge 84991 is a disruption but the plasma current is below 1 MA, therefore it has no relevance; discharge 85919 shows a low density event at the beginning of the plasma.

Following the previous case, it is necessary to catalogue separately the discharges where EFCC (Error Field Correction Coils) have been used. The EFCC are specified to compensate the main harmonic contribution to error field induced modes (Barlow et al., 2001). As it can be seen in figure 3.8, when EFCC are activated the mode lock signal follows the same trend as EFCC signal; however if EFCC are deactivated, there’s no effect on mode lock signal. Therefore it is better to not consider this discharge for the database because it would mean false information for the models. Despite of this, these discharges are listed and stored separately in order to carry out other studies.



(a) Discharge with EFCC activated



(b) Discharge without EFCC

Figure 3.8: Examples of EFCC effect on mode lock amplitude

Finally, if there is no disruption, the discharge is stored as non-disruptive, but if the discharge is disruptive, we have to look up if it is intentional or

non-intentional. In one hand, if the disruption is intentional, it is usually indicated by session leader. This is a common experiment to study the mitigation of disruption, electron runaways, vertical displacement disruptions, etc. In the case the session leader has not written it, it can be seen if the DMV (disruption mitigation valve, a gas is introduced to mitigate the disruption) has been programmed to be triggered. On the other hand, the discharge will be a non-intentional disruption. Both cases will be considered in the database although intentional disruptions must not be considered in prediction studies because they can introduce false information in the model. In figure 3.9, examples of intentional and non-intentional disruptions are illustrated. In the intentional disruption example it can be observed there is no precursors or previous disruptive behaviour until the disruption time, the disruption suddenly happens ending the plasma as it had been programmed. On the other hand, the non-intentional example shows how the plasma becomes unstable before disruption time. Because of this reason, intentional disruption are useless on disruption studies.

All of this must be done by hand, the different comments of the session leaders, outputs of the diagnostics and different types of actions and events make almost impossible to implement an automatic analysis. Despite of this, some parts of this analysis has been automatized: the use of EFCC, the triggers of the DMV, triggers of APODIS and JPS.

Once all the discharges are classified, the disruption time has to be obtained. This is the most important step because all the disruptive information is around this time, hence the most precise is the time determined, the best information we could extract as disruptive samples. This time is usually determined by hand, but automatic tools have been developed in this thesis to make easier this work (Moreno et al., 2014a). Therefore, this topic will be described in section 4.

3.2.2 Signal processing

This time the analysis is done, the discharges are listed in two files. The first file contains two columns with the number of the discharge and a value 0, 1 or 2, indicating if the shot is non-disruptive, non-intentional or intentional, respectively. The second file contains two columns with the number of the disruptive discharge and the disruption time. Then the raw data of the signals mentioned in 3.4 is downloaded from JET database.

We will have downloaded 9 signals (table 3.4) for every discharge. Three more signals are calculated: plasma inductance derivative (signal 10), beta poloidal derivative (signal 11) and vertical centroid position derivative (signal 12). For every discharge, it is selected the time for the beginning and the end. As initial time is selected the maximum value of the set formed by the minimum temporal values of each signal; the last time is selected as the minimum temporal value of the set compound of the maximum temporal values of each signal.

A threshold is applied in the amplitude of some signals:

- Maximum value of plasma current signal (signal 1) is set to 0. Any positive value is changed to 0.
- Minimum value of mode lock signal (signal 2) is set to 0. Any negative value is changed to 0.
- Minimum value of density signal (signal 4) is set to 0. Any negative value is changed to 0.
- Minimum value of radiated power signal (signal 6) is set to 1000. Any amplitude value lower than 1000 is changed to 1000.
- Minimum value of total input power signal (signal 7) is set to 1. Any amplitude value lower than 1 is changed to 1.

After that an analysis to detect outliers is carried out. If the amplitude of the signal is higher than a threshold, the discharge is considered an outlier unless:

- The threshold is exceed after disruption time.
- The threshold is exceed before reaching the value of 750 kA in the plasma current at the beginning of the discharge.
- The threshold is exceed after reaching the value of 750 kA in the plasma current at the end of the discharge.

In these cases there is no problem cause the data is out of the useful boundaries to develop studies. The signals and thresholds considered in the outliers analysis are:

- Mode lock amplitude $> 6 \cdot 10^{-3}$
- Plasma inductance amplitude > 13
- Plasma density amplitude $> 1 \cdot 10^{22}$
- Poloidal beta amplitude > 30

Then every signal is interpolated increasing the original sampling rate of the sequence to 1 kHz. In this way, we can process data in different temporal windows depending our needs. The interpolation used in these databases is a linear interpolation, as it is explained in section 2.1, this is not the best interpolation but it is a realistic approach thinking in real time application, where we would have to interpolate using nearest left-neighbour, and good enough for offline studies.

Following it has to be done data standardisation, so that the maximum and minimum for each signal is calculated. As it is explained in section 2.1, the data is normalized, although the right way would be to calculate the values for all the data, it has been determined maximum and minimum for campaigns $C28 - C30$, and campaigns $C31 - C34$ has not been taken in account for these values. Under the basis that the values should be similar it is thought this way is better for more realistic approaches. If we use the first campaigns to generate models and the later ones as test, it reproduces better a process in real time. Therefore, the standardisation parameters yield:

| Signal name | Minimum | Maximum |
|--|--------------------|------------------------|
| Plasma current | $-3.55 \cdot 10^6$ | 0 |
| Mode lock amplitude | 0 | $6 \cdot 10^{-3}$ |
| Plasma inductance | -1 | 12.9698 |
| Plasma density | 0 | $9.5031 \cdot 10^{21}$ |
| Diamagnetic energy time derivative | -28460799 | 15269191 |
| Radiated power | 1000 | $1.8297 \cdot 10^9$ |
| Total input power | 1 | 55606192 |
| Poloidal beta | -1 | 30.5776 |
| Plasma vertical centroid position | -1 | 1.7989 |
| Plasma inductance derivative | -12.6588 | 11.9698 |
| Poloidal beta derivative | -16.1936 | 14.3687 |
| Plasma vertical centroid position derivative | -1.4129 | 1.1175 |

Table 3.6: Standardisation parameters for databases during ILW campaigns $C28 - C34$

The final step is data processing, following the calculus mentioned in section 2.1, for each signal the mean value and the standard deviation of the Fourier transform is determined every 32 ms. The database is finally compound of 24 features for every discharge, figure 3.10.

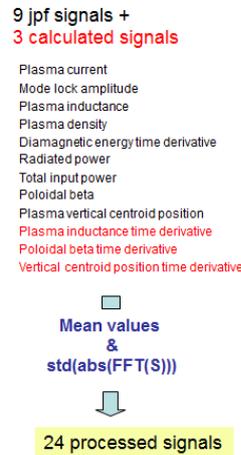


Figure 3.10: Signal processing overview to generate databases

3.2.3 Summary of ILW databases

The ILW campaigns $C28 - C34$ (September 2011 - October 2014) are compound from discharge 81867 until discharge 87918, it means 6051 shots. After doing the data analysis and signal processing, some discharges are deleted:

- Campaigns $C28 - C30$: discharges range 81852 – 83793
 - Due to sampling problems, 2 non-disruptive discharges are deleted.
 - After threshold analysis: 2 non-disruptive discharges, 2 intentional and 1 non-intentional disruptions are deleted due to exceed inductance threshold; and 1 non-disruptive shot due to exceed density threshold.
- Campaigns $C31 - C32$: discharges range 84442 – 85964
 - After threshold analysis: 4 non-disruptive discharges and 2 non-intentional disruptions are deleted due to exceed density threshold.
- Campaigns $C33$: discharges range 86453 – 87583
 - Due to sampling problems, 1 non-intentional disruption is deleted.
 - Due to low plasma current, 2 non-disruptive discharges are deleted.
 - After threshold analysis: 18 non-disruptive discharges are deleted due to exceed inductance threshold; and 1 non-disruptive shot due to exceed density threshold.

- Campaigns *C34*: discharges range 87606 – 87918
 - After threshold analysis: 2 non-disruptive discharges are deleted due to exceed inductance threshold.

The final databases yield:

| Campaign | (0) | (1) | (2) | Total |
|------------------|------------|------------|------------|--------------|
| <i>C28 – C30</i> | 1036 | 201 | 56 | 1293 |
| <i>C31 – C32</i> | 401 | 200 | 45 | 646 |
| <i>C33</i> | 491 | 151 | 106 | 748 |
| <i>C34</i> | 139 | 37 | 2 | 178 |

Table 3.7: Count of discharges from database if ILW campaigns *C28 – C34*. 0-Non-disruptive shot, 1-non-intentional disruption, 2-intentional disruption

Although the discharges from table 3.8 are not included in the final database (table 3.7), as previously mentioned, they are listed to have a strict control of every discharge and the reason it has been excluded. In addition it could be interesting to take this discharges in account for different studies.

| Campaign | Total |
|------------------|--------------|
| <i>C28 – C30</i> | 129 |
| <i>C31 – C32</i> | 86 |
| <i>C33</i> | 69 |
| <i>C34</i> | 10 |

Table 3.8: Count of discharges from ILW campaigns *C28 – C34* that cannot be considered safe neither disruptive

3.3 Assessment of the APODIS performance in ILW campaigns

An important task realised during this thesis has been the assessment of the predictor APODIS, explained in section 3.1, during ILW campaigns. The assessment during the first ILW campaigns *C28 – C30* was published in (Vega et al., 2013a). In this work, it is mentioned that high performance plasmas (with approximately plasma current 3.5 MA, input power 25 MW, electron temperature 1 keV and density $2 \times 10^{20} \text{ m}^{-3}$), as the non-disruptive discharge 83479 and the non-intentional disruption 83480, were identified correctly. The results show a success rate of 98.36% and false alarm rate of 0.92%. They are illustrated in figure 3.11, where Y-axis is the accumulative fraction of detected disruptions, X-axis is the warning time and the green line represents the rates with a warning time of 30 ms. It is important to note that this work considers good predictions any alarm triggered before the disruption time and after disruption time if there was more than 30 ms from the alarm time to the end of the plasma, that is, if there was enough time to carry out any mitigation action. In addition, during these campaigns the mitigation valves required at least 30 ms to carry out any action.

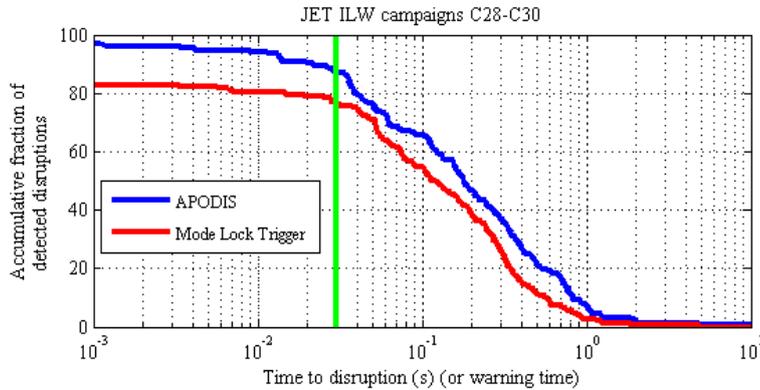


Figure 3.11: The logarithmic graphic shows the APODIS results during *C28 – C30* campaigns

The performance of APODIS during ILW campaigns *C31 – C34* is presented in (Moreno et al., 2015). Albeit, the same criteria has been used to present the performance during the ILW campaigns *C28 – C34*. During the ILW campaigns *C28 – C34* (September 2011 - October 2014), before showing APODIS performance, it is necessary to mention the discharges excluded and the reason. As mentioned, APODIS works using the signals showed in table 3.1, so if the EFCC (Error Field Corrections Coils) (Barlow et al., 2001) are activated, an anomalous behaviour appears on mode lock amplitude, which

follows the same pattern as the EFCC signal (it could be clearly seen in figure 3.8). These discharges are excluded from the results due to APODIS depend on this signal and it is not a natural evolution of the mode lock amplitude.

However, there are 218 discharges where EFCC have been activated and they have this effect on the mode lock amplitude. There are 148 discharges in campaign *C28 – C30* where EFCC does not affect, so they can be used. In campaign *C34* there is no discharge with EFCC activated, and the remaining 70 discharges are divided in: 35 safe shots without APODIS alarm; 17 safe discharges with APODIS alarm; 13 disruptive discharges where only 1 are missed; and 5 shots which can be saved because the EFCC affects when the plasma current ends.

Moreover all the discharges which do not follow a natural evolution (non-disruptive, intentional or non-intentional disruptive) are considered separately. It means all the discharges where any action from the control system has been taken or any failure on diagnoses which affects a signal used by the predictor. These anomalous discharges formed a total of 294 shots. Then, the count of the campaigns *C28 – C34* yields (table 3.9):

| Campaign | (0) | (1) | (2) | (3) | Total |
|------------------|------------|------------|------------|------------|--------------|
| <i>C28 – C30</i> | 713 | 177 | 35 | 129 | 1054 |
| <i>C31</i> | 319 | 151 | 14 | 72 | 556 |
| <i>C32</i> | 86 | 51 | 31 | 14 | 182 |
| <i>C33</i> | 513 | 151 | 106 | 69 | 839 |
| <i>C34</i> | 141 | 37 | 2 | 10 | 190 |
| <i>C28 – C34</i> | 1772 | 567 | 188 | 294 | 2821 |

Table 3.9: Count of discharges from ILW experimental campaigns *C28 – C34* for APODIS assessment. 0-Non-disruptive shot, 1-non-intentional disruption, 2-intentional disruption, 3-anomalous shot

Evaluation metrics play an important role in machine learning. They are used to assess and guide the learning algorithms. In the case of disruption prediction, which corresponds to unbalanced datasets, if a particular metric is chosen and it does not properly evaluate the minority class, then the learning algorithms will not be able to efficiently handle the unbalanced problem. A typical metric that is quite common in machine learning is the overall classification rate (i.e. accuracy). However, on an unbalanced dataset, the overall classification is no longer a suitable metric, since the small class has less effect on accuracy as compared with the prevalent class (Weiss and Provost, 2003). In our case, a multiobjective complex optimization problem is considered: the achievement of high success rates in the disruption predictions and simultaneously the highest reduction of the false alarm rates. Therefore, to present the results, the warning time is called the difference: $\Delta t = t_D - t_a$, where t_a is the time when the alarm is triggered, and t_D is

the disruption time. Taking into account that at least 10ms are needed to carry out mitigation action when a disruption happens, the results are going to be presented under the following criterion:

- Valid detections: $10\text{ ms} \leq \Delta t \leq 1.5\text{ s}$
- Early detections: $\Delta t > 1.5\text{ ms}$
- Tardy detections: $0 \leq \Delta t \leq 10\text{ ms}$
- False alarms: an alarm is triggered in a safe discharge.
- Missed alarms: in a disruptive discharge, the alarm happens after de disruption time ($\Delta t < 0$) or there is no alarm triggered.

Therefore the performance of APODIS during these campaigns yields (table 3.10):

| Campaign | Valid | Early | Tardy | False | Missed |
|------------------|------------------|----------------|----------------|-----------------|-----------------|
| <i>C31</i> | 81.46% (123/151) | 5.96% (9/151) | 3.31% (5/151) | 1.88% (6/319) | 9.27% (14/151) |
| <i>C32</i> | 84.31% (43/51) | 0% (0/51) | 0% (0/51) | 0% (0/86) | 15.69% (8/51) |
| <i>C33</i> | 74.17% (112/151) | 2.65% (4/151) | 3.31% (5/151) | 3.51% (18/513) | 19.87% (30/151) |
| <i>C34</i> | 86.49% (32/37) | 0% (13/37) | 0% (0/37) | 1.42% (2/141) | 13.51% (5/37) |
| <i>C28 – C34</i> | 79.37% (450/567) | 2.82% (16/567) | 2.47% (14/567) | 1.81% (32/1772) | 15.34% (87/567) |

Table 3.10: APODIS results during ILW campaigns *C28 – C34*

It is important to note that the average warning time of valid alarms is 350 ms and the standard deviation is 595 ms . If we compare with the mode lock trigger from the JPS (JET Protection System), we can see that APODIS keep high success rate and low false alarm rate taking into account that it is a system trained with data from C-wall campaigns and tested with data from a wall completely different. Considering the success rate the union of valid, early and tardy detections, the mode lock trigger reaches a success rate of $69.66\%(395/567)$ while APODIS obtains $84.66\%(480/567)$ and $1.81\%(32/1772)$ of false alarm rate. The false alarm rate cannot be considered for mode lock trigger because almost always it is used in close loop. In figure 3.12, the results previously mentioned are illustrated in a logarithmic graphic showing the accumulative fraction of detected disruptions and the warning time. The green lines show 10 ms and 1.5 s before the disruption time respectively.

Finally, APODIS has been used in close loop in some experiments during *C31 – C34* campaigns; it means that control system has carried out mitigation action after APODIS alarm. It has been used in a total of 443 discharges, where 147 are safe shots, 70 non-intentional and 5 intentional disruptions, 12 anomalous discharges and 209 outliers. In addition, from all the non-intentional disruptions where APODIS has not been taken in account by control system, it means all the non-intentional shots in table 3.9, the

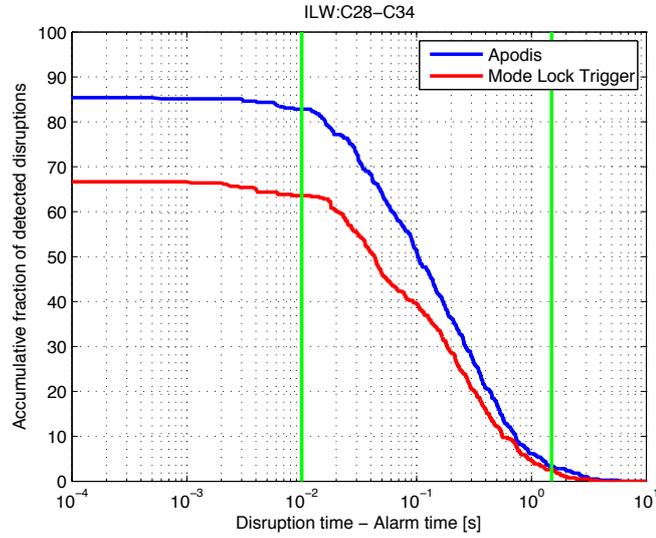


Figure 3.12: APODIS performance compare with mode lock trigger from JPS, during ILW campaigns *C28 – C34*

anticipation of APODIS alarm respect the trigger of the DMV (Disruption Mitigation Valve) has been analysed. In table 3.11 are shown the number of non-intentional disruptions where DMV has been triggered; the difference between the DMV triggered and APODIS alarm; and the mean value of these times. It can be seen that unless APODIS is not been used in close loop, it triggers an alarm before the DMV is opened in the 94.43% of the cases. The 5.57% of the discharges is compound of 8 shots without APODIS alarm; 1 discharge where APODIS has detected the disruption before disruption time; and 7 missed alarms.

| Campaign | DMV Triggers | $T = t_{DMV} - t_{APODIS} > 0$ | \bar{T} |
|------------------|--------------|--------------------------------|-----------|
| <i>C28 – C30</i> | 73/177 | 98.63% (72/73) | 972.5 ms |
| <i>C31</i> | 52/151 | 94.23% (49/52) | 374.6 ms |
| <i>C32</i> | 26/151 | 96.15% (25/26) | 262 ms |
| <i>C33</i> | 132/151 | 91.67% (121/132) | 273.9 ms |
| <i>C34</i> | 4/37 | 100% (4/4) | 4.45 s |
| <i>C28 – C34</i> | 287/567 | 94.43% (271/287) | 1.267 s |

Table 3.11: Relation between DMV triggers and APODIS alarms in non-intentional disruptions without APODIS in close loop

3.4 Robustness of APODIS

As previously explained, disruption prediction with enough warning time is necessary in order to carry out mitigation actions. This section explains the APODIS robustness analysis (Moreno et al., 2014b). Robustness is defined as the predictor reliability when a signal fails. The reliability has to be understood in terms of the success rate and the false alarm rate. In this study, the success rate is defined as the fraction of disruptive discharges that have been predicted with enough anticipation time. Enough anticipation time means to trigger an alarm well in advance to be able to put in operation mitigation actions. When this study was done, the minimum time in JET to activate the disruption mitigation valve was about 30 ms. Therefore, only predictions whose anticipation times are greater than or equal to 30 ms will be considered a success. However, it is important to mention that disruptions predicted with less than 30 ms cannot be considered missed alarms. Therefore, disruptions with anticipation times between 1 and 30 ms have been considered tardy detections. To accomplish the robustness analysis, simulations have been performed. These simulations replace one by one the signals of table 3.1 (except the plasma current) for a synthetic one. Two different scenarios have been considered. In the first one, a synthetic signal with mean value 0 and Gaussian noise is used during the whole discharge. This simulation can correspond to a situation in which some kind of instrumentation (power supplies, amplifiers, analogue to digital converters, data transmission lines or so) fails without notice. In the second scenario, a different situation is generated. The test simulates the failure of a signal from a certain time of a discharge and the amplitude remains the same but with a Gaussian noise added.

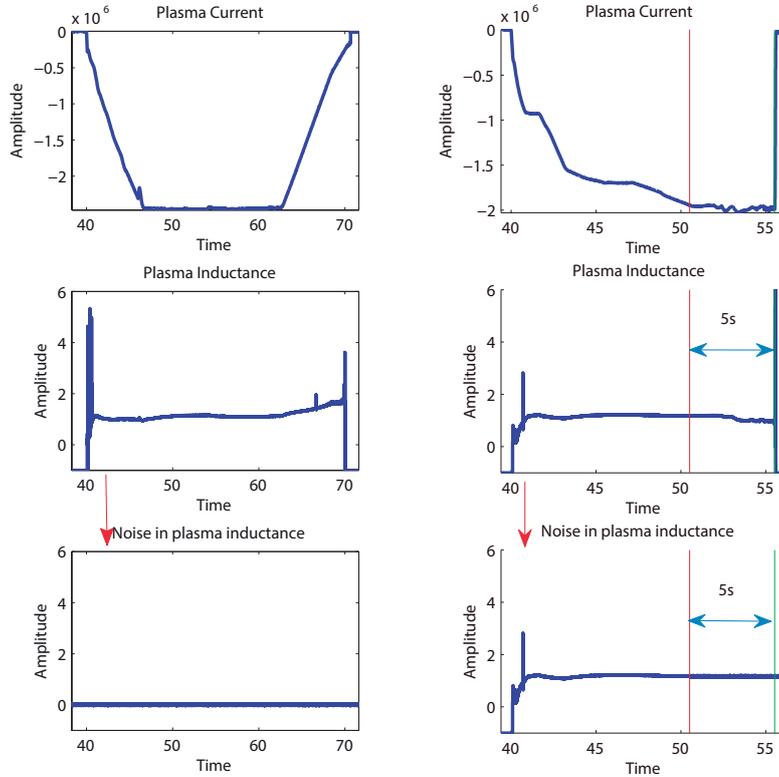
The purpose of the reliability analysis is not to select the best signals for APODIS (Rattá et al., 2012), but to determine the changes in the success and false alarm rates when a signal is in failure. As mentioned, the simulations replace each signal for a synthetic one. The plasma current has not been substituted because it is the signal that switches on/off the APODIS predictor when it crosses a threshold of 750 kA. It should be noted that the feature vectors have dimension 14 and, therefore, the simulation of a signal in failure means to have two wrong features in each predictor input. This is a consequence of using both time and frequency domains for each signal, as explained in section 3.2. To carry out the simulations, an off-line APODIS version and a huge database of discharges have been used (from both CFC and ILW campaigns). Firstly, 3578 non-disruptive discharges and 228 unintentional disruptions, from CFC campaigns C23-C27b (September 2008 - October 2009), have been considered. Then, 1036 non-disruptive discharges and 201 unintentional disruptions from ILW campaigns C28-C30 (July 2013 - October 2014) have been analysed (table 3.2 and 3.7). As mentioned, the

robustness analysis performs two different simulations, where the difference resides on the different amplitudes that use the synthetic signals:

1. The first simulation replaces the whole signal for a Gaussian noise distribution $N(0, 1)$, where the notation $N(\mu, \sigma)$ is the usual one to represent a Gaussian distribution with mean μ and standard deviation σ . Figure 3.13a shows the replacement of the whole plasma inductance signal with a Gaussian noise distribution $N(0, 1)$ in a non-disruptive discharge.
2. The second simulation changes the latest 5 s of the signals by a Gaussian noise distribution $N(0, 1)$. The latest 5 s are selected in different ways for non-disruptive and disruptive discharges. For non-disruptive discharges, they correspond to the latest 5 s before the plasma current crosses the 750 kA threshold in the ramp down. In the case of disruptive discharges, the 5 s interval coincides with the previous 5 s before the disruption. Figure 3.13a illustrates the latest 5 s of the plasma inductance signal are replaced with a Gaussian noise distribution $N(0, 1)$ in a disruptive discharge.

The results (table 3.12 for CFC campaigns and table 3.13 for ILW campaigns) show the same statistics in both simulations (figures 3.13a and 3.13b), it means, either replacing the whole signal for a Gaussian noise synthetic signal or replacing the signal at a certain time for a Gaussian noise synthetic signal. The worst results appear when the mode lock and the plasma inductance signals are in failure (figures 3.14a and 3.14b respectively). Therefore, these signals are absolutely necessary for APODIS. If either the mode lock or the plasma inductance signal is in failure, APODIS produces wrong predictions. It should be emphasized that the results of tables 3.12 and 3.13 in relation to the simulations with all signals in the ILW discharges do not coincide with the rates given at the end of this section. The APODIS real-time software is in execution in the JET real-time network from the discharge 82429. However, the off-line simulation starts in shot 81852.

Then, the other signals, plasma density, diamagnetic energy time derivative, radiated power and total input power, are also important but have less impact in the prediction than the mode lock and the plasma inductance. Whereas the mode lock and the plasma inductance signals are essential, failures in the remaining signals slightly decrease the success rate. For example, figure 3.15 shows the case of the diamagnetic energy time derivative in which the success rate remains around 65% (instead of 69%) in the CFC campaigns and 73% (instead of 77%) in ILW discharges. Also, the AWT (average warning time) has been analysed. Table 3.14 shows the AWT obtained in CFC and ILW campaigns by simulating both the whole signal in failure and the latest 5 s of the signal in failure. It is important to note in this table that

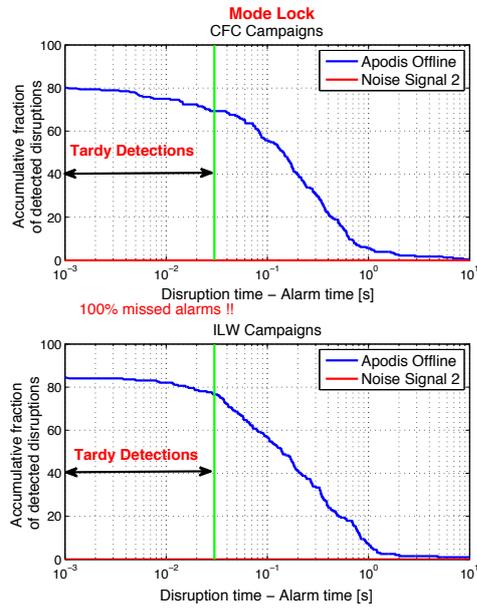


(a) Simulation 1: replacement of the whole signal. (b) Simulation 2: replacement of the last 5s.

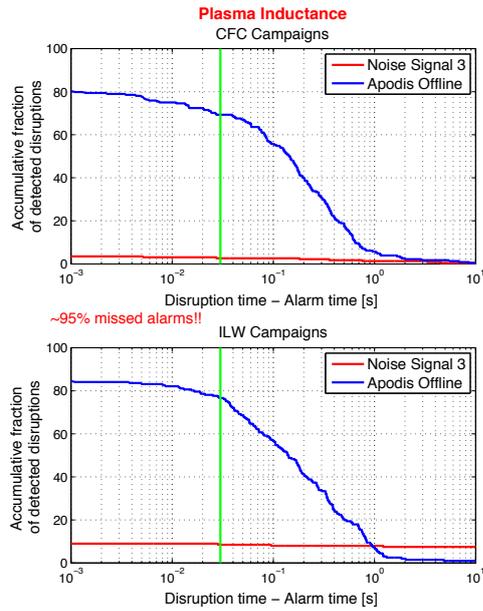
Figure 3.13: Example of both simulations carried out in the robustness analysis

the off-line analysis with all the signals shows an AWT value of 491 ms (with a standard deviation of 1.320 s). The simulation of failure in the mode lock and in the plasma inductance gives very bad success rate and, therefore, in these cases the computation of the AWT does not make sense. Focusing the attention on the rest of the signals, the average AWT corresponding to failures in plasma density, FWDIA (stored diamagnetic energy time derivative), radiated power and total input power in CFC campaigns is 489 ms (with a mean standard deviation of 1.093 s). With regard to the simulations corresponding to the ILW campaigns, the average AWT is 561 ms and the mean standard deviation is 1.695 s.

Finally the robustness analysis shows that the mode lock and the plasma inductance signals are essential for APODIS. Any failure in these signals



(a) Failure on mode lock amplitude



(b) Failure on plasma inductance

Figure 3.14: Results for CFC and ILW campaigns with mode lock and plasma inductance signals in failure

produces very low success rates, (0% for the mode lock signal failure and

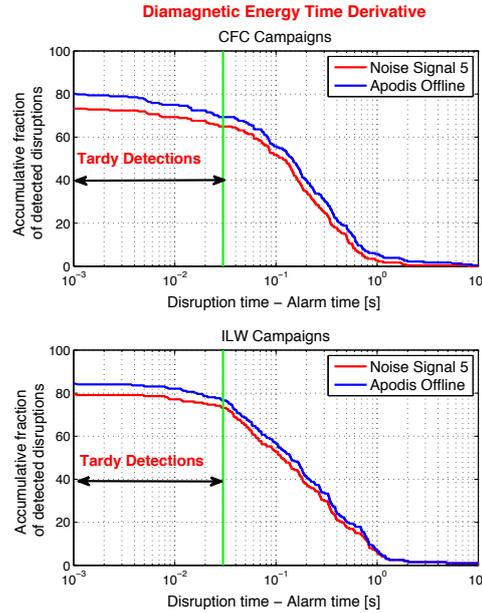


Figure 3.15: Results for CFC and ILW campaigns with diamagnetic energy time derivative signal failure

| CFC campaigns 2008 – 2010 | | | | |
|---|-----------------|------------------|-------------------|------------------|
| Signals | Success Rate | Tardy alarm Rate | Missed alarm Rate | False alarm Rate |
| Replacement of the whole signal (figure 3.13a) | | | | |
| Offline analysis with all signals | 69.3% (158/228) | 11.4% (26/228) | 19.3% (44/228) | 7.5% (267/3578) |
| Mode lock failure | 0% (0/228) | 0% (0/228) | 100% (228/228) | 0% (0/3578) |
| Plasma inductance failure | 1.3% (3/228) | 1.3% (6/228) | 97.4% (222/228) | 0.3% (11/3578) |
| Plasma density failure | 68.4% (156/228) | 11.4% (26/228) | 20.2% (46/228) | 6.7% (239/3578) |
| FWDIA failure | 64.9% (148/228) | 8.8% (20/228) | 26.3% (60/228) | 4.1% (147/3578) |
| Radiated power failure | 69.3% (158/228) | 11.4% (26/228) | 19.3% (44/228) | 7.4% (266/3578) |
| Total input power failure | 68.0% (155/228) | 11.0% (25/228) | 21.0% (48/228) | 7.1% (254/3578) |
| Replacement of the last 5s of the signal (figure 3.13b) | | | | |
| Offline analysis with all signals | 69.3% (158/228) | 11.4% (26/228) | 19.3% (44/228) | 7.5% (267/3578) |
| Mode lock failure | 0% (0/228) | 0% (0/228) | 100% (228/228) | 0% (0/3578) |
| Plasma inductance failure | 2.6% (6/228) | 0.9% (2/228) | 96.5% (220/228) | 0.6% (20/3578) |
| Plasma density failure | 68.4% (156/228) | 11.4% (26/228) | 20.2% (46/228) | 6.7% (239/3578) |
| FWDIA failure | 64.9% (148/228) | 8.8% (20/228) | 26.3% (60/228) | 4.1% (147/3578) |
| Radiated power failure | 69.3% (158/228) | 11.4% (26/228) | 19.3% (44/228) | 7.4% (266/3578) |
| Total input power failure | 68.0% (155/228) | 11.0% (25/228) | 21.0% (48/228) | 7.1% (254/3578) |

Table 3.12: CFC: *C23 – C27b* robustness result. Both types of signal failure provide similar outcomes

5 – 10% for the plasma inductance signal failure). The other signals (plasma density, diamagnetic energy time derivative, radiated power and total input power) are also important but have a less impact. Failures in them produce success rates around 75%. The simulations that add Gaussian noise provide similar results. Focusing on the radiated power signal, it can be observed in tables 3.12 and 3.13 that its failure provides the same success rate and lower

| ILW campaigns 2011 – 2012 | | | | |
|---|-----------------|------------------|-------------------|------------------|
| Signals | Success Rate | Tardy alarm Rate | Missed alarm Rate | False alarm Rate |
| Replacement of the whole signal (figure 3.13a) | | | | |
| Offline analysis with all signals | 76.6% (154/201) | 8.5% (17/201) | 14.9% (30/201) | 2.2% (23/1036) |
| Mode lock failure | 0% (0/201) | 0% (0/201) | 100% (201/201) | 0% (0/1036) |
| Plasma inductance failure | 3.5% (7/201) | 1.0% (2/201) | 95.5% (192/201) | 0% (0/1036) |
| Plasma density failure | 72.6% (146/201) | 9.5% (19/201) | 17.9% (36/201) | 1.5% (16/1036) |
| FWDIA failure | 73.1% (147/201) | 7.0% (14/201) | 19.9% (40/201) | 1.5% (15/1036) |
| Radiated power failure | 75.6% (152/201) | 9.0% (18/201) | 15.4% (31/201) | 2.0% (21/1036) |
| Total input power failure | 78.1% (157/201) | 8.5% (17/201) | 13.4% (27/201) | 3.0% (31/1036) |
| Replacement of the last 5s of the signal (figure 3.13b) | | | | |
| Offline analysis with all signals | 76.6% (154/201) | 8.5% (17/201) | 14.9% (30/201) | 2.2% (23/1036) |
| Mode lock failure | 0% (0/201) | 0% (0/201) | 0% (0/201) | 0% (0/1036) |
| Plasma inductance failure | 8.5% (17/201) | 0.5% (1/201) | 91.0% (183/201) | 6.3% (65/1036) |
| Plasma density failure | 72.6% (146/201) | 9.5% (19/201) | 17.9% (36/201) | 1.5% (16/1036) |
| FWDIA failure | 73.1% (147/201) | 7.0% (14/201) | 19.9% (40/201) | 1.5% (15/1036) |
| Radiated power failure | 75.6% (152/201) | 9.0% (18/201) | 15.4% (31/201) | 2.0% (21/1036) |
| Total input power failure | 78.1% (157/201) | 8.5% (17/201) | 13.4% (27/201) | 3.0% (31/1036) |

Table 3.13: ILW: C28 – C30 robustness result

| Substitution for synthetic signal | | |
|-----------------------------------|----------|---------|
| CFC campaigns 2008 – 2010 | | |
| Signals | AWT (ms) | STD (s) |
| Offline analysis with all signals | 491 | 1.320 |
| Mode lock failure | - | - |
| Plasma inductance failure | - | - |
| Plasma density failure | 531 | 1.342 |
| FWDIA failure | 371 | 0.707 |
| Radiated power failure | 590 | 1.440 |
| Total input power failure | 463 | 0.884 |
| ILW campaigns 2011 – 2012 | | |
| Offline analysis with all signals | 573 | 1.667 |
| Mode lock failure | - | - |
| Plasma inductance failure | - | - |
| Plasma density failure | 515 | 1.703 |
| FWDIA failure | 561 | 1.700 |
| Radiated power failure | 561 | 1.678 |
| Total input power failure | 608 | 1.700 |

Table 3.14: AWTs and standard deviation (STD) for CFC and ILW campaigns results

false alarm rate than the rest of the signals. It is interpreted as a specific result on the range of pulses used in the analysis, and more analysis with the incoming campaigns should be done to better understand this behaviour. Anyway it could indicate that this signal for these experiments provides less characteristic information for the predictor even false information. A potential alternative to use the APODIS model with one signal in failure is to have trained different models with only six quantities (it is assumed that

is in failure either the plasma density or the FWDIA or the radiated power or the total input power). This possibility has not yet been explored because the first objective was to know the limitations of the present APODIS predictor.

3.5 Increased time resolution of APODIS

During the ILW campaigns, APODIS has missed some disruptions due to the lack of time resolution (Vega et al., 2013a). The predictor is enabled whenever the plasma current is above the threshold of 750 kA. Figure 3.16 shows a typical example of missed alarm as a consequence of a prediction period of 32 ms. Point A in figure 3.16 indicates the last prediction of APODIS for the given discharge. The prediction at that time is “non-disruptive”. During the next 32 ms, a disruption takes place and the plasma current crosses the threshold of 750 kA. Therefore, APODIS is disabled and no more predictions are carried out. Consequently, the alarm is missed. The effect of increasing the APODIS temporal resolution is analysed in this section.

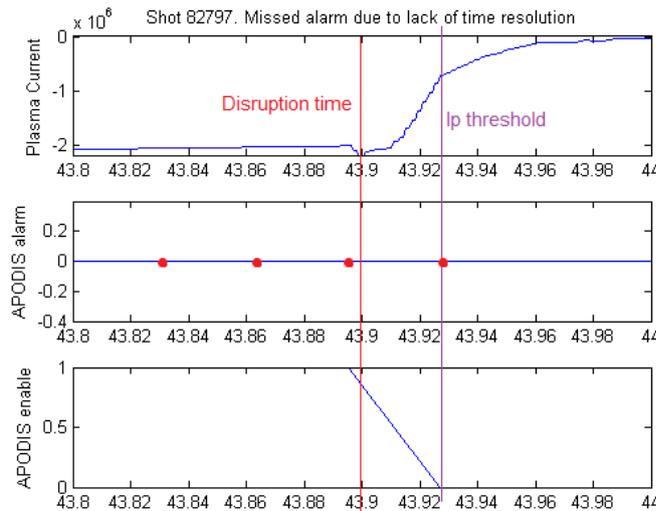


Figure 3.16: Example of missed alarm due to a lack of resolution

Following it will be explained the database used for these purposes and described the simulations performed to increase the time resolution by using a sliding window mechanism. As it was explained previously, APODIS makes a prediction every 32 ms. This prediction period should be shortened as much as possible in order to detect any disruption signature as soon as possible. (López et al., 2012) shows that the computation time in the JET real-time network to form a feature vector is hundreds of μ s. Therefore, a time resolution of 1 ms can be achieved. The objective in the present

analysis is not to change the sampling rate of the APODIS signals (which remain limited to $1 \text{ ksamples s}^{-1}$), but to change the time resolution of the prediction by implementing a sliding window mechanism. This mechanism allows APODIS implementing different time resolutions (16, 8, 4, 2 or 1 ms). In this way, APODIS is able to trigger an alarm every 16, 8, 4, 2 or 1 ms instead of the current temporal resolution of 32 ms (figure 3.17).

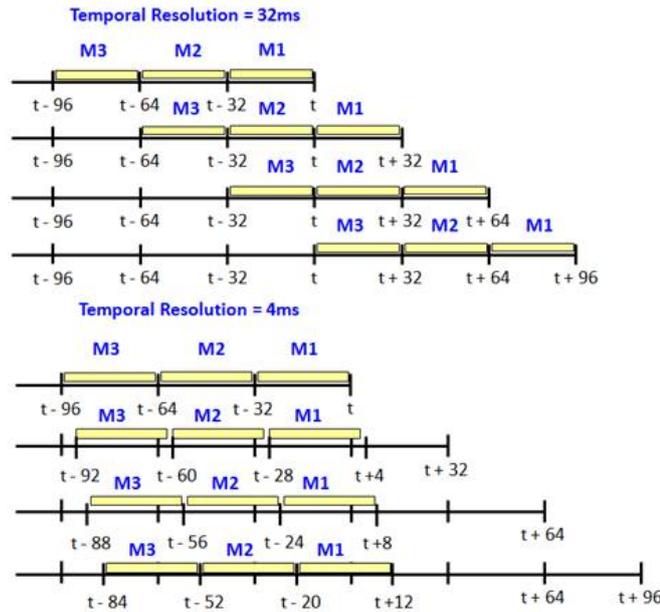


Figure 3.17: Example of temporal resolution of 32 ms (how APODIS works currently at JET) and temporal resolution of 4 ms

The database used in this analysis is made up of discharges from the three first ILW campaigns *C28 – C30*. It is the same database used in section 3.4 for the ILW campaigns (table 3.7), which is made up of 1036 non-disruptive discharges and 201 unintentional disruptions. At this point, it will be shown the results obtained from the simulations. The results (table 3.15 and figure 3.18) show how the success rate increases for higher temporal resolutions, reaching a success rate of 83% for 1 ms of temporal resolution. Furthermore, higher temporal resolutions allow achieving better success rate and also reducing the missed alarm rates (table 3.15 and figure 3.18). On the other hand, while the success rate is increased for higher temporal resolutions, the false alarm rate also increases (table 3.15).

Sliding window analysis shows that higher temporal resolutions can help to achieve better success rates, reaching 83% success rate and 8.5% tardy alarm rate for 1 ms resolutions. Despite of this, the false alarm rate slightly increases for higher temporal resolutions. Focusing on false alarm rate in ta-

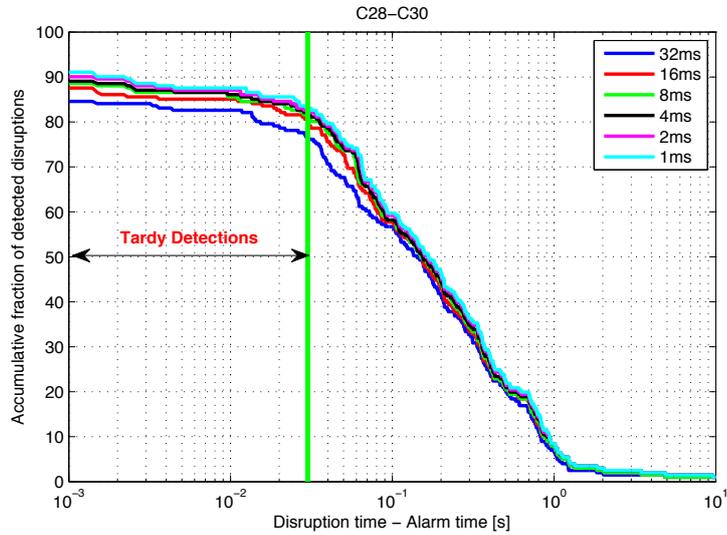


Figure 3.18: Sliding window results

| Temporal resolution (ms) | Success Rate | Tardy alarm Rate | Missed alarm Rate | False alarm Rate |
|--------------------------|-----------------|------------------|-------------------|------------------|
| 32 | 76.6% (154/201) | 8.0% (16/201) | 15.4% (31/201) | 2.9% (30/1036) |
| 16 | 80.6% (162/201) | 7.0% (14/201) | 12.4% (25/201) | 4.4% (46/1036) |
| 8 | 81.6% (164/201) | 7.5% (15/201) | 10.9% (22/201) | 5.0% (52/1036) |
| 4 | 82.1% (165/201) | 7.5% (15/201) | 10.4% (21/201) | 5.6% (58/1036) |
| 2 | 82.5% (166/201) | 8.5% (17/201) | 9.0% (18/201) | 6.0% (62/1036) |
| 1 | 83.0% (167/201) | 8.5% (17/201) | 8.5% (17/201) | 6.0% (62/1036) |

Table 3.15: Sliding window rates

ble 3.15, 32 ms and 1ms resolutions show 2.9% and 5.98% of false alarm rate respectively, which means an important difference. Higher time resolutions provide better warning times, so if the main problem to overcome disruptions is to have enough time to carry out mitigation actions, we can conclude that increasing the temporal resolution of APODIS would be a good improvement in terms of both success rate and warning times.

Chapter 4

Automatic location of disruption times in JET

It was introduced in section 1.2.1 that the stability and confinement of the plasma can be lost in a few hundred microseconds. The electromagnetic forces and thermal loads produced by disruptions can damage the components of the devices. Disruptions are difficult to understand from a theoretical point of view due to event complexity, highly non-linear interactions, and diversity of causes. Furthermore, many different behaviours and current quench scenarios are possible in disruptions. In the past JET C-wall campaigns, disruptions were identified as the fast decay of the plasma current (typically called “current quench”) produced by the increase in plasma resistivity that thermal quench cooling generated (Wesson and Campbell, 1987; Schuller, 1995) (figure 4.1 (Wesson and Campbell, 1987)).

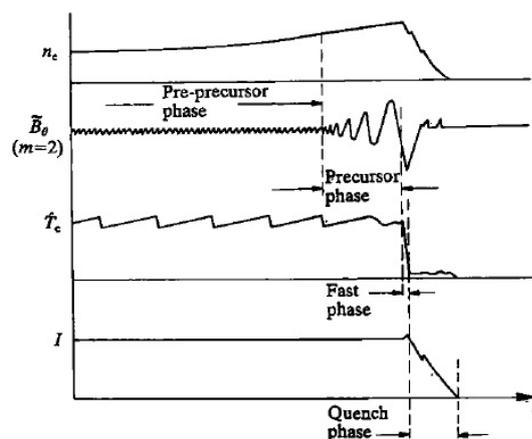


Figure 4.1: Example of a disruption during campaigns with the JET-C wall

However, disruptions at experiments with the JET-ILW show a different behaviour. It is not unusual for these plasmas to present several thermal quenches previous and after the current quench. In many instances the temperature recovers and the plasma survives until the next thermal quench, in figure 4.2 it can be seen several thermal quenches and how the temperature recovers. Therefore, the criterion used during the JET-C experiments is usually not valid and therefore, a new one has to be defined. A possibility is to use the start of the current quench as disruption time, since around this time the plasma loses completely its stability, becomes uncontrollable, and ends abruptly.

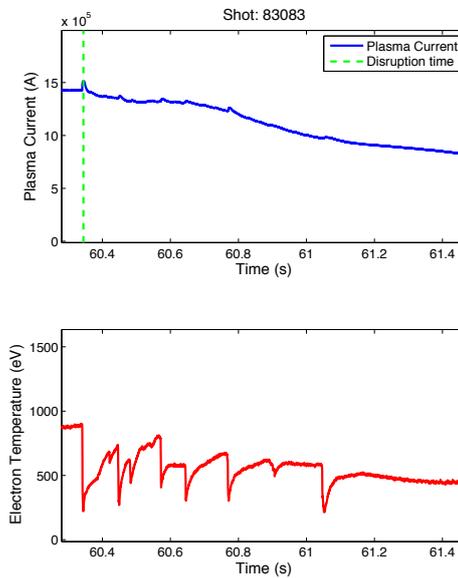


Figure 4.2: Disruption from JET-ILW experiments

In addition, as it was explained in section 3.2, once all the discharges are classified, the disruption time has to be obtained. This is the most important step because all the disruptive information is around this time, hence the most precise is the time determined, the best information we could extract as disruptive samples. This time is usually determined by hand, but automatic tools have been developed in this thesis to make easier this work (Moreno et al., 2014a). The purpose in this chapter is to provide an analysis tool to locate the disruption time defining a new criterion.

The first section 4.1 exposes previous studies carried out to locate the disruption time. In section 4.2, it is explained the wavelet analysis applied to this problem.

4.1 Previous studies

Several methods have been employed to detect disruption times before the use of wavelets achieving great results. The database used for every work in this chapter is formed by discharges from the ILW experimental campaigns *C28 – C30*, see table 3.7. Firstly, in section 4.1.1 the problem has been tackled as a classification problem, where SVM models have been trained using different signals. After that, in section 4.1.3, Universal Multi-Event Locator (UMEL), which has been explained in section 2.4.2, is used to detect the disruption time. Then another technique known as martingales, explained in section 2.5, is considered to this problem in section 4.1.2.

4.1.1 Classifier

Considering the location of disruption time as a classification problem to detect only the disruption time, two set of samples are differentiated: safe samples and disruptive samples. The idea is similar as the procedure to develop predictors, however, the base is considering as disruptive sample only the last one before the disruption time in order to trigger an alarm as close as possible to the disruption time. At this point, we consider the disruptive discharge is known and the purpose is only to determine the disruption time, therefore false alarms on safe discharge are not a problem because it will be analysed only disruptive discharges. From the ILW campaigns *C28 – C30*, it has been taken the first 50 disruptive discharges and 510 non-disruptive shots for the training, and 151 disruptive discharges for the test. The features used in this classifier are shown in table 4.1, where mean is the mean value of the signal amplitude and FFT is the standard deviation of the Fast Fourier Transform. In figure 4.3 illustrates the five signals used: plasma current, mode lock amplitude, plasma internal inductance, loop voltage and peaking factor. These signals have been considered because they show a distinguishable behaviour close to the disruption, as it can be seen in figure 4.3.

The data are pre-processed similarly to the method used for the databases in section 3.2:

1. Signals are downloaded from JET database.
2. The upper and lower bounds are determined for the signals. As lower bound is the threshold of 750 kA on the plasma current, and the upper bound is the disruption time for disruptive discharges or the time when the plasma current crosses 50 kA amplitude for safe shots.
3. The data is sampled to 1 kHz of frequency and normalized to the values showed in 3.6.
4. Finally the features in table 4.1 are calculated.

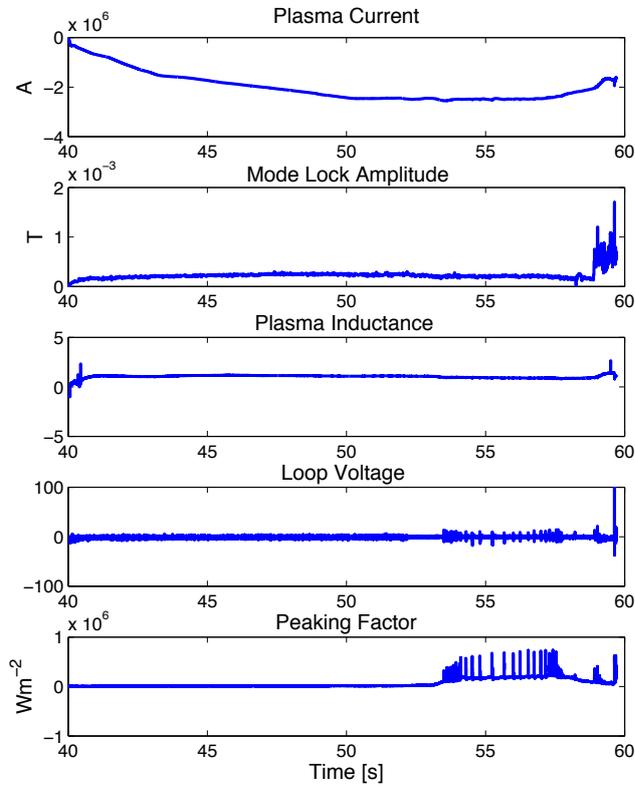


Figure 4.3: Signals used on disruption time location as a classification approach

| Signals | Feature |
|----------------------------|--------------|
| Plasma current | FFT |
| Mode lock amplitude | mean and FFT |
| Plasma internal inductance | mean and FFT |
| Loop voltage | mean |
| Peaking factor | mean |

Table 4.1: Features for disruption time location as a classification approach

The results of this approach shows that only the 19.33% of the disruptions are detected in a temporal window of 32 ms around the disruption time. This bad results are due to the classifier works better as predictor than detector, so that this approach were discarded.

4.1.2 Martingales

Following the methodology explained in section 2.5, the location of disruption time is focused as a problem of change detection in data streams. The simulations will follow the diagram in figure 4.4.

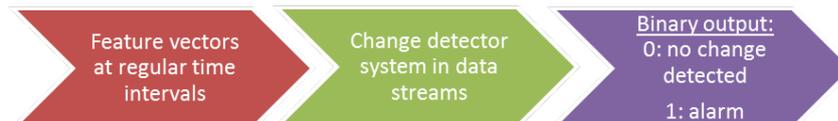


Figure 4.4: Diagram of martingale for change detection in data streams

Input samples are sequential feature vectors that correspond to a non-perturbed plasma state. At a certain point, feature vectors from a perturbed emission begin to reach the detector system. After some delay in relation to the first perturbed feature vector, the output triggers an alarm. This detection system analysis of the output has to be able to identify (with as small error as possible) the time instant in which the change of plasma state happened.

This analysis has been carried out following the data pre-processing explained in section 3.2, using discharges from campaigns *C28 – C30* (3.7). Three different signal sets have been used:

- Plasma current and loop voltage (S1).
- Peaking factor and plasma density (S2).
- Temporal evolution of the plasma current time derivative (S3).

Firstly, three different martingales were tested: RPM, SMM, PIM (see section 2.5). Changing the different parameters, it was observed that the RPM was the most efficient martingale to develop the analysis. Better results could be reached with an optimal selection of the parameters for PIM, however it was discarded because of the high computational cost and low improvement of the results. It is important to note that only the last 1.5 – 3 s of the discharge are considered in this analysis, due to the martingale test is extremely sensitive to plasma changes. Therefore, considering this is an offline analysis and its purpose is determining the exact time of the disruption, there is no inconvenient on taking the last seconds of the discharge. Furthermore. this way the analysis will be faster. Let see and example, given the plasma density (signal 1) and peaking factor (signal 2) from the

discharge 81867, the temporal evolution and the standard deviation of the FFT are illustrated in figure 4.5. The green line shows the disruption time. The standard deviation of the FFT of both signals are taken as input data on martingale test, the scatter plot of the features can be seen in figure 4.6.

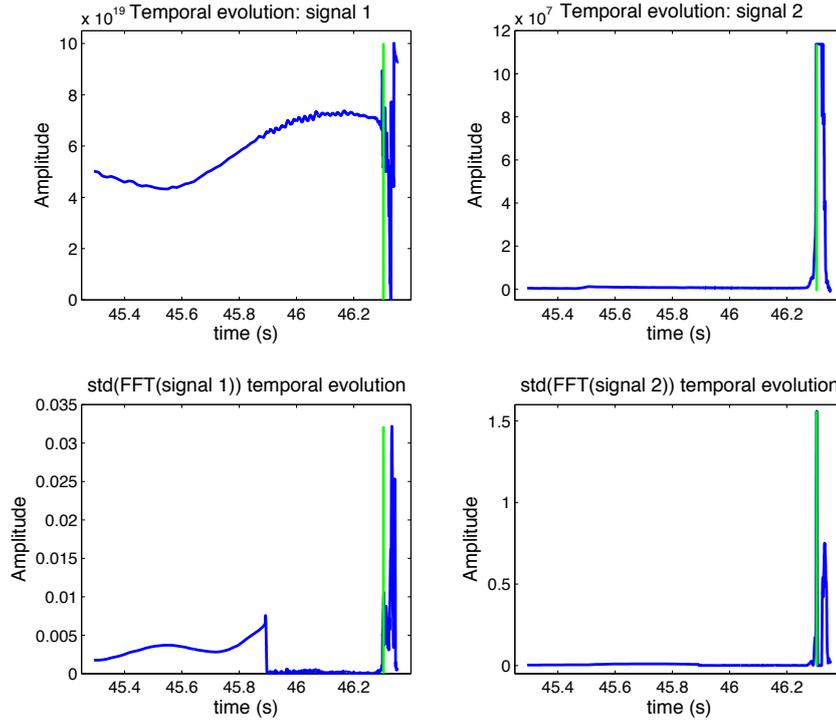


Figure 4.5: Density and Peaking Factor in discharge 81867

Following the MTA for the different martingales (RPM, SMM, PIM), the best parameters to fit the input data have been selected. Figure 4.7 shows the results for each martingale with their respective parameters. It can be seen that SMM (figure 4.7b) is extremely sensitive and many changes are detected. On the other hand, RPM (figure 4.7a) and PIM (figure 4.7c) detect a change close to the disruption time. The time where p -values distribution is not uniform on $[0, 1]$, the change happens but the alarm is triggered with a delay, when the value λ reaches the threshold. We are going to determine the disruption time as the sample where λ reaches the threshold selected. Despite of this assumption is theoretically incorrect, this approach provides better times in our problem and it allows detecting automatically the change. The case of RPM shows more uniform distribution of the p -values before the change and it is much faster than PIM, therefore it has been decided to use

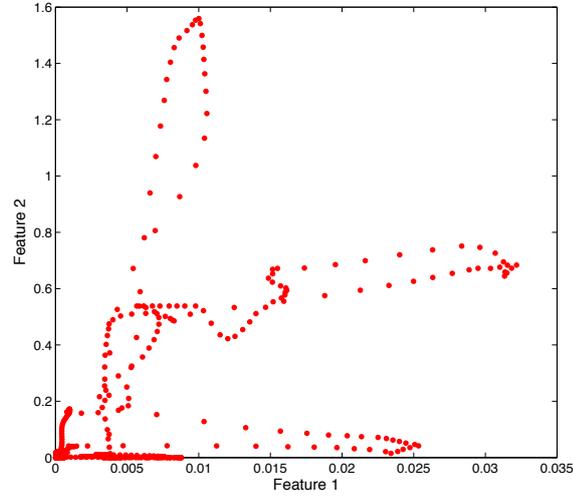
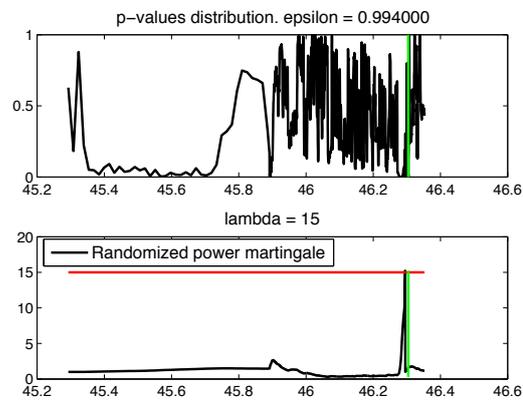


Figure 4.6: Scatter plot of the Density and Peaking Factor in discharge 81867

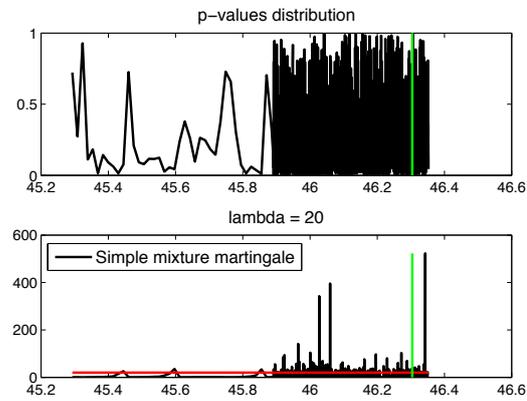
the RPM for the rest of the analysis.

Many values of λ and ϵ parameters have been tested for each set of signals. In tables 4.2 and 4.3, the results, for sets S1 and S2 respectively, are shown, with $\lambda = 10$ and selecting the best ϵ in a range $[0.990, 0.999]$ for each discharge. It can be seen in table 4.2 that S1 recognizes a change in the 20 discharges analysed, the average window around the disruption time is 113 ms. Table 4.3 shows that S2 detects a change in 19 discharges, and three changes are detected in one discharge. The average window around the disruption time is 124 ms. S1 obtains better results than S2, however the location of disruption time is not good enough, the time determined by the martingale needs to be more similar to the real disruption time.

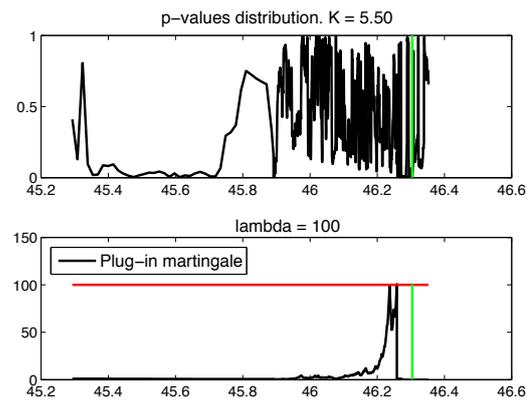
Considering the set S1, in order to smooth the signals and eliminate possible premature changes, a wavelet Haar was applied to both signals. Different levels of the wavelet were tested and the best results were obtained to level 2. In figure 4.8 the results using a wavelet Haar decomposition of level 2 in S1 are shown. It can be seen that a change is detected in 60 discharges of 72, and taking a window of 48 ms around disruption time, only the 51.39% (37/72) are detected. As an alternative, instead considering the change when the value of λ reached the threshold, it was considered a change if the value of λ reached a certain percentage of the threshold selected. However, the results were not improved.



(a) RPM



(b) SMM



(c) PIM

Figure 4.7: Martingale change detection to locate disruption time

| Shot | T_{dis} | ϵ | $T_{detected}$ | $T_{dis}-T_{detected}$ |
|-------|-----------|------------|----------------|------------------------|
| 81867 | 46.304 | 0.994 | 46.2962 | 0.0078 |
| 81916 | 51.706 | 0.997 | 51.7006 | 0.0054 |
| 81921 | 59.532 | 0.999 | 59.4976 | 0.0344 |
| 81922 | 55.5198 | 0.998 | 55.4996 | 0.0202 |
| 81923 | 55.514 | 0.998 | 54.8958 | 0.6182 |
| 81924 | 55.524 | 0.998 | 55.5028 | 0.0212 |
| 81925 | 59.518 | 0.997 | 59.5006 | 0.0174 |
| 81926 | 60.5032 | 0.999 | 60.464 | 0.0392 |
| 81927 | 60.5172 | 0.998 | 60.4948 | 0.0224 |
| 81928 | 60.5302 | 0.999 | 60.5014 | 0.0288 |
| 81929 | 54.366 | 0.998 | 54.3606 | 0.0054 |
| 81930 | 57.0754 | 0.997 | 57.0456 | 0.0298 |
| 81953 | 59.762 | 0.992 | 59.7624 | -0.0004 |
| 81978 | 53.7788 | 0.995 | 53.7818 | -0.003 |
| 81983 | 59.452 | 0.999 | 58.1958 | 1.2562 |
| 81985 | 52.27 | 0.996 | 52.1902 | 0.0798 |
| 81990 | 52.374 | 0.997 | 52.3392 | 0.0348 |
| 82001 | 53.362 | 0.997 | 53.3382 | 0.0238 |
| 82004 | 48.788 | 0.997 | 48.7618 | 0.0262 |
| 82005 | 53.16 | 0.997 | 53.1634 | -0.0034 |

Table 4.2: RPM-S1 results

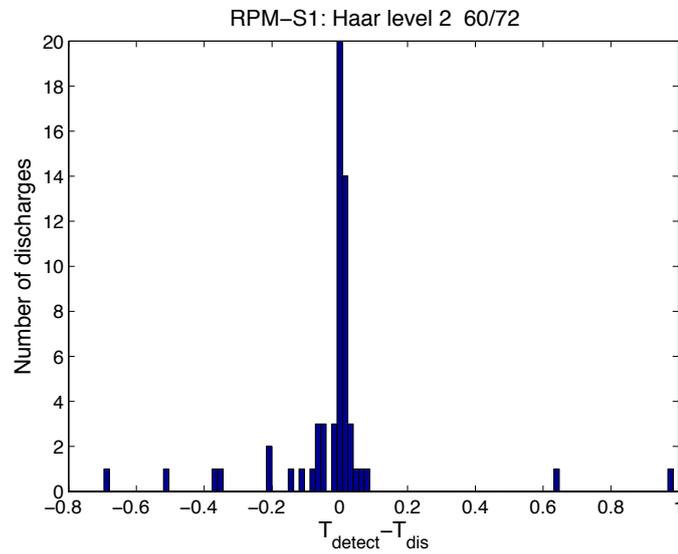


Figure 4.8: Histogram of RPM-S1 results with wavelet Haar level 2

| Shot | T_{dis} | ϵ | $T_{detected}$ | $T_{dis}-T_{detected}$ |
|-------|-----------|-----------------|----------------|------------------------|
| 81867 | 46.304 | 0.997 | 46.29 | 0.014 |
| 81916 | 51.706 | 0.995 | 51.7032 | 0.0028 |
| 81921 | 59.532 | 0.995 | 59.5262 | 0.0058 |
| 81922 | 55.5198 | 0.998 | 55.5106 | 0.0092 |
| 81923 | 55.514 | 0.996 | 55.5344 | -0.0204 |
| 81924 | 55.524 | — | — | — |
| 81925 | 59.518 | 0.998 | 59.511 | 0.007 |
| 81926 | 60.5032 | 0.998 | 60.4604 | 0.0428 |
| 81927 | 60.5172 | 0.998 | 60.4406 | 0.0766 |
| 81928 | 60.5302 | 0.997 | 60.526 | 0.0042 |
| 81929 | 54.366 | 0.998 | 54.3632 | 0.0028 |
| 81930 | 57.0754 | 0.998 | 56.9276 | 0.1478 |
| 81953 | 59.762 | 0.994 | 59.8018 | -0.0398 |
| 81978 | 53.7788 | 0.995 | 53.6558 | 0.123 |
| 81983 | 59.452 | 0.993(3changes) | 58.0464 | 1.4056 |
| 81985 | 52.27 | 0.997 | 52.0904 | 0.1796 |
| 81990 | 52.374 | 0.998 | 52.2424 | 0.1316 |
| 82001 | 53.362 | 0.999 | 53.11 | 0.252 |
| 82004 | 48.788 | 0.996 | 48.7994 | -0.0114 |
| 82005 | 53.16 | 0.999 | 53.1312 | 0.0288 |

Table 4.3: RPM-S2 results

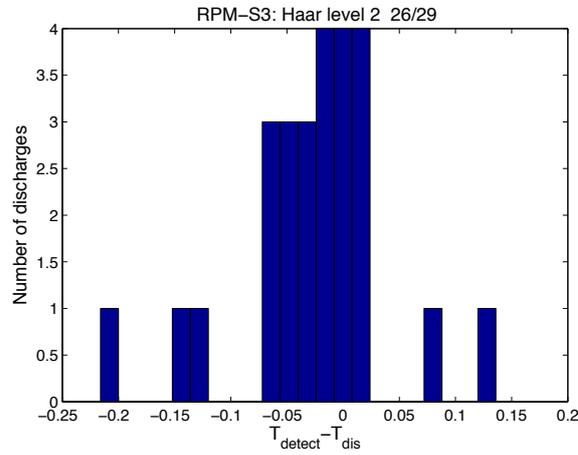


Figure 4.9: Histogram of RPM-S3 results with wavelet Haar level 2

Finally the third set, S3 was tested following the same process. Over 29 discharges, after several combinations of the different parameters the best results reached were 26 detections of 29 discharges. Figure 4.9 show the

histogram of the results for RPM-S3 with Haar level 2, where it can be seen that taking a window of 48 ms around disruption time, only the 41.38%(12/29) are detected.

Although the martingale test for change detection in data streams is a useful and efficient tool, it does not fit correctly our problem.

4.1.3 UMEL

After the previous analysis, it was thought that Universal Multi-Event Locator (UMEL), which has been explained in section 2.4.2, could be a useful tool for this purpose. Taking some characteristic signals which show a clearly distinguishable behaviour at disruption time, UMEL is applied to detect the ESP(external support vectors) out of the boundaries. In this study three signals have been used, figure 4.10:

- Plasma current which shows an abrupt decay of the amplitude from several MA (its operational value) to zero in some ms.
- Loop voltage (“jpf/gs/bl-nflopv<s”) is the voltage at plasma bound, and it shows a characteristic spike.
- Peaking factor also presents a characteristic spike. The main overview bolometer measures the radiated power density, and the processed line-integral bolometer signals are stored as a function of channel number and time. Two signals from the vertical camera, which are processed data on the basis of a weighted sum of representative channels, can be used: the total radiated power (“ppf/bolo/topi”) or the total bulk radiation (“ppf/bolo/tobu”). The first one is selected to be employed in this analysis although both signals practically follow the same pattern.

Considering the explanation at the beginning of this section 4, the plasma current and temperature signals should be good enough to detect disruption time locating the thermal and current quench; but the problem is the lack of temperature signal in many discharges. Because of this reason, temperature signal has been avoided. Other possible signal which were discarded is mode lock amplitude. Mode lock shows a typical behaviour close to the disruption but it is suitable to predict not to locate the disruption time.

In the work (Vega et al., 2010), UMEL was applied to differentiate disruptive and non-disruptive discharges without providing the times where the disruption was detected. As earlier pointed out, this analysis does not aim to classify between disruptive or non-disruptive discharges, it is intended to locate with high accuracy the disruption time. Therefore, it is assumed that the discharge has been identified as a disruptive shot previously.

The signals are downloaded and re-sampled to 1 kHz through a linear interpolation. The discharge is analysed from the beginning, if an ESV is

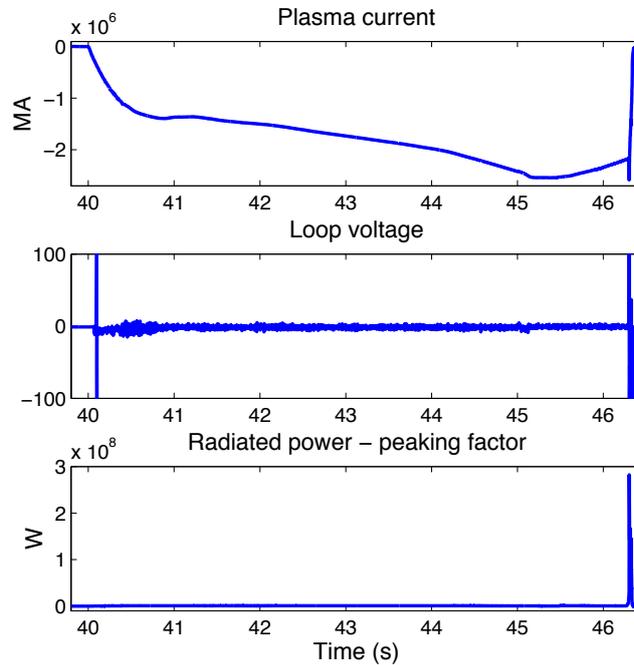


Figure 4.10: Signals used by UMEL on disruption time location

detected in the loop voltage, it is checked if the other signal/s contain a ESV in a temporal window of 300 ms around the time of the loop voltage ESV, figure 4.11. After that, the plasma current is selected in a range of time from 200 ms (200 samples) before the ESV chosen to 200 ms (200 samples) after the end of the discharge. If the disruption suddenly happens, the high frequency samples could not be detected in the next step, so that the last 200 samples are added. The wavelet of the selected plasma current data is calculated and UMEL is applied to the detail coefficients; locating the disruption at the time of the first ESV, figure 4.12.

Figure 4.13 shows two examples with different values of σ , C , ϵ for UMEL algorithm, the discharge 81867 disrupts at 46.304 s and the time detected is 46.2972 s in both cases. Several values for UMEL (σ , C , ϵ) and wavelet (level) parameters have been tested, and the best results have been reached using two signal combinations: plasma current with loop voltage; plasma current, loop voltage and peaking factor. The results are shown in table 4.4.

It is concluded that UMEL could correctly recognise a discharge as safe or disruptive. However if the purpose is to locate the disruption time, this technique is not enough accurate, therefore it would be necessary other methodology which allow to obtain disruption times with errors smaller than the results obtained at this analysis.

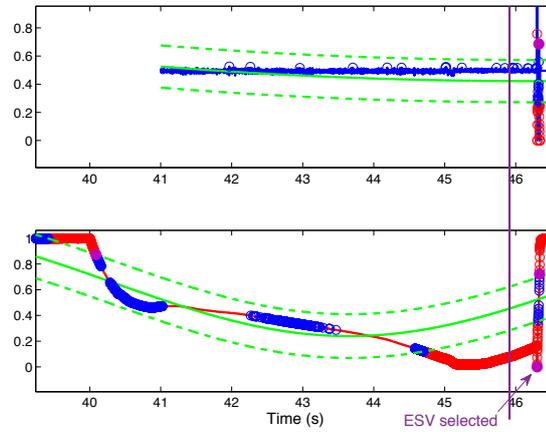


Figure 4.11: First step: UMEL is applied to the signals and the first shared ESV is detected

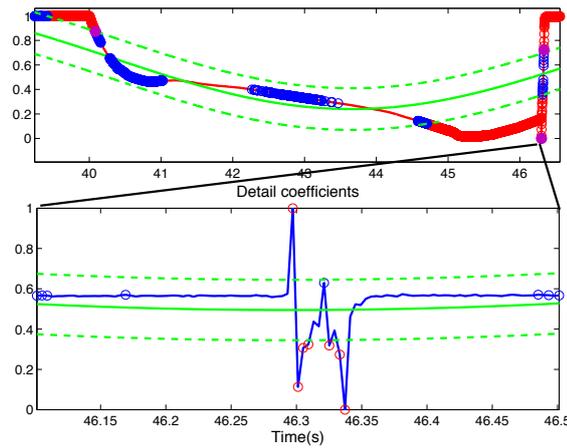
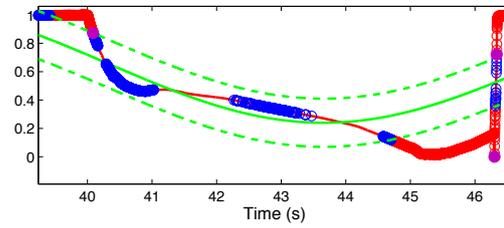
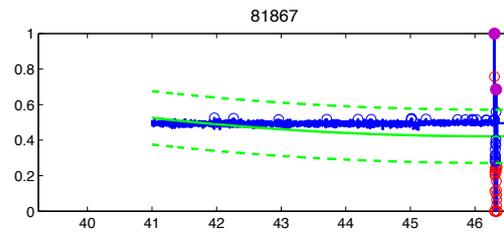


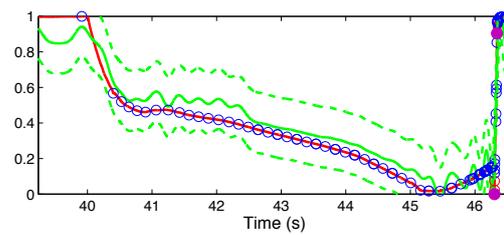
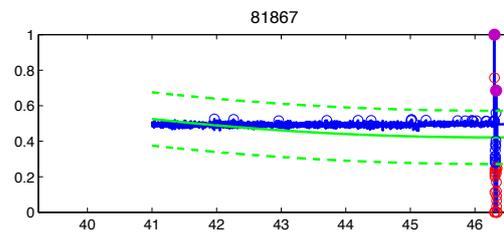
Figure 4.12: Second step: UMEL is applied to the detail coefficients of the plasma current data selection

| Simulation | $ \Delta t \leq 30 \text{ ms}$ | $ \Delta t > 30 \text{ ms}$ |
|------------|---------------------------------|------------------------------|
| 1 | 52.53%(135/257) | 94.94%(244/256) |
| 2 | 63.59%(166/257) | 95.72%(247/256) |

Table 4.4: Results for disruption time location using UMEL. Simulation 1: plasma current and loop voltage. Simulation 2: plasma current, loop voltage and peaking factor. $|\Delta t| = |t_{dis} - t_{detect}|$



(a) Parameters 1



(b) Parameters 2

Figure 4.13: Example of different parameters for UMEL to locate disruption time

4.2 Wavelets

After developing many analysis as previously explained, this final work using wavelets has obtained great results and it has been chosen as the automatic tool to locate disruption time (Moreno et al., 2014a).

Wavelets are basis functions which can be used to approximate a signal or extract information from data. They will be briefly described in this section, however a wide explanation can be found in section 2.2. They are similar to Fourier transforms but wavelets are able to represent a signal in the time and frequency. The discrete wavelet transform (DWT), proposed by Mallat (Mallat, 1989) is an efficient algorithm for calculating the wavelet coefficients of a discrete series. The idea is to filter the series, using the high and low pass filters associated with the wavelet basis to obtain the wavelet coefficients. In DWT, the signal is convolved and decimated. Therefore, a modified version of the traditional wavelet transform DWT has been used in this study. Non-Decimated Wavelet Transform (NDWT) or stationary wavelet transform has no sub-sampling step so it keeps the same number of coefficients of each level. The basis or family functions chosen in this study is the Haar wavelet. A Haar wavelet is the simplest type of wavelet, a sequence of rescaled “square-shaped” functions. As a special case of the Daubechies wavelet, the Haar wavelet is also known as $D2$. The main disadvantage of this family is that it is not continuous and not differentiable; but this is an advantage for the analysis of signals with sudden transitions.

The data processing algorithm is based on following the temporal evolution of the plasma current time derivative during the last 3 s of the discharge before the plasma current crosses the value of 50 kA. To this end, its evolution is analysed through a multilevel non-decimated wavelet decomposition looking for the temporal location of the components that determine the current quench. Two sets of coefficients are provided: detail and approximation. The analysis has been done using both coefficients but the best results have been obtained with the approximation coefficients. Therefore, the location is carried out with the coefficients of the approximation of level L . This latter signal shows large-scale features at the current quench times. To discriminate these features from others, due to phenomena different from disruptions, it is necessary to identify an appropriate threshold. This threshold allows selecting only the main peaks that correspond to the biggest changes in plasma evolution. It is defined by the bounds of the band:

$$\bar{W}_L - k\sigma_W, \bar{W}_L + k\sigma_W,$$

where \bar{W}_L and σ_W are the mean value and the standard deviation of the approximation coefficients obtained after applying non-decimated wavelet of level L to current time derivative, respectively, and k is a small positive integer. With regard to the developed interface and create the ground truth, the

user can select the discharge to analyse; the program will show the plasma current, the time derivative of plasma current, and the approximation coefficients with the threshold. The possible disruption times are shown and the user chooses the point that best represents the beginning of the disruption. In figure 4.14, it can be seen the process of the data processing algorithm. The program shows the best options under the parameters selected (in this case a level of 3 and for σ , $k = 1$) and, user selects the proper time. In this example the difference between the current quench time estimated and the disruption time is 1 ms.

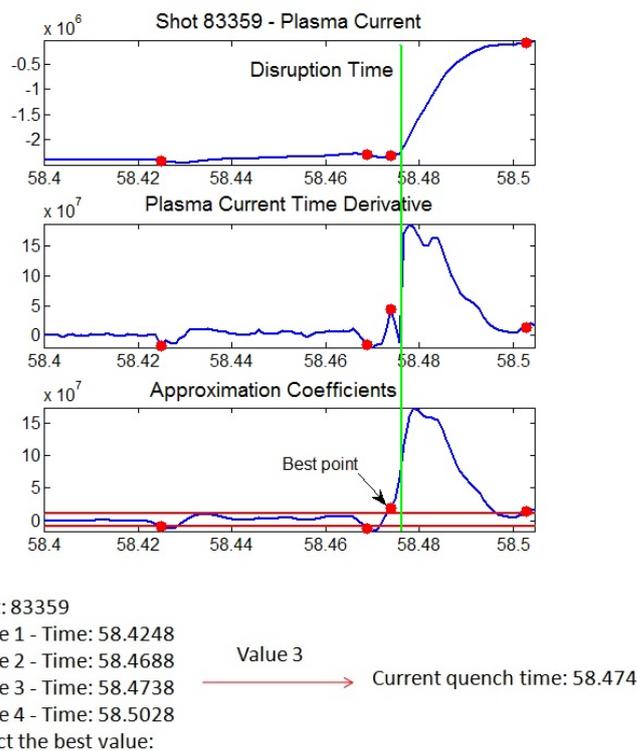


Figure 4.14: Example of the data processing algorithm for disruption time location

Very promising results are obtained from the final analysis of the disruption database; which is formed by 256 non-intentional and intentional disruptions from 2011 to 2012 JET-ILW campaigns. The discharges have been analysed with different values of levels for non decimated wavelet (levels 1 – 6) and threshold (values of k from 1 to 6). Summarizing, the best results are obtained for levels 1, 2, 3, and 4 with sigma $k = 1$ (Table 4.5). A window of 16 ms around the current quench provides a success rate of 100% and a window of 8 ms around the current quench shows a success rate of

99.61% (only 1 discharge is outside the window of 8 ms).

| Level (L) | Sigma (k) | Window of 16 ms | Window of 8 ms |
|-----------|-----------|-----------------|-----------------|
| 1 | 1 | 100%(256/256) | 99.61%(255/256) |
| 2 | 1 | 100%(256/256) | 99.61%(255/256) |
| 3 | 1 | 100%(256/256) | 99.61%(255/256) |
| 4 | 1 | 100%(256/256) | 99.61%(255/256) |

Table 4.5: Values of levels for non-decimated wavelet and threshold that provide the best results

Taking into account that many disruptions occur during JET experiments and that analysis and estimation of disruption times are also carried out manually, it is not unusual to find human errors when a big database is analysed. The proposed automatic data processing algorithm allows both reducing human efforts to locate the disruption times and standardizing the estimations (with the benefit of being less vulnerable to human errors). Locating disruption times, minimizing human errors, and establishing a general criterion are important issues which must be addressed. Actually, disruption predictors are developed using the disruption time to characterize the disruptive features of the training samples, accordingly, enhancing the importance of the precise estimation of disruption times. If the disruption time is not estimated correctly, the samples that define the disruptive behaviour could be confusing and provide false information to classifiers. This algorithm shows a possible way to generalize and calculate the time of disruptions, which could help in different studies: benchmarking of theoretical models, development of physics-driven models, and training of disruption predictors. After numerous simulations with different values of the parameters, the best results shown in 4.5 provide a success rate of 100% in a window of 16 ms around the current quench and 99.61% in a window of 8 ms. Disruption location is sometimes confused, so user selection lets determine which point corresponds better to the beginning of disruption. Future works could include other signals and parameters to define deeply disruption time, such as electron temperature, radiated power, loop voltage, etc. Furthermore, it could be studied whether including the time derivative of the plasma current in a classifier can provide useful information about disruptions.

Chapter 5

Disruption predictions for the next generation of tokamaks

This chapter covers all the disruption prediction approaches thinking on ITER. The first section 5.1 analyses plasma dynamics to detect the approach to the disruption boundaries (see (Vega et al., 2015a)); and describes the performance of some ILW predictors trained and tested with ILW data from JET experimental campaigns (see (Moreno et al., 2015)). Then, in section 5.2 the main disruption predictor from scratch works are explained: section 5.2.1 shows the probabilistic Venn predictors applied to disruption prediction from scratch, in section 5.2.2 the APODIS structure is applied to this approach, and finally in 5.2.3 the disruption prediction from scratch process is applied to to the whole ILW campaigns databases currently available.

5.1 ITER-like Wall (ILW) predictors

5.1.1 Analysis of plasma dynamics to detect the approach to the disruption boundaries

The analysis of the plasma dynamics in disruptive discharges shows that the plasma evolves quietly in the safe zone of the parameter space when it is far from the disruption. Near the disruption, the plasma evolution can alternate between disruptive and non-disruptive zones during a variable period of time. From a certain time instant, the plasma transits for the last time to the disruptive zone, remains in this zone also a variable time and, finally, the plasma disrupts. The last time instant in which the plasma was in the safe zone has been called “no-return” point.

Currently, disruption prediction is carried out by means of machine learning methods that distinguish between disruptive and non-disruptive behaviours in the multi-dimensional operational space. A training dataset made up of disruptive and non-disruptive examples allows determining the

separation frontier between both zones, as shown in figure 5.1. During the execution of a discharge, inputs are provided to the model on a periodic basis and an alarm is triggered when the output is “disruptive”. In figure 5.1, cyan curve represents a possible trajectory of the plasma behaviour in the operational space during a discharge.

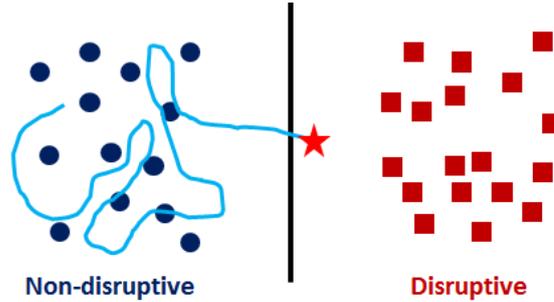


Figure 5.1: Operational space during a discharge

Therefore, an approach to characterize the plasma dynamics in the operational space could be the temporal evolution of the distance to the APODIS separating hyper-plane (see section 3.1). Figure 5.2 shows some examples of this evolution. In general the plasma evolves in a steady way during the plasma current flat top, it means there is no erratic trajectories in the operational space. The plasma state remains at a “constant” distance from the separating hyper-plane, the furthest the better, and the transit is fast. However, sometimes the plasma transits and comes back to the non-disruptive state but there exist “no-return points”. This behaviour is shown at discharge 87355 in figure 5.2, and it would be a false alarm triggered by APODIS if the discharge was non-disruptive. Despite of some events can push the plasma towards the hyper-plane, most of times the plasma stays in the safe region and recovers the initial distance.

The transit speed between the non-disruptive and the disruptive zones, see figure 5.3, can be determined: $v = \Delta d / \Delta t = (d(t + 32) - d(t)) / 32$. From ILW campaigns $C31 - C33$, 297 discharges between shots 84628 and 87532 have been analysed. Figure 5.3 illustrates the transit speed distribution, which follows a gamma probability distribution with shape parameter $t = 1.6274$ and scale parameter $\lambda = 3.0516$. The mean value is 0.533 and variance 0.175.

The temporal evolution of the APODIS distance can be used for the creation of specialised databases to identify events that produce loss of stability. In this way, APODIS distances are used to create a conformal predictor (see section 2.3.2) and provide a value of confidence and credibility for each distance value. The nonconformity measure is the distance to the hyperplane, so the closer the samples to the hyper-plane the lesser credibility for the

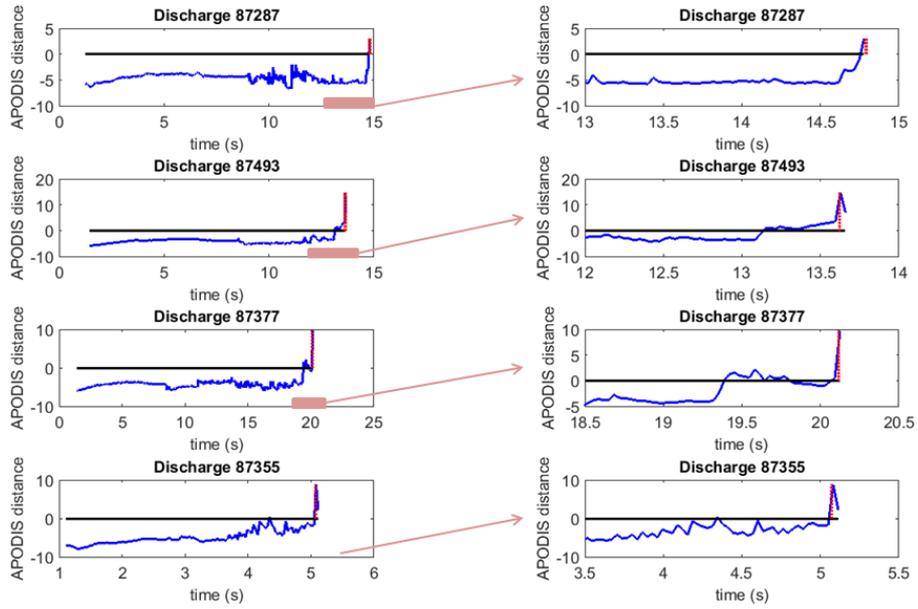


Figure 5.2: Examples of temporal evolution of the distance to the APODIS separating hyper-plane

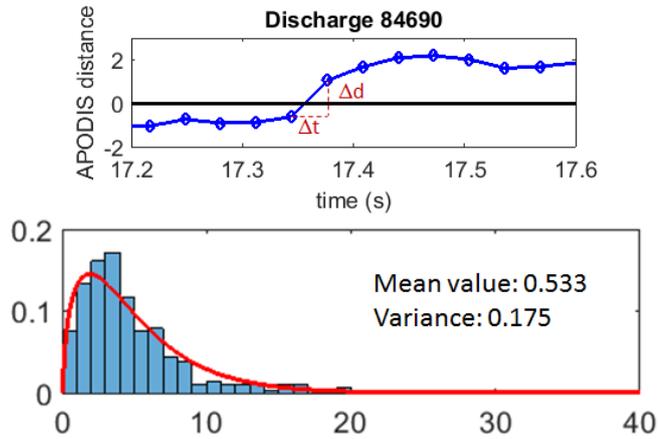


Figure 5.3: Example and distribuyion of transit speed

prediction. Therefore the samples are strange for low credibility values, as illustrated in figure 5.4.

To compute the credibility, a conformal prediction framework is used:

- The first feature vector of a discharge uses the initial training dataset.

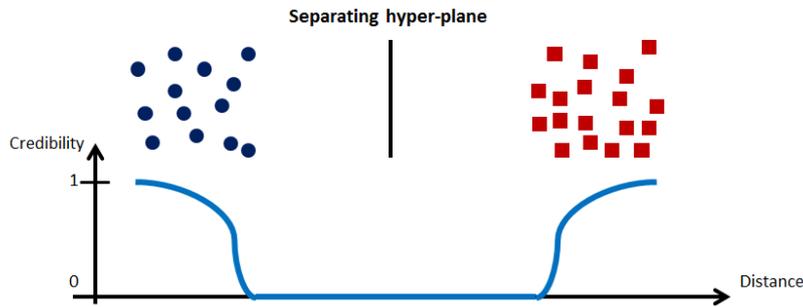


Figure 5.4: Credibility diagram

- After computing the credibility of the first feature vector, the vector is added to the training set.
- With each new feature vector, both the separating hyper-plane and the credibility are computed. As new feature vectors are added, the hyper-plane can change.

In figure 5.5 some examples are shown and it is observed the credibility is more sensitive than the APODIS distance to diagnose the plasma dynamics. It can be noted that a decreasing credibility means that the plasma approaches to the separating hyper-plane; and an increasing credibility means that the plasma moves away from the separating hyper-plane.

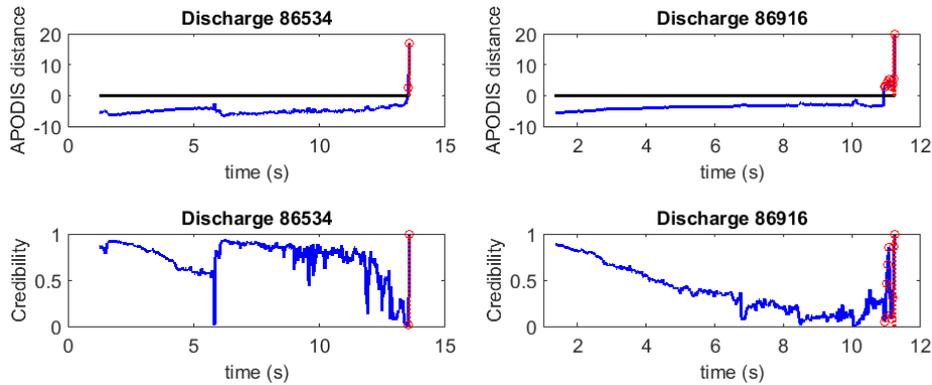


Figure 5.5: Plasma dynamics examples with conformal prediction

If the analysis is focused to the last seconds of the discharge, see figure 5.6, it is observed a behaviour where the plasma crosses the time t_1 defined as “no-return” point. After that, the credibility decreases, it means the plasma approaches to the separating-hyperplane until time t_2 . At this point, the plasma evolves in a steady way around credibility 0 until the time t_3 is

reached, thus, the samples are around the separating-hyperplane. From t_3 the credibility increases until disruption time, t_4 . It is understood as the plasma moves away from the separating-hyperplane as the samples go deep into disruptive space.

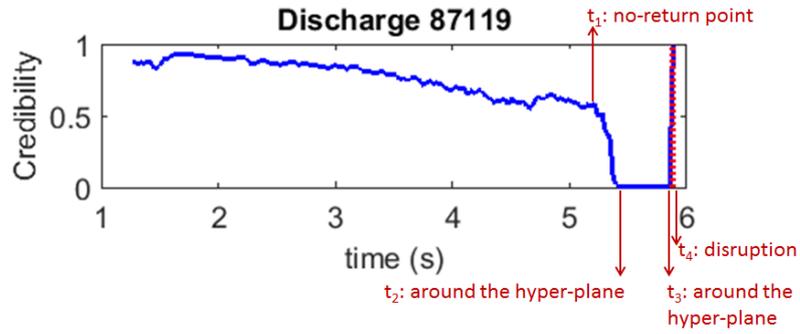


Figure 5.6: Plasma dynamics at non-return point

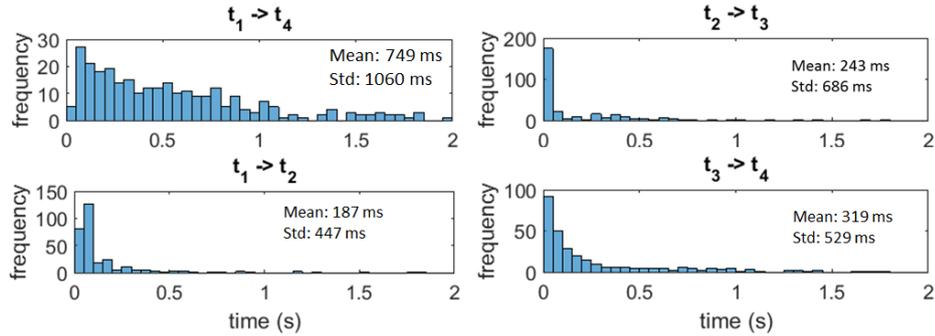


Figure 5.7: No-return times histogram

Finally, figure 5.7 shows some histograms. The first histogram illustrates interval time $t_1 - t_4$, where the mean is 749 ms and the standard deviation is 1060 ms. Intervals $t_1 - t_2$, $t_2 - t_3$ and $t_3 - t_4$ are shown in the remaining plots showing different means and standard deviations. These results allow to have a first approach about the evolution and duration of the plasma dynamics around disruption boundaries. This can be a special point because the plasma does not recover its non-disruptive character. Therefore, no-return points have to be investigated. Physics models around these times can provide knowledge about the disruptive behaviour and, perhaps, these models could be used as disruption predictors. The analysis of no-return points is a clear problem of data mining. The plasma has to be characterized around them.

5.1.2 Comparison between predictors trained and test with ILW data

In view of ITER, every study in JET is extremely important due to the similarity between the two devices, even more with the new ILW. Disruptions have gained relevance because of the damage that the device can receive, hence the importance of disruptions prediction and avoidance. Currently, the only predictor implemented in real time with a high performance is APODIS. Despite it was trained with data from C-wall experimental campaigns, it has kept a good performance during all the ILW experimental campaigns. In particular, during campaigns $C31 - C34$ (July 2013- October 2014), APODIS obtained 85.38%(333/390) of success rate with 2.46% of false alarm rate, while mode lock trigger reached a success rate of 66.67%(333/390). Focusing on ITER, it is very important to obtain results with predictors trained and tested with data from ILW campaigns. For this purpose, a comparison between an APODIS version and a simpler one layer predictor has been carried out, using ILW data for training and testing.

In 3.3 results are obtained with APODIS trained with C wall data (see 3.1) and without any retraining in spite of its use with metallic wall discharges. Taking into account ITER will work with a similar wall than the current ITER-like wall (ILW) at JET, the purpose of this study is to compare predictors trained with data from JET ILW campaigns. An APODIS version has been trained with data from campaigns $C28 - C30$ (1036 non-disruptive discharges and 201 non intentional disruptions), and it has been compared with a simpler one layer predictor trained with the same data. They have been tested with experimental data from campaigns $C31 - C34$ (1051 non-disruptive discharges and 390 non intentional disruptions).

It is important to note that this is useful in ITER view because the models are trained and tested with ILW data, as it will be done at ITER. The idea of a simpler predictor comes from the fact that a huge variety of experiments have been done along C-wall campaigns (see table 3.2), while ILW campaigns $C28 - C30$ (see table 3.7), which have been used to train, have a much smaller amount of data and less variety of experiments. As mentioned, APODIS online version on JET was trained with a huge database from C-wall campaigns (see 3.1). The database used in this study corresponds to the table 3.7:

- *Train*: campaigns $C28 - C30$ (September 2011 - July 2012), 201 non-intentional disruptions and 1036 safe shots.
- *Test*: campaigns $C31 - C34$ (July 2013 - October 2014), 388 non-intentional disruptions and 1031 safe discharges.

An sliding window mechanism applied to APODIS is developed in other study that is explained in 3.15. In (Moreno et al., 2014b), the sliding window

mechanism shows a better success rate but a slightly increase in the false alarm rates for higher temporal resolutions. Despite of this, in order to compare with APODIS structure, the same temporal resolution as APODIS (analyse consecutive temporal windows of 32 ms) has been taken for this simpler predictor. The selected structure in this work is a single one layer formed by SVM classifiers. As explained in 2.3.1, SVM maps the input space of the feature vectors into a new feature space, simplifying the separating hyper plane (or decision function) in the feature space. This transformation is done by a kernel function. The radial basis kernel function is the one selected to train our SVM classifiers:

$$K(x, x_i) = \exp \left\{ -\frac{|x - x_i|^2}{2\sigma^2} \right\}$$

This kernel function needs as inputs the regularization parameter C and the kernel parameter γ . SVM theory has been exposed in 2.3.1. As the structure is simpler, *high-computing performance*(HPC) has not been necessary and all the models have been trained and tested in Matlab using the *libsvm*¹ library. Although in some studies (Moreno et al., 2014b; Pereira et al., 2015) is demonstrated that mode lock amplitude and plasma inductance signals are essential for APODIS, in this work it has been used two set of signals to develop the predictor: the 7 signals (14 features after data processing) from table 3.1 in order to use the same data as APODIS; and following the results in (Moreno et al., 2014b), a set formed by 3 signals(6 features after data processing), mode lock amplitude, plasma inductance and plasma current.

Several numbers of predictors have been developed combining values of C and γ parameters: $C = 10^4 : 10^1 : 10^8$ and $\gamma = 0.01 : 0.05 : 10$. Furthermore three different predictors have been considered taking into account the disruptive samples selected, it means:

- *P1*: this predictor considers as disruptive samples only the 32 ms before the disruption, it means there is only one disruptive sample ($[-32 \text{ ms}, 0]$) for every disruptive discharge.
- *P2*: it is considered as disruptive samples 64 ms before the disruption, it means there are two disruptive samples ($[-64 \text{ ms}, -32 \text{ ms}]$, $[-32 \text{ ms}, 0]$) in a disruptive discharge.
- *P3*: it is selected as disruptive samples 96 ms before the disruption; it means there are three disruptive samples ($[-96 \text{ ms}, -64 \text{ ms}]$, $[-64 \text{ ms}, -32 \text{ ms}]$, $[-32 \text{ ms}, 0]$) in a disruptive discharge.

At this point it is necessary to remark that APODIS does not consider the last 32 ms as disruptive sample, on the contrary, as it is shown in figure

¹<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

3.2, it takes as disruptive information from 32 ms to 128 ms before the disruption. Therefore combining 2 set of signals, 3 predictors, 5 values of C parameters and 200 values of γ parameter, 6000 different models have been developed. From all these models, the best results for each predictor and set of signal are shown in table 5.1.

To present the results, the same criteria in 3.3 has been followed. The warning time is called the difference: $\Delta t = t_D - t_a$, where t_a is the time when the alarm is triggered, and t_D is the disruption time. Taking into account that at least 10 ms are needed to carry out mitigation action when a disruption happens, the results are presented in terms of:

- Valid detections: $10 \text{ ms} \leq \Delta t \leq 1.5 \text{ s}$
- Early detections: $\Delta t > 1.5 \text{ ms}$
- Tardy detections: $0 \leq \Delta t \leq 10 \text{ ms}$
- False alarms: an alarm is triggered in a safe discharge.
- Missed alarms: in a disruptive discharge, the alarm happens after de disruption time ($\Delta t < 0$) or there is no alarm triggered.

| Predictor | Success | Valid | Early | Tardy | False | Missed |
|---------------------|---------------------|---------------------|-------------------|--------------------|--------------------|-------------------|
| $P3$ 14 features | 92.53% (359/388) | 80.16% (311/388) | 4.12% (16/388) | 8.25% (32/388) | 2.72% (28/1031) | 7.47% (29/388) |
| $P3$ 6 features | 92.53% (359/388) | 79.38% (308/388) | 2.32% (9/388) | 10.82% (42/388) | 1.94% (20/1031) | 7.47% (29/388) |
| $P2$ 14 features | 91.75% (356/388) | 78.87% (306/388) | 9.02% (15/388) | 3.87% (35/388) | 2.23% (23/1031) | 8.25% (32/388) |
| $P2$ 6 features | 92.53% (359/388) | 79.64% (309/388) | 2.32% (9/388) | 10.57% (41/388) | 1.65% (17/1031) | 7.47% (29/388) |
| $P1$ 14 features | 90.72% (352/388) | 76.55% (297/388) | 3.87% (15/388) | 10.31% (40/388) | 1.84% (19/1031) | 9.28% (36/388) |
| $P1$ 6 features | 92.27% (358/388) | 78.09% (303/388) | 2.84% (11/388) | 11.34% (44/388) | 2.13% (22/1031) | 7.73% (30/388) |

Table 5.1: Best models for each set of signals and type of predictor

The best model is $P2$ with 6 features, which obtains a success rate of 92.53% and the 81.6% of the detections are at least 10 ms before the disruption time; while a 10.57% is detected before the disruption time but too late to carry out any mitigation action; and 7.47% are missed alarms. The false alarm rate is low, only 1.65%. To make a realistic comparison an APODIS online version has been trained and tested over the same database, however due to the high computational cost the parameters used to APODIS offline training has been selected between a reduced ranges of values. For the first layer is taken $C = 1000$ and $\gamma = 0.01$ as values of the RBF kernel, and the

second layer uses $C = 10^7$ as parameter of the linear kernel. In addition APODIS offline has taken as safe discharges for the training set, the 300 first non-disruptive shots from campaigns $C28 - C30$, this decision has been taken due to the computational cost if the datasets are higher. Table 5.2 shows the results of the offline APODIS version trained and tested with the same data as the new predictor; and the results of the best model selected for the new predictor. In figure 5.8, the comparison between both predictors is illustrated. The green lines represent 10 ms and 1.5 s before the disruption time respectively.

| Predictor | Success | Valid | Early | Tardy | False | Missed |
|--------------------|---------------------|---------------------|-------------------|--------------------|--------------------|--------------------|
| $P2$ 6 features | 92.53% (359/388) | 79.64% (309/388) | 2.32% (9/388) | 10.57% (41/388) | 1.65% (17/1031) | 7.47% (29/388) |
| APODIS offline | 89.43% (347/388) | 74.74% (290/388) | 4.12% (16/388) | 10.57% (41/388) | 3.01% (31/1031) | 10.57% (41/388) |

Table 5.2: Results of APODIS offline and $P2$ -6 features

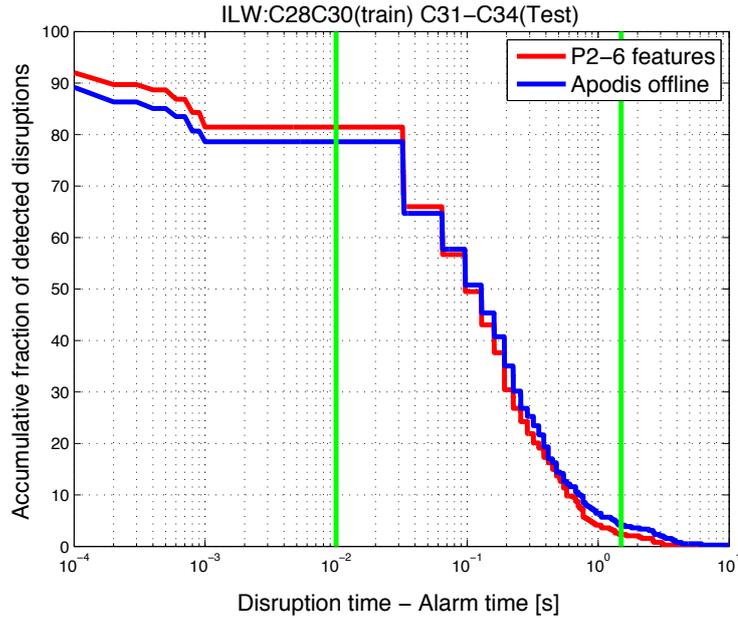


Figure 5.8: Comparison between ILW predictors

To obtain the best results for the simpler one layer predictor, as explained in this work, multiple models have been trained achieving a predictor with a good performance. This predictor formed by only one layer of a SVM classifier and only 3 signals, reaches 79.64% of valid detections and 1.65% of false alarm rate. It can predict with a warning time less than 10 ms, the 81.6% of

the disruptions. The offline version of APODIS, trained with the same data, reached a 74.74% of valid detections and 3.01% of false alarm rate. We can conclude that the simpler one layer predictor has a great performance, even better than the offline version of APODIS. However it cannot be said this is a better approach due to APODIS online was trained with a higher variety of experimental campaigns and bigger amount of data training; on the contrary the campaigns $C28 - C30$ are the first campaigns of the ILW at JET, and the experiments are quite similar. However, the one layer predictor achieves two important requirements for disruption predictors: simplicity, this means to develop the simplest possible classifier to distinguish between disruptive and non-disruptive behaviours at any moment; and fast training process (thinking in a disruption predictor from scratch), the training process should be fast enough (from a computational point of view).

As future work, APODIS could be trained with a higher variety of campaigns in the future when JET operates more ILW campaigns; even it could be considered the 32 ms before the disruption time as disruptive sample. On the other hand, the simpler one layer predictor, due to its speed could be checked from scratch as an approach to ITER.

5.1.3 Deuterium-Hydrogen ILW campaign

The last ILW campaign $C34$, formed by 139 non-disruptive discharges and 37 non-intentional disruptions see table 3.7), was operated using deuterium-hydrogen, instead of deuterium-deuterium. The rest of the ILW campaigns were developed with deuterium-deuterium discharges. Hydrogen campaigns are interesting for physics studies but not to reach ignition or develop data-driven models for disruption predictions. The reason is these experiments are much more stable and the operation conditions are quite different from the deuterium-deuterium experiments or even deuterium-tritium. However, ITER would start the operation with an hydrogen campaign. Furthermore, ITER and DEMO, disruption predictors have to be ready to work from the beginning of the operations and have to be able to recognize an incoming disruption. Therefore the first databases available to develop an efficient disruption predictor will come from an hydrogen campaign.

The database used in this study corresponds to the table 3.7:

- *Train*: campaign $C34$, 37 non-intentional disruptions and 137 safe shots.
- *Test*: campaigns $C28 - C33$, 552 non-intentional disruptions and 1928 safe discharges.

Following the same criterion in 5.1, several numbers of predictors have been developed: combining values of C and γ parameters ($C = 10^4 : 10^1 : 10^8$ and $\gamma = 0.01 : 0.05 : 10$); three types of predictors ($P1, P2, P3$); and two

set of signals (14 and 6 features). In this case the best results are obtained with $P1-6$ features obtaining the results shown in table 5.3. The results show 70.11% of valid detections and 14.1% of false alarm rate. The predictor can predict with a warning time less than 10 ms, the 72.83% of the disruptions.

| Predictor | Success | Valid | Early | Tardy | False | Missed |
|------------|-----------|-----------|----------|----------|------------|----------|
| $P2$ | 87.32% | 70.11% | 2.72% | 14.49% | 14.21% | 12.68% |
| 6 features | (482/552) | (387/552) | (15/552) | (80/552) | (274/1928) | (70/552) |

Table 5.3: Results of disruption predictor trained with the hydrogen campaign

At this point, it can be said that the results are quite unsatisfactory. In order to analyse these results, in figures 5.9 and 5.10 it is shown a scatter plot of the features (standard deviation of the FFT and mean value) of plasma inductance and mode lock signals. It can be seen in figure 5.9 that the disruptive (cyan, red and green circles) and non-disruptive (blue circles) features on campaign $C34$ are practically linearly separable. On the other hand, figure 5.10 shows an overlap of disruptive and non-disruptive features for campaign $C28 - C30$. It has been plotted only campaign $C28 - C30$ to compare because there will be too many samples if all campaigns were plotted. Therefore if a model is generated with data linearly separable the decision boundary would be a linear hyperplane which cannot be able to classify data overlapped such the data shown in figure 5.10.

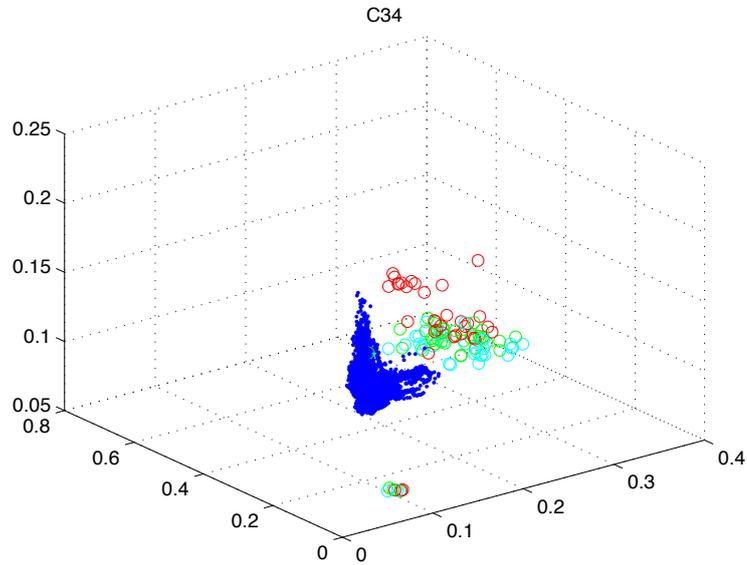


Figure 5.9: Scatter plot: $C34$ features

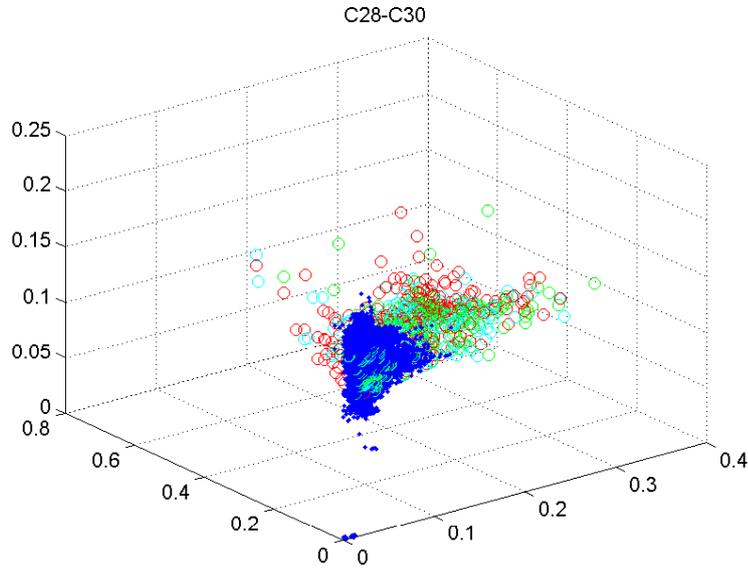


Figure 5.10: Scatter plot: $C28 - C30$ features

Despite the previous analysis, ITER cannot wait until the end of the first campaign, in this way it would be developed a disruption predictor from scratch. This approach will be explained in the following section.

5.2 Disruption predictors from scratch

All the methodologies about disruption prediction that have been worked in the previous chapters and sections depend on a database of past discharges to create models able to predict disruptions. However, next generation of tokamaks such as ITER and DEMO can suffer irreversible must be able to predict disruption from scratch, it means from the beginning of the operation when there are not past discharges. Recently, two different works have dealt with the development of adaptive data-driven predictors from scratch to learn from the incoming data. As previous predictors based in data-driven models, these both predictors are not based on plasma physics, but on extracting knowledge to identify the incoming disruption from the different signals provided by the device during the experiments. As explained in 3.3, APODIS has achieved high success rates and low false alarm rate during the ILW campaigns $C28 - C34$ (Vega et al., 2013a; Moreno et al., 2015). Furthermore, APODIS has proven no *ageing* effect, it means that the predictor has not suffered deterioration derived from operating the device in different operational regions from those used for training. Due to these performance, APODIS structure has been used to develop a predictor from scratch and

to estimate the minimum number of disruptions to have a reliable predictor (Dormido-Canto et al., 2013), which is showed in 5.2.2. In (Vega et al., 2014b), a particular algorithm based on probabilistic classifiers has been developed and it has been applied to the database of the first ILW campaigns of JET $C28 - C30$ (see table 3.7); this predictor is showed in 5.2.1.

Other predictor from scratch (Aledda et al., 2013) makes reference to fault detection and isolation (FDI) techniques (Patton et al., 1989). In (Aledda et al., 2013), the disruption prediction is formalized as a fault detection problem, where the discharges which are correctly terminated (non-disruptive discharges) are assumed as the normal operation conditions and the disruptions are assumed as status of fault. The normal operation conditions model was built with non-disruptive discharges from ASDEX Upgrade (AUG) and the dynamic structure of the data was estimated through the fitting of a multivariate AutoRegressive with eXogenous inputs (ARX) model. The datasets are composed of time series of the radiated fraction of the total input power, the internal inductance and the poloidal beta. The disruption prediction system is based on the analysis of residuals in the multidimensional space of the selected variables. The discrepancy between the outputs provided by the ARX model and the actual measurements is an indication of process fault (disruption). The predictor was applied to AUG data between 2002 and 2009. Results are promising but lower false alarm rates are needed. It should be noted that the methodology is not applied to conditions *from scratch* but the technique is susceptible of such a development. However, thinking of ITER, the authors state that, perhaps, the method cannot be applied during the very first pulses of ITER due to the need of a sufficient number of pulses safely landed.

As mentioned previously, in ITER and DEMO, disruption predictors have to be ready to work from the beginning of the operations and have to be able to recognize an incoming disruption. As mentioned, nowadays, there are not satisfactory physic-driven systems and data-driven models are the only viable option. A disruption predictor is a pre-requisite for any mitigation system, therefore, at least, the following operation requirements should be considered for any disruption predictor from scratch: learning from scratch, real-time operation, high success rate, high learning rate, early recognition of disruptions, low rate of false alarms, controlled ageing effect, predictor simplicity, fast training process and reliable predictions.

- *Learning from scratch*: disruption predictors require data from past experiments to train a model. However, there will be a complete absence of previous experimental data in ITER and DEMO. The training has to be carried out from scratch as the discharges are produced. This lack of previous information is not only limited to new fusion devices but it also happens in existing devices when significant changes are implemented (for example, in JET after the installation of the metallic

wall).

- *Real-time operation*: the predictor should be capable to work in a real-time network in order to provide a decision with enough time to mitigate the disruption.
- *High success rate*: it means a low number of missed alarms. For example ITER requires a success rate $\geq 95\%$.
- *High learning rate*: a high success rate is not enough for a disruption predictor from scratch, due to a predictor from scratch has to work efficiently from the beginning, a high success rate should be reached as soon as possible.
- *Early recognition of disruptions*: in any disruption mitigation technique, a delay between the alarm triggered by the predictor and the start of the mitigation action exists, this is the reaction time. Then the warning time has to be higher than the reaction time, in order to carry out an efficient mitigation.
- *Low false alarm rate*: in a real-time network, a false alarm means that control system carries out a mitigation action when an alarm is triggered during a non-disruptive discharge but it is not possible to know whether or not the alarm was false. Therefore, the device is stopped and the experiment is not finished. Current and future devices cannot allow this, they need the minimum number of false alarm.
- *Controlled ageing effect*: a disruption predictor should not suffer deterioration derived from operating the device in different operational regions from those used for training.
- *Simplicity*: this means to develop the simplest possible classifier to distinguish between disruptive and non-disruptive behaviours at any moment. To this end, the disruption configuration space will have to be characterized with a reduced number of features compatible with the best possible generalization capability.
- *Fast training process*: due to the fact that an adaptive classifier is needed to continuously incorporate new relevant information, the training process should be fast enough (from a computational point of view) to allow inter-shot trainings when necessary.
- *Reliable predictions*: any training process with low number of samples is an issue. For this reason, each individual prediction should be qualified with estimation about its reliability.

The compliance of the requirements establish the suitability of the predictor from scratch. Furthermore, the training process should be done in chronological order. The available information for training purposes is very limited and it depends crucially on the chronological order of the discharges. This means that the experimental program of a new device has to envisage a discharge sequence that takes into account the necessary continuous learning.

5.2.1 Probabilistic Venn predictor

In general, most of the machine learning algorithms provide a prediction without any information about its reliability. However predictions from different samples, it means from different feature vectors, probably contain different values of confidence and probability. Probabilistic classifiers express the probability of a sample, represented by a specific feature vector, to belong to a particular class. Following the Bayes' theorem (Theodoridis and Koutroumbas, 2008), it can be written as:

$$P(c_i|\vec{x}) = \frac{P(c_i)P(\vec{x}|c_i)}{P(\vec{x})} = \frac{P(c_i)P(\vec{x}|c_i)}{\sum_{i=1}^{N_c} P(c_i)P(\vec{x}|c_i)} = \frac{\textit{Priori} \cdot \textit{Authenticity}}{\textit{Evidence}} \quad (5.1)$$

This formula gives the posterior probability that the sample (feature vector), \vec{x} , belongs to the respective class c_i . This probability is strictly given under statistical randomness assumption. Feature vectors must accomplish the i.i.d. assumption (Vovk et al., 2005). However, to apply Bayes' theorem the likelihood and prior probability of each class must be known. Then, the estimation of the likelihood is an issue if this information is not known, which is the case of disruption prediction from scratch. Therefore, in order to avoid unjustified assumptions about the form of the probability distribution function, non-parametric approaches are used. The Parzen window method (Duda et al., 2000) is the most used non-parametric density estimator, which is based on kernel methods but it needs a minimum number of samples to produce reliable estimations. Due to there are not initial samples, Bayesian methods are discarded.

On the contrary, Venn predictors (Vovk et al., 2005) provide a probability prediction interval for each prediction and well calibrated outputs under the i.i.d. assumption. A well calibrated output means that the accuracy of the Venn predictors is bound to fall within the upper and lower probability intervals. These predictors belong to the family of conformal predictors (Vovk et al., 1999; Saunders et al., 1999), which have been explained in 2.3.2 and provide additional information about the credibility and confidence of a prediction. Basically, Venn predictors can be understand as the process to find the most likely hypothesis of all possible. Given a training set (x_i, y_i) with $i = 1, \dots, n - 1$, whose samples satisfy the i.i.d. assumption, the feature

vectors are x_i and the respective label is $y_i \in Y_1, \dots, Y_C$. The objective is to classify the new sample (x_n, y_n) , where the label y_n is unknown. The objective is to estimate the probability of the new sample of belonging to one class y_i . This way Venn predictor framework assigns each one of the possible classification Y_1, \dots, Y_C to the new sample x_n , and divides all the possible examples $(x_1, y_1), \dots, (x_{n-1}, y_{n-1}), (x_n, Y_j)$ into a number of categories based on a taxonomy $\tau_i, i = 1, \dots, T$. A taxonomy is a function A which classifies in T categories the relation between a sample (x_k, y_k) and the rest of the samples:

$$\tau_i = A((x_1, y_1), \dots, (x_n, y_n), (x_k, y_k)). \quad (5.2)$$

Many taxonomies have been used in Venn predictors, such as neural networks in (Papadopoulos, 2013), logistic regression in (Nouretdinov et al., 2012), SVM in (Lambrou et al., 2012), etc. For the disruption predictor from scratch using Venn predictors, the nearest centroid taxonomy (NCT) (Dashevskiy and Luo, 2008) has been selected. This taxonomy is given by:

$$\tau_i = A((x_1, y_1), \dots, (x_n, y_n), (x_k, y_k)) = Y_j, \quad (5.3)$$

$$j = \arg \min_{j=1, \dots, C} (\|x_i - C_j\|) \quad (5.4)$$

where C_j are the centroids of the C classes and $\|\cdot\|$ is a metric which corresponds to Euclidean distance.

The first step to predict the label of a new sample x_n with unknown label y_n using Venn predictors formulation is to choose a taxonomy with T categories, NCT in this case.

The next step is to assume that $Y_j = Y_1$ (i.e. $y_n = Y_1$) and partition the examples into categories using the NCT. The empirical probability distribution of the labels in the category τ that contains (x_n, Y_1) is:

$$p^{Y_1}(Y_k) = \frac{|(x^*, y^*) \in \tau : y^* = Y_k|}{C}, \quad k = 1, \dots, C \quad (5.5)$$

Therefore, computing the probabilities of each label within the category τ a row vector $(p^{Y_1}(Y_1), \dots, p^{Y_1}(Y_C))$ is obtained. Equally, the probability distribution that contains (x_n, Y_2) is:

$$p^{Y_2}(Y_k) = \frac{|(x^*, y^*) \in \tau : y^* = Y_k|}{C}, \quad k = 1, \dots, C \quad (5.6)$$

and a similar vector is obtained $(p^{Y_2}(Y_1), \dots, p^{Y_2}(Y_C))$. After all possible classifications to x_n have been assigned, a set of probability distributions P_C (a square matrix of dimension C) is generated:

$$P_C = \begin{pmatrix} p^{Y_1}(Y_1) & p^{Y_1}(Y_2) & \dots & p^{Y_1}(Y_C) \\ p^{Y_2}(Y_1) & p^{Y_2}(Y_2) & \dots & p^{Y_2}(Y_C) \\ \dots & \dots & \dots & \dots \\ p^{Y_C}(Y_1) & p^{Y_C}(Y_2) & \dots & p^{Y_C}(Y_C) \end{pmatrix} \quad (5.7)$$

The last step in the process is to assign a label to the sample x_n , which will be $y_n = Y_{k_{best}}$. The maximum and minimum probabilities obtained for each label Y_k , it means the maximum and minimum probability value in each column of P_C , define the interval for the probability of x_n belonging to Y_k .

$$k_{best} = \arg \max_{k=1, \dots, C} (p(\bar{k})) \quad (5.8)$$

where $p(\bar{k})$ corresponds to the mean value of the probabilities obtained for label Y_k among all probability distributions (the mean of every column of matrix P_C). The probability interval for this prediction is $[L(Y_k), U(Y_k)]$, which is the maximum and minimum probability of column with the maximum mean value.

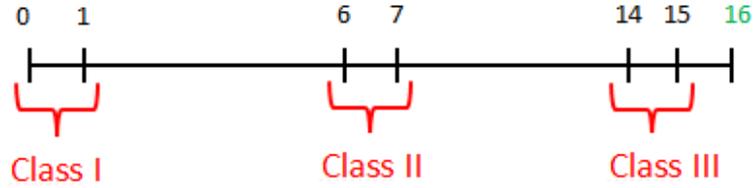


Figure 5.11: Data samples of Venn predictor example

For example, given the data samples in figure 5.11 with the classes $Y = I, II, III$, the objective is to classify the new sample $x_n = 16$. The taxonomy used is the NCT, it means that the category of a sample is the same as the label of its nearest centroid. There will be three categories $\tau_I, \tau_{II}, \tau_{III}$. Firstly we assume that x_n belongs to the first class I , $(16, I)$. The centroids will be:

$$Centroid(I) = \frac{0 + 1 + 16}{3} = 5.6 \quad (5.9)$$

$$Centroid(II) = \frac{6 + 7}{2} = 6.5 \quad (5.10)$$

$$Centroid(III) = \frac{14 + 15}{2} = 14.5 \quad (5.11)$$

The label of the nearest centroid of the sample $(16, I)$ is $(14.5, III)$, which means that $(16, I)$ is in category τ_{III} . The categories in this case yield:

$$\tau_I = (0, I), (1, I), (6, II) \quad (5.12)$$

$$\tau_{II} = (7, II) \quad (5.13)$$

$$\tau_{III} = (14, III), (15, III), (16, I) \quad (5.14)$$

Therefore, the first row of P_C is computed with the data corresponding to category τ_{III} .

$$P^I(I) = 1/3, P^I(II) = 0, P^I(III) = 2/3 \quad (5.15)$$

Now we assume that x_n belongs to the second class II , $(16, II)$. The centroids will be:

$$\text{Centroid}(I) = \frac{0 + 1}{2} = 0.5 \quad (5.16)$$

$$\text{Centroid}(II) = \frac{6 + 7 + 16}{3} = 9.67 \quad (5.17)$$

$$\text{Centroid}(III) = \frac{14 + 15}{2} = 14.5 \quad (5.18)$$

The label of the nearest centroid of the sample $(16, II)$ is $(14.5, III)$, which means that $(16, II)$ is in category τ_{III} . The categories in this case yield:

$$\tau_I = (0, I), (1, I) \quad (5.19)$$

$$\tau_{II} = (6, II), (7, II) \quad (5.20)$$

$$\tau_{III} = (14, III), (15, III), (16, II) \quad (5.21)$$

Therefore, the second row of P_C is computed with the data corresponding to category τ_{III} .

$$P^{II}(I) = 0, P^{II}(II) = 1/3, P^{II}(III) = 2/3 \quad (5.22)$$

Finally, we assume that x_n belongs to the last class III , $(16, III)$. The centroids will be:

$$\text{Centroid}(I) = \frac{0 + 1}{2} = 0.5 \quad (5.23)$$

$$\text{Centroid}(II) = \frac{6 + 7}{2} = 6.5 \quad (5.24)$$

$$\text{Centroid}(III) = \frac{14 + 15 + 16}{3} = 15 \quad (5.25)$$

The label of the nearest centroid of the sample $(16, III)$ is $(15, III)$, which means that $(16, III)$ is in category τ_{III} . The categories in this case yield:

$$\tau_I = (0, I), (1, I) \quad (5.26)$$

$$\tau_{II} = (6, II), (7, II) \quad (5.27)$$

$$\tau_{III} = (14, III), (15, III), (16, III) \quad (5.28)$$

Therefore, the third row of P_C is computed with the data corresponding to category τ_{III} .

$$P^{II}(I) = 0, P^{II}(II) = 0, P^{II}(III) = 1 \quad (5.29)$$

Then the P_C square matrix of dimension $C = 3$ which contains all the probability distributions is given by:

$$P_3 = \begin{pmatrix} 1/3 & 0 & 2/3 \\ 0 & 1/3 & 2/3 \\ 0 & 0 & 1 \end{pmatrix} \quad (5.30)$$

The mean of the probabilities for each class is:

$$P(\bar{I}) = 1/9 = 0.11, P(\bar{II}) = 1/9 = 0.11, P(\bar{III}) = 7/9 = 0.78 \quad (5.31)$$

it can be concluded that the new sample $x_n = 16$ belongs to class $y_n = III$ with probability 0.78 in an error bar of $[0.67, 1]$.

In (Vega et al., 2014b; Pereira et al., 2015), Venn predictors are used with the NCT. Both works deal with two classes, disruptive and non-disruptive, to predict disruptions; and they use the databases explained in 3.2, corresponding to the first ILW campaigns of JET $C28 - C30$ (1036 safe discharges and 201 unintentional disruptions, see table 3.7). The signals used are shown in table 3.4, and 14 features are employed by computing the mean value and standard deviation of the FFT of each signal, see 3.2. The selection of NCT allows to summarize the initial input data and reduce significantly all the disruptive and non-disruptive information. The predictions are carried out sequentially, the predictor begins with only two samples, one disruptive and one non-disruptive. Venn predictors learn by transduction (Vapnik, 2000), they do not need to train models as inductive classifiers.

The first predictor is obtained after the first disruption and, from that moment, all discharges are analysed in chronological order. Each discharge is analysed by simulating a real-time data processing, as explained in 3.2. After a missed alarm, a new training set is created to incorporate new knowledge. All possible combinations between 2 and 7 features, have been tested (9893 predictors have been developed). Given a specific combination of features, the algorithm showed in (Vega et al., 2014b) is executed for the whole dataset of discharges. The first predictor is generated after the first disruption and it is used with all posterior discharges, subject to retrain after every missed alarm. The results presented in figure 5.12 correspond to the evaluation of the predictors after the analysis of the whole database, it means, the success and false alarm rates are the cumulative results of the process after

| Feature Id. | | | | | | | | | | | | | | SR | FA | AVP |
|-------------|---|---|---|---|---|---|---|---|----|----|----|----|----|-------|------|-------------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | (%) | (%) | |
| | | x | x | | | | | | | | | | | 94.00 | 4.70 | 0.813±0.187 |
| | x | | x | | | | | | | | | | | 92.50 | 5.09 | 0.831±0.169 |
| | x | x | x | | | | | | | | | | | 94.00 | 4.31 | 0.809±0.191 |
| | | x | x | | | | | | | x | | | | 94.00 | 4.70 | 0.813±0.187 |
| | x | x | x | x | | | | | | | | | | 94.00 | 4.21 | 0.811±0.189 |
| | x | x | x | | | | | | | x | | | | 94.00 | 4.21 | 0.810±0.190 |
| | x | x | x | x | | | | | | x | | | | 94.00 | 4.21 | 0.811±0.189 |
| | x | x | x | | | x | x | | | | | | | 94.00 | 4.21 | 0.803±0.197 |
| | x | x | x | | | x | x | | | x | | | | 94.00 | 4.21 | 0.803±0.197 |
| | x | x | x | x | | | | | | x | x | | | 94.00 | 4.31 | 0.810±0.190 |
| | x | x | x | x | | | x | x | | x | | | | 94.00 | 4.31 | 0.802±0.198 |
| | x | x | x | | | x | x | | | x | x | | | 94.00 | 4.31 | 0.802±0.198 |

Figure 5.12: Features with the best results using Venn predictors

1237 discharges. The average prediction probability (AVP) represents the average probability interval in Venn predictors.

An important drawback in (Vega et al., 2014b) was the time required to compute the results (figure 5.12) that can be estimated in 1731 hours (≈ 2.4 months). Genetic algorithms (GA) are searching algorithms that simulate the process of natural selection. In (Pereira et al., 2015), the GA and the Venn predictors are combined with the objective not only of finding good enough features within the 14 available ones but also of reducing the computational time requirements. Five different performance metrics as measures of the GA fitness function were evaluated. The best metric was the measurement called Informedness, with just 6 generations (168 predictors at 29.4 hours).

5.2.2 APODIS from scratch

APODIS has been used to predict from scratch and to estimate the minimum number of disruptions to have a reliable predictor (Dormido-Canto et al., 2013). The first ILW campaigns of JET *C28 – C30* (1036 safe discharges and 201 unintentional disruptions, see table 3.7) have been used in chronological order. The first predictor is created after the first disruption and re-trainings are carried out after each missed alarm. The main result is that APODIS reproduces its good prediction capabilities and low rate of false alarms after including in the training process about 40 disruptions.

General methodology for disruption prediction from scratch

The problem to be solved is a binary classification but this methodology can be extrapolated to multi-class problems. As explained in section 3.2, the number of non-disruptive discharges is much higher than the number of

disruptive discharges. In this sense, the fusion databases to develop disruption predictors are highly unbalanced. A predictor from scratch have to be trained as discharges are produced, this means that discharges are used in chronological order as they occur. Regardless of the balanced or unbalanced data, it has to be established when a retraining is required to improve the predictor. A first criterion has been to generate a new predictor just after the occurrence of every new disruptive discharge. The reason for this is to incorporate any new information from disruptive discharges to be used in the future prediction of disruptions. To explain this criterion, let us consider the simplified example in figure 5.13, where a set of 50 discharges (in chronological order) is represented. The example considers 45 non-disruptive discharges (represented by circles) and 5 disruptive discharges (denoted by squares). It is just an illustrative example and therefore, for didactic purposes, the disruptions are chosen evenly spaced. From a practical point of view, the methodology is exactly the same regardless of the number of disruptive discharges or the variable number of non-disruptive discharges between disruptive ones. A greater number of disruptive discharges will result in a larger number of trained models. According to the criteria of generating a predictor after every disruptive discharge, the situation simulated in figure 5.13 requires the generation of five predictors.

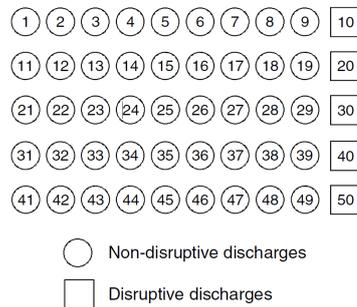


Figure 5.13: Simplified example of prediction from scratch

Two training data sets approaches have been used:

- *Unbalanced training data sets*: in this case, it is well known that the decision boundary established by the machine learning algorithm tends to be biased towards the majority class. Therefore the minority class samples suffer a certain risk of misclassification. In nuclear fusion databases, the disruptive and non-disruptive samples are highly unbalanced as mentioned previously. In this case, each predictor from figure 5.13 will take all the disruptive and non-disruptive discharges available until this moment. For example, $M(1)$ training set will be formed by disruptive discharge number 10 and non-disruptive dis-

charges 1 – 9; $M(2)$ training set will be formed by disruptive discharges number (10, 20) and non-disruptive discharges (1 – 9, 11 – 19); etc. It is important to note that we are talking about unbalanced number of discharges, the number of samples used in each disruptive and non-disruptive discharge follow the process explain in section 3.1.

- *Balanced training data sets*: the balance can be understood as a guarantee of fairness. Most machine learning algorithms performs well due to they aim to optimize the overall classification accuracy. It is important to note that classification performance also depends on the training set size and the problem complexity. For example, a linearly separable problem is not susceptible to any amount of unbalance (Japkowicz and Stephen, 2002). As the degree of complexity increases, the class unbalance factor starts influencing the classifier generalization ability. In this case, each predictor from figure 5.13 will take all the disruptive discharges available until this moment, and the same number of non-disruptive discharges selected randomly between all the safe discharges available. For example, $M(1)$ training set will be formed by disruptive discharge number 10 and one non-disruptive discharge, which is chosen randomly between the safe discharges available 1 – 9; $M(2)$ training set will be formed by disruptive discharges number (10, 20) and two non-disruptive discharges, which are chosen randomly between the safe discharges available (1 – 9, 11 – 19); etc.

The test set used for each model correspond to all the discharges produced after the ones used for training. A similar metric than the one used in section 3.3 has been used, the success rate is not divided in valid, early and tardy detections, it is defined as $\frac{\text{disruption predictions}}{\text{disruptive discharges}}$, where the disruption predictions are understood as every disruption detected before the disruption time. The objective is to achieve the highest success rate and the lowest false alarm rate.

Twelve signals are chosen for training and testing models (see table 5.4). The majority or all of these signals have also been used in previous research on disruptions. This work uses two signal sets: an expanded set (that includes all the signals in table 5.4) and a reduced set (that includes the signals from (1) to (7). These signals are the same shown in table 3.4) and three new signals which are the time derivative of three signals from the list.

| Id number | Signal name |
|------------------|---|
| (1) | Plasma current |
| (2) | Mode lock amplitude |
| (3) | Plasma inductance |
| (4) | Plasma density |
| (5) | Diamagnetic energy time derivative |
| (6) | Radiated power |
| (7) | Total input power |
| (8) | Poloidal beta |
| (9) | Plasma vertical centroid position |
| (10) | Plasma inductance time derivative |
| (11) | Poloidal beta time derivative |
| (12) | Plasma vertical centroid position time derivative |

Table 5.4: List of signals

Results

The results for the unbalanced approach with both set of signals are shown in figure 5.14. The expanded set is represented by black lines and the reduced set is represented by grey lines. The x-axis shows the number of disruptive discharges in each model. Each model is tested with the rest of the discharges produced after it is generated. It can be seen that the prediction is unstable until model 24, which determines a clear frontier. In table 5.5 it can be seen the average rates of success and false alarms rate with the standard deviation from model 24 onwards with both expanded and reduced signal sets.

| Signal set | Success rate (%) | False alarm rate (%) |
|------------|-------------------------|-----------------------------|
| Expanded | 85.99 ± 8.43 | 1.37 ± 1.19 |
| Reduced | 85.65 ± 7.78 | 2.27 ± 3.06 |

Table 5.5: Results for unbalanced approach from model 24 onwards

In figure 5.15 it is illustrated the results for the balanced approach with both set of signals. As the previous results, the expanded set is represented by black lines and the reduced set is represented by grey lines. Unlike what happened with the unbalanced approach, this case obtains better results from the first models in terms of stability, it means that the number of oscillation has significantly been reduced. The results are very stable from model 42. In table 5.6 it can be seen the average rates of success and false alarms rate with the standard deviation from model 42 onwards with both expanded and reduced signal sets.

Finally a last approach were developed, an hybrid approach based on starting the generation of models with balanced datasets (to avoid the un-

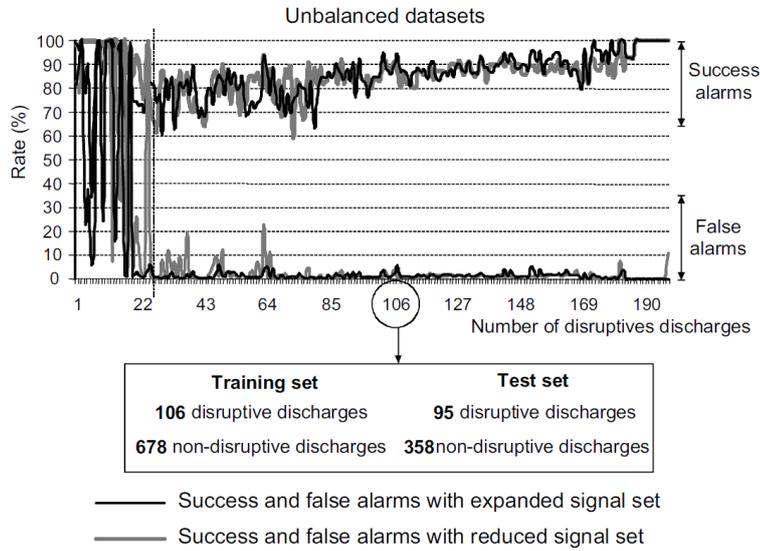


Figure 5.14: Unbalanced approach results

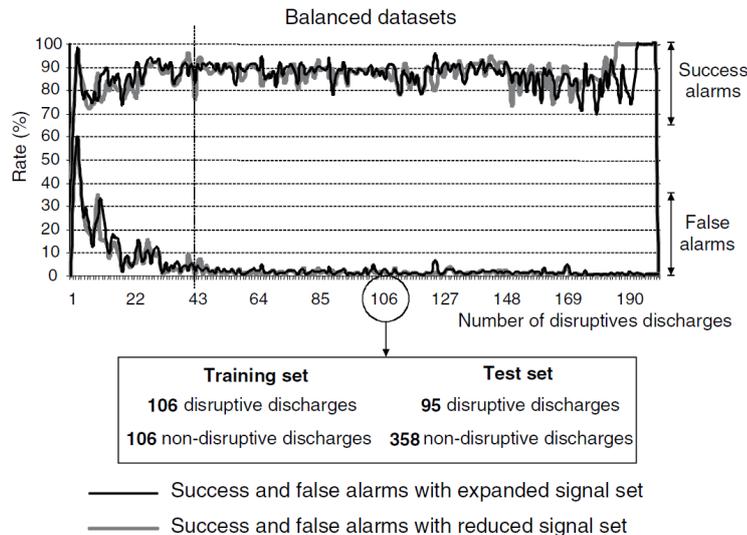


Figure 5.15: Balanced approach results

stable and erratic predictions, see figure 5.14) and switching to unbalanced datasets at a certain point with the aim of keeping the false alarms rate as low as possible. This point establishes the number of non-disruptive discharges which will be used in the training datasets. It means from that point onwards, the training sets will be unbalanced with the same number of non-

| Signal set | Success rate (%) | False alarm rate (%) |
|------------|------------------|----------------------|
| Expanded | 87.63 ± 5.27 | 1.76 ± 1.01 |
| Reduced | 88.80 ± 5.41 | 1.48 ± 0.98 |

Table 5.6: Results for balanced approach from model 24 onwards

disruptive discharges (randomly chosen from the available safe discharges) and a increasing number of disruptive discharges (each new disruption is included in the new model). The switching point has been selected at model 42 due the results are quite stable after model 42 in the balanced approach, see figure 5.15. Therefore the hybrid approach uses balanced datasets until model 42 and unbalanced dataset from that moment, in favour of the disruptive discharges. In table 5.7 it can be seen the average rates of success and false alarms rate with the standard deviation from model 42 onwards with both expanded and reduced signal sets. In figure 5.16, the sharp fall in the success rate at the end is due to the last models use very reduced test sets. For example, model 200 is tested with only one disruptive discharge, then if it is missed the success rate falls to 0%.

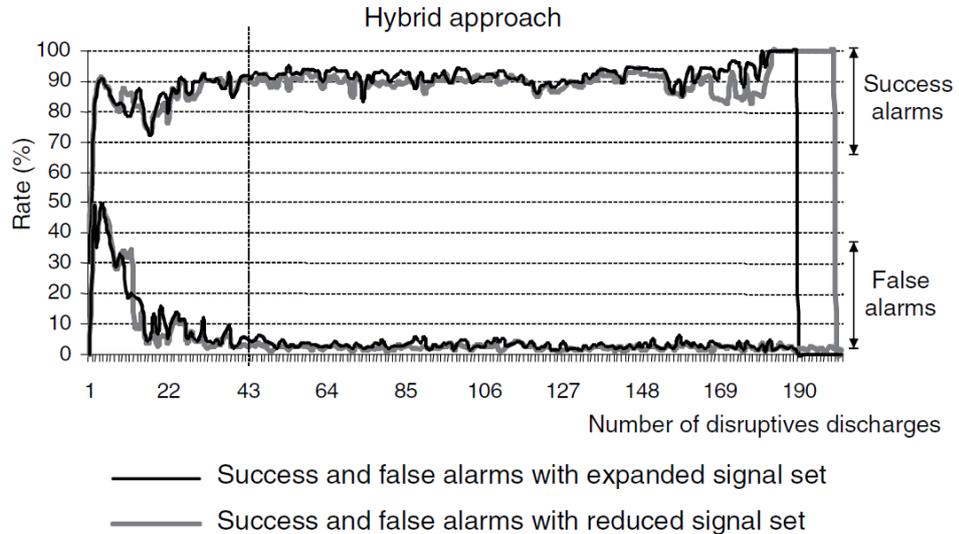


Figure 5.16: Hybrid approach results

| Signal set | Success rate (%) | False alarm rate (%) |
|------------|------------------|----------------------|
| Expanded | 93.03 ± 2.91 | 3.14 ± 1.14 |
| Reduced | 91.27 ± 3.99 | 2.25 ± 0.80 |

Table 5.7: Results for hybrid approach from model 24 onwards

At this point, it is important to discuss that generating a new model every new disruption arrives is not really necessary. If the model identifies correctly the new incoming disruptions, a new model which includes this disruption does not incorporate new knowledge to the classifier. Therefore this training could be avoided by generating a new model when the new incoming disruption is missed, in this way the new model incorporates new knowledge from the missed disruption.

Figure 5.17 shows the success rate of an hybrid approach, using the expanded signal set, that follows the criterion of retraining the system only when a disruption is missed. For the JET ILW campaigns $C28 - C30$, only 10 retrains after disruption 42 have been necessary (instead of 159 developed in figure 5.16). In figure 5.17, each retraining is represented by a circle. The numbers at the top of the figure show the number of disruptive discharges successfully recognized by each predictor before missing a disruption. A total number of 160 disruptions happened and 150 of them have been recognized (93.75%). The average rate of false alarms is 2.79%.

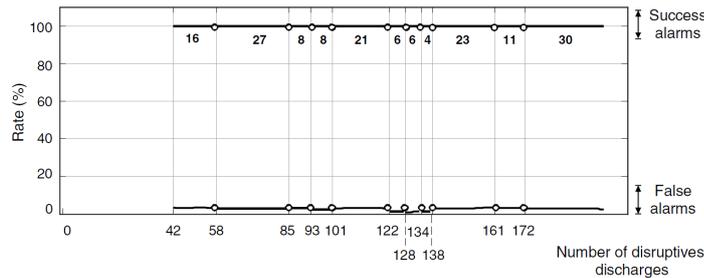


Figure 5.17: Hybrid approach results retraining every missed disruption

5.2.3 Disruption prediction from scratch during ILW campaigns C31-C34

The previous studies about disruption prediction from scratch were implemented over the ILW campaigns $C28 - C30$ (September 2011 - July 2012). In order to extend an analysis from scratch, in this section a predictor from scratch is applied to ILW campaigns $C28 - C30$ (September 2011 - July 2012) and $C31 - C34$ (July 2013 - October 2014). The database from these campaigns is formed by 1036 non-disruptive discharges and 201 non intentional disruptions in campaigns $C28 - C30$; and 1051 non-disruptive discharges and 390 non intentional disruptions during campaigns $C31 - C34$ (table 3.7).

Due to the APODIS structure requires a high computational cost and the one layer predictor developed in section 5.1 shows a great performance with a simpler structure, the last one has been chosen for this analysis.

The best model from section 5.1 is *P2* with 6 features. It is taken the unbalanced approach from section 5.2.2, it means, each predictor from figure will take all the disruptive and non-disruptive discharges available until the last disruptive discharge in the training set.

Firstly, the performances of the models trained from scratch with the unbalanced and balanced approaches shown in section 5.2.2 are compared in campaigns *C28 – C30*. The disruption prediction for the unbalance dataset approach is illustrated in figure 5.18. The x-axis represents the number of disruptive discharges in each model. In order to compare with the previous work in section 5.2.2, the model 24 has been selected as a reference point; however it can be seen in figure 5.18 that there is no erratic and unstable behaviour for the unbalanced approach. On the contrary, balanced approach shows an erratic and unstable behaviour (figure 5.19). Due to the random selection of non-disruptive discharges, the balanced approach has been repeated several times with the same behaviour.

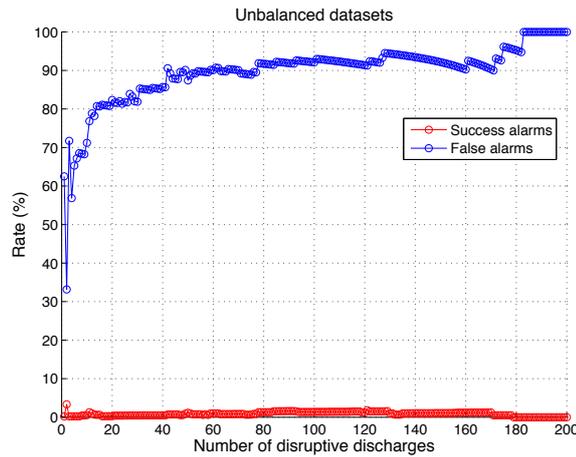


Figure 5.18: Unbalanced approach results with one layer predictor

| Approach | Success rate (%) | False alarm rate (%) |
|------------|-------------------|----------------------|
| Unbalanced | 91.82 ± 3.96 | 0.90 ± 0.49 |
| Balanced | 80.99 ± 11.72 | 2.18 ± 10.62 |

Table 5.8: Results for one layer predictor from scratch from model 24 onwards

Table 5.8 shows the average rates of success and false alarms with the standard deviation from model 24 onwards. The standard deviation in the balanced approach is high because there are several fluctuations, as shown in figure 5.19. Unlike what happened with the balanced datasets, the unbal-

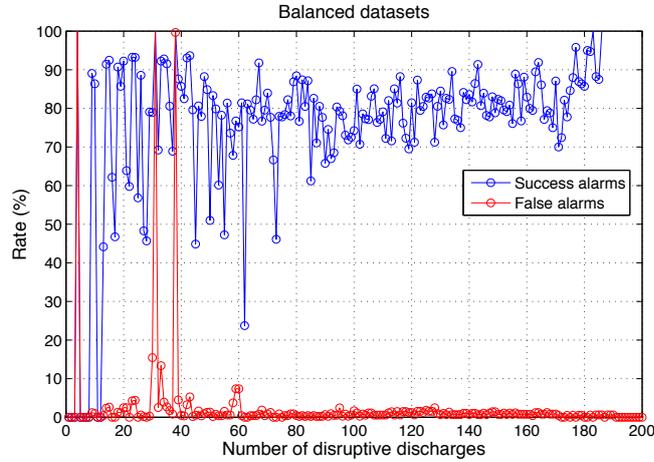


Figure 5.19: Balanced approach results with one layer predictor

anced approach obtains better results in general and from the first models. The results are quite similar than APODIS performance from scratch with the hybrid approach in section 5.2.2. Therefore, the hybrid approach is not necessary to be developed in this study and the unbalanced approach has been chosen for the following analysis.

In ITER views, the first operations will be carried out in hydrogen and helium; helium plasmas have a lower power threshold to create the first ITER H-mode plasmas. At this point, an analysis from scratch taking as the initial campaign the JET campaign *C34* and then the rest ILW campaigns in chronological order *C28* to *C33* is developed to test a possible application to ITER. Therefore this analysis train and test from scratch the previous one layer predictor with the unbalanced approach taking the JET ILW campaigns in the following order: *C34*, *C28* – *C30*, *C31*, *C32*, *C33*.

| Approach | Success rate (%) | False alarm rate (%) |
|------------|------------------|----------------------|
| Unbalanced | 93.40 ± 1.73 | 3.92 ± 13.70 |

Table 5.9: Results for all ILW campaigns from scratch from model 40 onwards

The disruption prediction for all ILW campaigns with the unbalanced dataset approach is illustrated in figure 5.20. It can be seen in figure 5.20 the erratic and unstable behaviour approximately from the beginning until model 40. From model 40 onwards the results are quite stable and this is the number of disruptive discharges to be used as reference. Furthermore the number of non-intentional disruptions is 37 during campaign *C34*, therefore the difference between experiments with deuterium-hydrogen and deuterium-deuterium could be the explanation of the unstable results. The difference

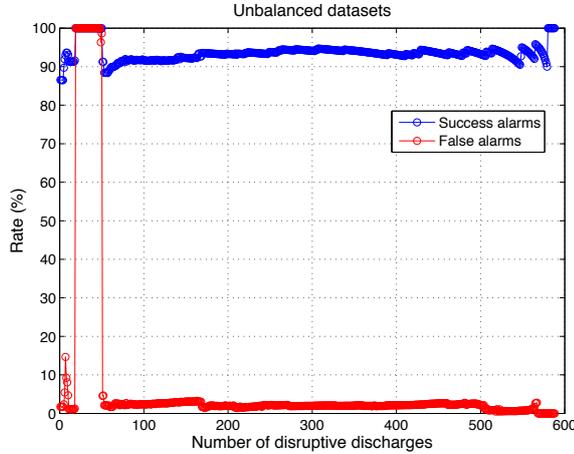


Figure 5.20: Unbalanced approach results from scratch for all ILW campaigns

between campaigns is illustrated in section 5.1.3, where figure 5.9 shows a dataset which is practically linearly separable. On the other hand, figure 5.10 shows an overlap of disruptive and non-disruptive features for campaign *C28 – C30*. In table 5.9 it can be seen the average rates of success and false alarms with the standard deviation from model 40 onwards.

To avoid the unstable and erratic behaviour on campaign *C34*, it has been tested from scratch the dataset formed by campaigns *C28 – C33*. In figure 5.21, it can be seen the results improve and the false alarm rate decays under 2.5% from model 24 onwards. Therefore, this is the model to be used as reference. Table 5.10 shows the average rates of success and false alarms with the standard deviation from model 24 onwards. It can be seen a similar success rate but satisfactory results are obtained from model 24 onwards while the case which includes campaign *C34* obtains stable and high success rate from model 40 onwards. On the other hand, without campaign *C34* the false alarm rate has been significantly improved.

| Approach | Success rate (%) | False alarm rate (%) |
|------------|--------------------|----------------------|
| Unbalanced | $92.76 \pm 2.2.20$ | 1.74 ± 0.70 |

Table 5.10: Results for ILW campaigns *C28 – C33* from scratch from model 24 onwards

To deal with deuterium-hydrogen campaign, a predictor from scratch only for this campaign is developed. At this point, in figure 5.22 it is shown the results of a predictor from scratch with the one layer predictor. Due to the reduced number of discharges in this campaign, it has been followed the same process in section 5.1, 1000 different models have been developed (5

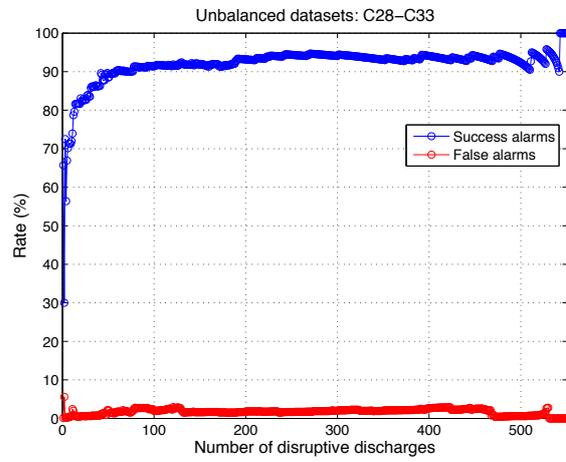


Figure 5.21: Unbalanced approach results from scratch for ILW campaigns $C28 - C33$

values of C and 200 values of γ parameters). Table 5.11 shows the average rates of success and false alarms with the standard deviation.

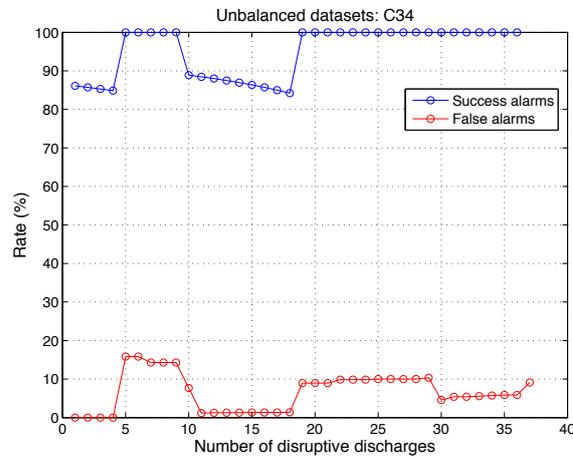


Figure 5.22: Unbalanced approach results from scratch for ILW campaign $C34$

| Approach | Success rate (%) | False alarm rate (%) |
|------------|------------------|----------------------|
| Unbalanced | 95.09 ± 6.69 | 6.67 ± 4.92 |

Table 5.11: Results for ILW campaign $C34$ from scratch

To conclude, a possible approach for ITER could be a predictor from scratch for the first deuterium-hydrogen campaign. Afterwards, a new predictor from scratch for deuterium-deuterium campaigns would be developed without considering data from the previous campaigns. However, in order to reach the best results a selection of the best parameters and samples should be carried out every time a model is trained. This selection could be developed by implementing genetic algorithms.

Chapter 6

Conclusions and Future Work

Conclusions

The work that have been carried out in this thesis have led to the following conclusions:

- An exhaustive and wide analysis of disruptions during the ILW campaigns at JET has been carried out. This analysis has allowed characterised the disruptive information in every discharge and to know the events and actions which have produced this phenomena.
- Large clean and processed databases for disruption prediction have been generated for the ILW campaigns in JET. These databases have been used in several works for disruption prediction and disruption analysis. Furthermore these databases are available to carry out different studies in nuclear fusion.
- Useful tools to analyse and identify important events during discharges have been developed. For example the automatic location of disruption time which allows to identify automatically the disruption time instead of determining it manually.
- The analysis of APODIS performance during the ILW campaigns where the predictor have been working in the real time network. In addition a study of the APODIS robustness and a version of the predictor with a sliding window mechanism have been done. These studies have shown the reliability and robustness of the predictor under signals in failure, identify the most important signals for the predictor and determine the best temporal resolution.
- In view of ITER, several studies have been carried out. Predictors trained and tested with ILW data have been developed and the comparison with an APODIS version trained and tested with the same

data. Following this study it has been studied the disruption prediction from scratch, where it has been done a probabilistic Venn predictor, an APODIS version from scratch and a simpler one layer predictor. All these studies are very useful to generate an efficient predictor for a future device such ITER or DEMO. Furthermore, an important task is gaining physics knowledge about disruptions. At this point it has been developed an analysis about the no-return points.

Future Work

- Data mining on nuclear fusion databases: the understanding of disruption physics is an extremely complex task. On the one hand, this task aims to gain physics knowledge about the disruptive event. Data-driven models, not necessarily focused on disruption prediction, will be deduced from data mining techniques. To this end, a set of data analysis could be accomplished: no-return points, false alarms studies, simulation tools to analyse predictor responses when changing the inputs, probabilistic models,... It is clear that data mining methods demand HPC to analyse and extract knowledge from TB of information (JET, AUG and DIII-D databases). On the other hand it is necessary to apply data mining techniques on nuclear fusion databases in order to deal with massive data and extract knowledge. In this sense several automatic tools can be developed.
- Predictors from scratch and anomaly detections: next generation fusion devices like ITER or DEMO cannot wait for hundreds of disruptions to have reliable predictions to trigger avoidance and mitigation methods. The possibility of avoiding information from past discharges to develop efficient predictors is of big interest. Reliable predictors with very few or even without previous information are necessary.
- Development of physics-based disruption predictors is an important challenge taking into account the lack of plasma theoretical models to handle disruption prediction. Binary classification predictors work well from an engineering point of view because they learn to recognise forthcoming disruptions although the real physics mechanism that triggers the disruptive event remains hidden. Typically, a binary classification predictor is developed for an individual fusion device and is put into operation in that device. Not all quantities that intervene in the predictor can have a direct physics interpretation and also can depend on specific implementations of diagnostics. Therefore, the development of predictors to be used as cross-predictors between different tokamaks is not a straightforward task. As an alternative, physics-based predictors are ideal predictors due to its universal character: models can be generated with data from either one or several tokamaks and are valid

for others. So, the creation of data-driven models based on physics is a possible choice to alleviate the non-existence of theory from first principles about disruptions.

Part I

Appendices

Appendix A

List of Journal Papers & Conferences

Journal Papers

- Dormido-Canto, S., Vega, J., Ramírez, J.M., Murari, A., **Moreno, R.**, López, J.M. and Pereira, A: (2013). *Development of an efficient real-time disruption predictor from scratch on JET and implications for ITER*. Nuclear Fusion 53(11): 113001.
- **Moreno, R.**, Vega, J. and Murari, A. (2014). *Automatic location of disruption times in JET*. Review of Scientific Instruments 85(11): 11D826.
- **Moreno, R.**, Vega, J., Murari, A., Dormido-Canto, S., López, J.M and Ramírez, J.M. (2014). *Robustness and increased time resolution of JET Advanced Predictor of Disruptions*. Plasma Physics and Controlled Fusion 56(11): 114003.
- **Moreno, R.**, Vega, J., Dormido-Canto, S., Pereira, A. and Murari, A. (2015). *Disruption prediction on JET during the ILW experimental campaigns*. Fusion Science and Technology. Submitted for publication.
- Murari, A., Boutot, P., Vega, J., Gelfusa, M., **Moreno, R.**, Verdoolaege, G. and de Vries, P.C. (2013). *Clustering based on the geodesic distance on Gaussian manifolds for the automatic classification of disruptions*. Nuclear Fusion 53(3): 033006.
- Pereira, A., Vega, J., **Moreno, R.**, Dormido-Canto, S., Rattá, G.A. and Pavón, F. (2015). *Feature selection for disruption prediction from scratch in JET by using genetic algorithms and probabilistic predictors*. Fusion Engineering and Design 96-97(907-911) doi:10.1016/j.fusengdes.-2015.04.040.

- Vega, J., Dormido-Canto, S., López, J.M., Murari, A., Ramírez, J.M., **Moreno, R.**, Ruiz, M., Alves, D. and Felton, R. (2013). *Results of the JET real-time disruption predictor in the ITER-like wall campaigns*. Fusion Engineering and Design 88(6-8): 1228-1231.
- Vega, J., Murari, A., Dormido-Canto, S., **Moreno, R.**, Pereira, A. and Acero, A. (2014). *Adaptive high learning rate probabilistic disruption predictors from scratch for the next generation of tokamaks*. Nuclear Fusion 54(12): 123001.
- Rattá, G. A., Vega, J., Murari, A., Dormido-Canto, S. and **Moreno, R.** (2015). *Global optimization driven by genetic algorithms for disruption predictors based on APODIS architecture*. Fusion Engineering and Design. Submitted for publication.

Conferences

- **Moreno, R.** (2014). *Predicción de disrupciones en JET*. Jornadas de investigación y difusión de la Ciencia (Ciclo de Seminarios “Jóvenes Investigadores” 2014), NIPO 721-15-035-1 Depósito Legal M-17056-2015.
- **Moreno, R.**, Vega, J., Dormido-Canto, S., Pereira, A. and Murari, A. (2015). *Overview of disruption prediction on JET during the ILW experimental campaigns*. 1st IAEA Technical Meeting on Fusion Data Processing, Validation and Analysis, Nice, France.
- **Moreno, R.**, Vega, J. and Murari, A. (2014). *Automatic location of disruption times in JET*. 20th Topical Conference on High-Temperature Plasma Diagnostics (HTPD), Atlanta, GA, USA.
- **Moreno, R.**, Vega, J., Murari, A., Dormido-Canto, S., López, J.M. and Ramírez, J.M. (2013). *Robustness of JET Advanced Predictor of Disruptions (APODIS)*. 8th Workshop on Fusion Data Processing, Validation and Analysis, Ghent, Belgium.
- Murari, A., Vega, J., Boutot, P., Cannas, B., Dormido-Canto, S., Fanni, A., López, J.M., **Moreno, R.**, Pau, A. and Sias, G. (2012). *Latest developments in data analysis tools for disruption prediction and for the exploration of multimachine operational spaces*. 24th International Conference on Fusion Energy, San Diego, USA.
- Pereira, A., Vega, J., **Moreno, R.**, Dormido-Canto, S. and Rattá, G.A. (2014). *Feature selection for disruption prediction from scratch in JET by using genetic algorithms and probabilistic predictors*. 28th Symposium on Fusion Technology (SOFT 2014), San Sebastian, Spain.

-
- Rattá, G.A., Vega, J., Murari, A., Dormido-Canto, S. and **Moreno, R.** (2015). *Global optimization driven by genetic algorithms for disruption predictors based on APODIS architecture*. 10th IAEA Technical Meeting on Control, Data Acquisition and Remote Participation for Fusion Research, Ahmedabad, Gujarat, INDIA.
 - Vega, J., Dormido-Canto, S., López, J.M., Murari, A., Ramírez, J.M., **Moreno, R.**, Ruiz, M., Alves, D., and Felton, R. (2012). *Results of the JET real-time disruption predictor in the ITER-like wall campaigns*. 27th Symposium on Fusion Technology (SOFT), Liege, Belgium.
 - Vega, J., **Moreno, R.**, Pereira, A., Dormido-Canto, S. and Murari, A. (2015). *Advanced Disruption Predictor Based on the Locked Mode signal: Application to JET*. 1st EPS Conference on Plasma Diagnostics, Frascati, Italy.
 - Vega, J., **Moreno, R.**, Pereira, A., Murari, A. and Dormido-Canto, S. (2015). *Investigation of plasma dynamics to detect the approach to the disruption boundaries*. 1st IAEA Technical Meeting on Fusion Data Processing, Validation and Analysis, Nice, France.
 - Vega, J., Murari, A., Dormido-Canto, S., Alves, D., Farias, G., López, J.M., **Moreno, R.**, Pereira, A., Ramírez, J.M. and Rattá, G.A. (2014). *Overview of real-time disruption prediction in JET: applicability to ITER*. 41st EPS Conference on Plasma Physics Berlin, Germany.
 - Vega, J., Murari, A., Dormido-Canto, S., **Moreno, R.**, Pereira, A., Ramírez, J.M., López, J.M. and Alves, D. (2013). *Real-time disruption prediction in JET with a view on ITER*. Fusion Reactor Diagnostics Conference, Varenna, Italy.
 - Vega, J., Murari, A., Dormido-Canto, S., **Moreno, R.**, Pereira, A. and Rattá, G.A. (2015). *Disruption precursor detection: combining the time and frequency domains*. 20th IEEE Pulsed Power Conference (PPC) and 26th IEEE Symposium on Fusion Engineering (SOFE) Austin, Texas (USA).
 - Vega, J., Murari, A., Dormido-Canto, S., **Moreno, R.**, Ramírez, J.M., López, J.M., Alves, D., Rattá, G.A. and Pereira, A. (2013). *Real-time prediction of disruptions: results in JET and research lines for ITER*. 8th Workshop on Fusion Data Processing, Validation and Analysis, Ghent, Belgium.
 - Vega, J., Murari, A., **Moreno, R.**, González, S., Pereira, A., Dormido-Canto, S., Ramírez, J.M., López, J.M. and Alves, D. (2013). *Advanced data analysis techniques for event identification and prediction*

in plasma experiments. International Conference on Research and Application of Plasmas, Warsaw, Poland.

- Talebzadeh, S., Murari, A., Gaudio, P., Gelfusa, M., **Moreno, R.**, Lungaroni, M., Peluso, E. and Vega, J. (2015). *From Machine Learning Tools to Mathematical Models: The Case of Disruption Prediction and Avoidance*. 1st IAEA Technical Meeting on Fusion Data Processing, Validation and Analysis, Nice, France.

Appendix B

Parkinson's disease database

Parkinson's disease data set was created by Max Little of the University of Oxford, in collaboration with the National Center for Voice and Speech, Denver, Colorado, who recorded the speech signals. The original study published the feature extraction methods for general voice disorders.

This data set is composed of a range of biomedical voice measurements from 31 people, 23 with Parkinson's disease (PD). Each column in the table is a particular voice measure, and each row corresponds one of 195 voice recording from these individuals ("name" column). The main aim of the data is to discriminate healthy people from those with PD, according to "status" column which is set to 0 for healthy and 1 for PD.

Every instance corresponds to one voice recording. There are around six recordings per patient, the name of the patient is identified in the first column. Further details are contained in the following references (Little et al., 2009, 2007).

The database can be downloaded from the *UCI machine learning repository*¹. The matrix column entries (attributes) are shown in table B.1.

¹<https://archive.ics.uci.edu/ml/datasets/Parkinsons>

| Attribute name | Content |
|---|---|
| name | ASCII subject name and recording number |
| MDVP:Fo(Hz) | Average vocal fundamental frequency |
| MDVP:Fhi(Hz) | Maximum vocal fundamental frequency |
| MDVP:Flo(Hz) | Minimum vocal fundamental frequency |
| MDVP:Jitter(%) MDVP:Jitter(Abs) MDVP:RAP MDVP:PPQ Jitter:DDP | Several measures of variation in fundamental frequency |
| MDVP:Shimmer MDVP:Shimmer(dB) Shimmer:APQ3 Shimmer:APQ5 MDVP:APQ Shimmer:DDA | Several measures of variation in amplitude |
| NHR HNR | Two measures of ratio of noise to tonal components in the voice |
| status | Health status of the subject: Parkinson's - 1, healthy-0 |
| RPDE D2 | Two nonlinear dynamical complexity measures |
| DFA | Signal fractal scaling exponent |
| spread1 spread2 PPE | Three nonlinear measures of fundamental frequency variation |

Table B.1: Parkinson's disease database

Appendix C

Breast cancer database

This breast cancer databases was obtained from the University of Wisconsin Hospitals, Madison from Dr. William H. Wolberg. Samples arrive periodically as Dr. Wolberg reports his clinical cases. The database used here reflects this chronological grouping of the data from January 1989 to November 1991. There are a total of 699 instances with 10 attributes which present 2 possible classes: benign (2) or malignant (4). The attribute information is commented in table C.1. There are 16 instances with missing information on some attributes, therefore the database remains in 683 samples. The class distribution is 67.06%(458) of benign cases and 35.29% of malignant cases.

| Attribute | Domain |
|-----------------------------|-------------------------|
| Sample code number | id number |
| Clump Thickness | 1 – 10 |
| Uniformity of Cell Size | 1 – 10 |
| Uniformity of Cell Shape | 1 – 10 |
| Marginal Adhesion | 1 – 10 |
| Single Epithelial Cell Size | 1 – 10 |
| Bare Nuclei | 1 – 10 |
| Bland Chromatin | 1 – 10 |
| Normal Nucleoli | 1 – 10 |
| Mitoses | 1 – 10 |
| Class: | 2(benign), 4(malignant) |

Table C.1: Breast cancer database

The database can be downloaded from the *UCI machine learning repository*¹. For further details it can be consulted the following references: (Wolberg and Mangasarian, 1990; Mangasarian et al., 1990, 1995; Bennett and Mangasarian, 1992).

¹[https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+\(Original\)](https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Original))

Appendix D

Martingale Test Algorithm

The martingale test algorithm is defined as:

Martingale Test Algorithm (MTA)

Initialize: $M(0) = 1$; $i = 1$; $T = \cdot$.

Set: λ .

1: loop

2: A new example x_i is observed.

3: if $T = \cdot$ then

4: Set strangeness of $x_i := 0$

5: else

6: Compute the strangeness of x_i and data points in T .

7: end if

8: Compute the \hat{p} -values \hat{p}_i using 2.88.

9: Compute $M(i)$ using 2.89.

10: if $M(i) > \lambda$ then

11: CHANGE DETECTED

12: Set $M(i) = 1$;

13: Re-initialize T to an empty set.

14: else

15: Add x_i into T .

16: end if

17: $i := i + 1$;

18: end loop

Bibliography

- ALEDDA, R., CANNAS, B., FANNI, A., SIAS, G. and PAUTASSO, G. Multivariate statistical models for disruption prediction at asdex upgrade. *Fusion Engineering and Design*, vol. 88(6-8), pages 1297–1301, 2013. ISSN 0920-3796.
- AZIZOV, E. A. Tokamaks: from a dâ sakharov to the present (the 60-year history of tokamaks). *Physics-Usp ekhi*, vol. 55(2), page 190, 2012. ISSN 1063-7869.
- BAKER, D. R., SNIDER, R. T. and NAGAMI, M. Observation of cold, high-density plasma near the doublet iii limiter. *Nuclear Fusion*, vol. 22(6), page 807, 1982. ISSN 0029-5515.
- BARLOW, I., BIGI, M., BIRD, J., BONIZZONI, G., BUTTERY, R., CLAY, R., DE BENEDETTI, M., DOBBING, T., GALLAGHER, T., GERVASINI, G., GITTINI, G., HENDER, T. C., HOTCHIN, S., LAST, J., LAZZARO, E., LENNHOLM, M., PERUZZO, S., RICCARDO, V., ROSSI, L., SANTAGIUSTINA, A., STARKEY, D., STEVENS, A., TABELLINI, M., TANGA, A., TERRINGTON, A., VINCENT, M., WATKINS, P., WAY, M. and WILLIAMS, M. The error field correction coils on the jet machine. *Fusion Engineering and Design*, vol. 58-59(0), pages 189–193, 2001. ISSN 0920-3796.
- BENNETT, K. P. and MANGASARIAN, O. L. Robust linear programming discrimination of two linearly inseparable sets. *Optimization Methods and Software*, vol. 1(1), pages 23–34, 1992. ISSN 1055-6788.
- BOSER, B. E., GUYON, I. M. and VAPNIK, V. A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*, pages 144–152. ACM, 130401, 1992.
- BROWN, J. and CHURCHILL, R. *Fourier Series and Boundary Value Problems*. McGraw-Hill, 2001. ISBN 9780072325706.
- BURKE, B. *The world according to wavelets: the story of a mathematical technique in the making*. A. K. Peters, Ltd., 1996. ISBN 1-56881-047-4.

- CANNAS, B., FANNI, A., MARONGIU, E. and SONATO, P. Disruption forecasting at jet using neural networks. *Nuclear Fusion*, vol. 44(1), pages 68–76, 2004a. ISSN 0029-5515 1741-4326.
- CANNAS, B., FANNI, A., MURARI, A., PAU, A. and SIAS, G. Automatic disruption classification based on manifold learning for real-time applications on jet. *Nuclear Fusion*, vol. 53(9), page 093023, 2013. ISSN 0029-5515.
- CANNAS, B., FANNI, A., PAUTASSO, G., SIAS, G. and SONATO, P. An adaptive real-time disruption predictor for asdex upgrade. *Nuclear Fusion*, vol. 50(7), page 075004, 2010. ISSN 0029-5515.
- CANNAS, B., FANNI, A., SONATO, P. and ZEDDA, M. A prediction tool for real-time application in the disruption protection system at jet. *Nuclear Fusion*, vol. 47(11), page 1559, 2007. ISSN 0029-5515.
- CANNAS, B., FANNI, B., SIAS, G., SONATO, P. and ZEDDA, M. Neural approaches to disruption prediction at jet. In *31st EPS Conference on Plasma Physics, ECA*, vol. 28G, page 1.167. 2004b.
- CARRERA, R., HAZELTINE, R. and KOTSCHENREUTHER, M. Island bootstrap current modification of the nonlinear dynamics of the tearing mode. *Physics of Fluids*, vol. 29, pages 899–902, 1986.
- CHANG, Z., CALLEN, J. D., FREDRICKSON, E. D., BUDNY, R. V., HEGNA, C. C., MCGUIRE, K. M., ZARNSTORFF, M. C. and GROUP, T. Observation of nonlinear neoclassical pressure-gradient-driven tearing modes in tftr. *Physical Review Letters*, vol. 74(23), pages 4663–4666, 1995. PRL.
- CHAPMAN, I. T. Controlling sawtooth oscillations in tokamak plasmas. *Plasma Physics and Controlled Fusion*, vol. 53(1), page 013001, 2011. ISSN 0741-3335.
- CHATTERJEE, S. and HADI, A. *Regression Analysis by Example*. Wiley, 2006. ISBN 9780470055458.
- CHERKASSKY, V. and MULIER, F. *Learning from Data: Concepts, Theory, and Methods*. Wiley, 2007. ISBN 9780470140512.
- CORTES, C. and VAPNIK, V. Support-vector networks. *Machine Learning*, vol. 20(3), pages 273–297, 1995. ISSN 0885-6125.
- DASHEVSKIY, M. and LUO, Z. *Reliable probabilistic classification and its application to internet traffic*, pages 380–388. Springer, 2008. ISBN 3540874402.
- DAUBECHIES, I. *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, 1992. ISBN 9780898712742.

- DENDY, R. *Plasma Physics: An Introductory Course*. Cambridge University Press, 1995. ISBN 9780521484527.
- ITER PHYSICS EXPERT GROUP ON DISRUPTIONS, P. C., MHD and EDITORS, I. P. B. Chapter 3: Mhd stability, operational limits and disruptions. *Nuclear Fusion*, vol. 39(12), page 2251, 1999. ISSN 0029-5515.
- DORMIDO-CANTO, S., VEGA, J., RAMÍREZ, J., MURARI, A., MORENO, R., LÓPEZ, J. M. and PEREIRA, A. Development of an efficient real-time disruption predictor from scratch on jet and implications for iter. *Nuclear Fusion*, vol. 53(11), page 113001, 2013. ISSN 0029-5515.
- DOYLE, E. J., GROEBNER, R. J., BURRELL, K. H., GOHIL, P., LEHECKA, T., LUHMANN, N. C., MATSUMOTO, H., OSBORNE, T. H., PEEBLES, W. A. and PHILIPONA, R. Modifications in turbulence and edge electric fields at the l-h transition in the diii-d tokamak. *Physics of Fluids B: Plasma Physics (1989-1993)*, vol. 3(8), pages 2300–2307, 1991.
- DUDA, R. O., HART, P. and STORK, D. G. *Pattern Classification (2nd Edition)*. Wiley-Interscience, 2000. ISBN 0471056693.
- EDDINGTON, A. S. S. *The internal constitution of the stars*. University Press, Cambridge, 1926.
- FARTHING, T., J.W. BUDD, CAPEL, A., COOK, N., EDWARDS, A., FELTON, R., GRIPH, F., JONES, E., LAYNE, R., MCCULLEN, P., MCDONALD, D., WHEATLEY, M., GREENWALD, M., LISTER, J. and VEGA, J. Data management at jet with a look forward to iter. In *ICALEPCS*. 2007.
- FEDOROVA, A., V. GAMMERMAN, NOURETDINOV, I. and VOVK, V. Plug-in martingales for testing exchangeability online. In *29th Int. Conf. on Machine Learning*. 2012.
- FITZPATRICK, R. *Plasma Physics: An Introduction*. Taylor & Francis, 2014. ISBN 9781466594265.
- GAMMERMAN, A. and VOVK, V. Hedging predictions in machine learning - the second computer journal lecture. *Computer Journal*, vol. 50(2), pages 151–163, 2007. ISSN 0010-4620.
- GAMMERMAN, A. AND VOVK, V. AND VAPNIK, V. Learning by transduction. 1998.
- GLENZER, S., MACGOWAN, B. J., MICHEL, P., MEEZAN, N. B., SUTER, L. J., DIXIT, S. N., KLINE, J. L., KYRALA, G. A., CALLAHAN, D., DEWALD, E., DIVOL, L., DZENITIS, E., EDWARDS, M. J., HAMZA, A. V., HAYNAM, C. A., HINKEL, D. E., KALANTAR, D. H., KILKENNY, J. D., LANDEN, O. L., LINDL, J. D., LEPAPE, S., MOODY, J. D., NIKROO,

- A., PARHAM, T., SCHNEIDER, M. B., TOWN, R. P. J., WEGNER, P., WIDMANN, K., WHITMAN, P., YOUNG, B. K. F., VAN WONTERGHEM, B., ATHERTON, L. J. and MOSES, E. I. Symmetric inertial confinement fusion implosions at ultra-high laser energies. *Science*, vol. 327 (5970), pages 1228–1231, 2010.
- VON GOELER, S., STODIEK, W. and SAUTHOFF, N. Studies of internal disruptions and $m = 1$ oscillations in tokamak discharges with soft—x-ray techniques. *Physical Review Letters*, vol. 33(20), pages 1201–1203, 1974. PRL.
- GOLDSTON, R. J. and RUTHERFORD, P. H. *Introduction to Plasma Physics*. IOP Publishing Ltd, 1995. ISBN 0-7503-0183-X.
- GONZÁLEZ, S., VEGA, J., MURARI, A., PEREIRA, A., DORMIDO-CANTO, S. and RAMÍREZ, J. Automatic location of $1/h$ transition times for physical studies with a large statistical basis. *Plasma Physics and Controlled Fusion*, vol. 54(6), page 065009, 2012. ISSN 0741-3335.
- GORBUNOV, E. P. and RAZUMOVA, K. A. Effect of a strong magnetic field on the magnetohydrodynamic stability of a plasma and the confinement of charged particles in the tokamak machine. *Journal of Nuclear Energy. Part C, Plasma Physics, Accelerators, Thermonuclear Research*, vol. 6(5), page 515, 1964. ISSN 0368-3281.
- HARRIS, J. H., CANTRELL, J. L., HENDER, T. C., CARRERAS, B. A. and MORRIS, R. A flexible heliac configuration. *Nuclear Fusion*, vol. 25(5), page 623, 1985. ISSN 0029-5515.
- HENDER, T. C., WESLEY, J. C., BIALEK, J., BONDESON, A., BOOZER, A. H., BUTTERY, R. J., GAROFALO, A., GOODMAN, T. P., GRANETZ, R. S., GRIBOV, Y., GRUBER, O., GRYAZNEVICH, M., GIRUZZI, G., GÄNTNER, S., HAYASHI, N., HELANDER, P., HEGNA, C. C., HOWELL, D. F., HUMPHREYS, D. A., HUYSMANS, G. T. A., HYATT, A. W., ISAYAMA, A., JARDIN, S. C., KAWANO, Y., KELLMAN, A., KESSEL, C., KOSLOWSKI, H. R., HAYE, R. J. L., LAZZARO, E., LIU, Y. Q., LUKASH, V., MANICKAM, J., MEDVEDEV, S., MERTENS, V., MIRNOV, S. V., NAKAMURA, Y., NAVRATIL, G., OKABAYASHI, M., OZEKI, T., PACCAGNELLA, R., PAUTASSO, G., PORCELLI, F., PUSTOVITOV, V. D., RICCARDO, V., SATO, M., SAUTER, O., SCHAFFER, M. J., SHIMADA, M., SONATO, P., STRAIT, E. J., SUGIHARA, M., TAKECHI, M., TURNBULL, A. D., WESTERHOF, E., WHYTE, D. G., YOSHINO, R., ZOHN, H., GROUP, T. I. M. H. D. D. and MAGNET. Chapter 3: Mhd stability, operational limits and disruptions. *Nuclear Fusion*, vol. 47(6), pages S128–S202, 2007. ISSN 0029-5515 1741-4326.

- HO, S. S. Query by transduction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30(9), pages 1557–1571, 2008. ISSN 0162-8828.
- HO, S. S. and WECHSLER, H. Detecting changes in unlabeled data streams using martingale. In *20th International Joint Conference Artificial Intelligence*, pages 1912–1917. 2007.
- HO, S. S. and WECHSLER, H. A martingale framework for detecting changes in data streams by testing exchangeability. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32(12), pages 2113–2127, 2010. ISSN 0162-8828.
- IEA. *Resources to Reserves*. OECD Publishing, 2013. ISBN 978-9264083547.
- JAPKOWICZ, N. and STEPHEN, S. The class imbalance problem: A systematic study. *Intell. Data Anal.*, vol. 6(5), pages 429–449, 2002. ISSN 1088-467X.
- JIAWEI, H., MICHELINE, K. and JIAN, P. *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers Inc., 2011. ISBN 0123814790, 9780123814791.
- JOHNSON, J. L. Stellarator and heliotron devices. *Nuclear Fusion*, vol. 39(2), page 293, 1999. ISSN 0029-5515.
- KEILHACKER, M. H-mode confinement in tokamaks. *Plasma Physics and Controlled Fusion*, vol. 29(10A), page 1401, 1987. ISSN 0741-3335.
- KEILHACKER, M., BECKER, G., BERNHARDI, K., EBERHAGEN, A., ELSHAER, M., FUBMANN, G., GEHRE, O., GERNHARDT, J., GIERKE, G. v., GLOCK, E., HAAS, G., KARGER, F., KISSEL, S., KLUBER, O., KORNHERR, K., LACKNER, K., LISITANO, G., LISTER, G. G., MASSIG, J., MAYER, H. M., MCCORMICK, K., MEISEL, D., MESERVEY, E., MULLER, E. R., MURMANN, H., NIEDERMEYER, H., POSCHENRIEDER, W., RAPP, H., RICHTER, B., ROHR, H., RYTER, F., SCHNEIDER, F., SILLER, S., SMEULDERS, P., SOLDNER, F., SPETH, E., STABLER, A., STEINMETZ, K., STEUER, K. H., SZYMANSKI, Z., VENUS, G., VOLLMER, O. and WAGNER, F. Confinement studies in l and h-type asdex discharges. *Plasma Physics and Controlled Fusion*, vol. 26(1A), page 49, 1984. ISSN 0741-3335.
- KIKUCHI, M. *Frontiers in Fusion Research: Physics and Fusion*. Springer London, 2011 edition, 2009. ISBN 9781849964104.
- KLEVA, R. G. and DRAKE, J. F. Density limit disruptions in tokamaks. *Physics of Fluids B: Plasma Physics (1989-1993)*, vol. 3(2), pages 372–383, 1991.

- KROM, J. G. The evolution of control and data acquisition at jet. *Fusion Engineering and Design*, vol. 43(3-4), pages 265–273, 1999. ISSN 0920-3796.
- LAMBROU, A., PAPADOPOULOS, H., NOURETDINOV, I. and GAMMERMAN, A. *Reliable probability estimates based on support vector machines for large multiclass datasets*, pages 182–191. Springer, 2012. ISBN 3642334113.
- LAYNE, R. and WHEATLEY, M. New data storage and retrieval systems for jet data. *Fusion Engineering and Design*, vol. 60(3), pages 333–339, 2002. ISSN 0920-3796.
- LEONTOVICH, A. *Plasma physics and the problem of controlled thermonuclear reactions*, vol. 1. Pergamon Press, 1961. Page 3,20,31. (Moscow:Izs. AN SSSR, 1958).
- LIPSCHULTZ, B., LABOMBARD, B., MARMAR, E. S., PICKRELL, M. M., TERRY, J. L., WATTERSON, R. and WOLFE, S. M. Marfe: an edge plasma phenomenon. *Nuclear Fusion*, vol. 24(8), page 977, 1984. ISSN 0029-5515.
- LITTLE, M. A., MCSHARRY, P. E., HUNTER, E. J., SPIELMAN, J. and RAMIG, L. O. Suitability of dysphonia measurements for telemonitoring of parkinson’s disease. *Biomedical Engineering, IEEE Transactions on*, vol. 56(4), pages 1015–1022, 2009. ISSN 0018-9294.
- LITTLE, M. A., MCSHARRY, P. E., ROBERTS, S. J., COSTELLO, D. A. E. and MOROZ, I. M. Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection. *BioMedical Engineering OnLine*, vol. 6, pages 23–23, 2007. ISSN 1475-925X.
- LÓPEZ, J. M., VEGA, J., ALVES, D., DORMIDO-CANTO, S., MURARI, A. and RAMÍREZ, J. Implementation of the disruption predictor apodis in jet’s real-time networkd using the marte framework. In *18th REAL-TIME Conference*. 2012.
- MALLAT, S. G. A theory for multiresolution signal decomposition: the wavelet representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 11(7), pages 674–693, 1989. ISSN 0162-8828.
- MALLAT, S. G. *A wavelet tour of signal processing*. Academic press, 1999. ISBN 0080520839.
- MANGASARIAN, O., SETIONO, R. and WOLBERG, W. H. *Pattern recognition via linear programming: Theory and application to medical diagnosis*, pages 22–23. SIAM Publications, Philadelphia, 1990.

- MANGASARIAN, O., STREET, W. N. and WOLBERG, W. H. Breast cancer diagnosis and prognosis via linear programming. *Operations Research*, vol. 43(4), pages 570–577, 1995.
- MARTINEZ, W. L. and MARTINEZ, A. R. *Computational statistics handbook with MATLAB*. Chapman & Hall/CRC Taylor & Francis Group, Boca Raton, 2008. ISBN 9781584885665 1584885661.
- MATTHEWS, G. F. Plasma operation with an all metal first-wall: Comparison of an iter-like wall with a carbon wall in jet. *Journal of Nuclear Materials*, vol. 438, Supplement(0), pages S2–S10, 2013. ISSN 0022-3115.
- MCCRACKEN, G. and STOTT, P. *Fusion: The Energy of the Universe*. Elsevier Science, 2012 edition, 2005. ISBN 9780123846570.
- MINH, H. Q., NIYOGI, P. and YAO, Y. *Mercer's theorem, feature maps, and smoothing*, pages 154–168. Springer, 2006. ISBN 3540352945.
- MIYAMOTO, K. *Fundamentals of Plasma Physics and Controlled Fusion*. K. Miyamoto, 2000.
- MORABITO, F. C., VERSACI, M., PAUTASSO, G. and TICHMANN, C. Fuzzy-neural approaches to the prediction of disruptions in asdex upgrade. *Nuclear Fusion*, vol. 41(11), page 1715, 2001. ISSN 0029-5515.
- MORENO, R., VEGA, J., DORMIDO-CANTO, S., PEREIRA, A. and MURARI, A. Disruption prediction on jet during the ilw experimental campaigns. *Fusion Science and Technology*, vol. Submitted for publication(0), 2015.
- MORENO, R., VEGA, J. and MURARI, A. Automatic location of disruption times in jet. *Review of Scientific Instruments*, vol. 85(11), page 11D826, 2014a.
- MORENO, R., VEGA, J., MURARI, A., DORMIDO-CANTO, S., LÓPEZ, J. and RAMÍREZ, J. Robustness and increased time resolution of jet advanced predictor of disruptions. *Plasma Physics and Controlled Fusion*, vol. 56(11), page 114003, 2014b. ISSN 0741-3335.
- MORRIS, A. W. Mhd instability control, disruptions, and error fields in tokamaks. *Plasma Physics and Controlled Fusion*, vol. 34(13), page 1871, 1992. ISSN 0741-3335.
- MURAKAMI, M., CALLEN, J. D. and BERRY, L. A. Some observations on maximum densities in tokamak experiments. *Nuclear Fusion*, vol. 16(2), page 347, 1976. ISSN 0029-5515.
- MURARI, A. Caracterización de plasmas termonucleares en reactores de fusión: Diagnósticos en jet y herramientas de análisis. In *Ciclo de Conferencias Doctorado UNED*. 2014.

- MURARI, A., BOUTOT, P., VEGA, J., GELFUSA, M., MORENO, R., VERDOOLAEGE, G. and DE VRIES, P. Clustering based on the geodesic distance on gaussian manifolds for the automatic classification of disruptions. *Nuclear Fusion*, vol. 53(3), page 033006, 2013. ISSN 0029-5515.
- MURARI, A., VAGLIASINDI, G., ARENA, P., FORTUNA, L., BARANA, O. and JOHNSON, M. Prototype of an adaptive disruption predictor for jet based on fuzzy logic and regression trees. *Nuclear Fusion*, vol. 48(3), page 035010, 2008a. ISSN 0029-5515.
- MURARI, A., VEGA, J., ALONSO, J. A., DE LALUNA, E., FARTHING, J., HIDALGO, C., RATTÁJ, G. A., SVENSSON, J. and VAGLIASINDI, G. How to extract information and knowledge from fusion massive databases. *AIP Conference Proceedings*, vol. 988(1), pages 457–470, 2008b.
- NOURETDINOV, I., DEVETYAROV, D., BURFORD, B., CAMUZEAX, S., GENTRY-MAHARAJ, A., TISS, A., SMITH, C., LUO, Z., CHERVONENKIS, A. and HALLETT, R. *Multiprobabilistic venn predictors with logistic regression*, pages 224–233. Springer, 2012. ISBN 3642334113.
- PAPADOPOULOS, H. *Inductive conformal prediction: Theory and application to neural networks*. INTECH Open Access Publisher, 2008. ISBN 9537619036.
- PAPADOPOULOS, H. Reliable probabilistic classification with neural networks. *Neurocomputing*, vol. 107(0), pages 59–68, 2013. ISSN 0925-2312.
- PAPADOPOULOS, H., PROEDROU, K., VOVK, V. and GAMMERMAN, A. *Inductive Confidence Machines for Regression*, vol. 2430 of *Lecture Notes in Computer Science*, book section 29, pages 345–356. Springer Berlin Heidelberg, 2002a. ISBN 978-3-540-44036-9.
- PAPADOPOULOS, H., VOVK, V. and GAMMERMAN, A. Qualified predictions for large data sets in the case of pattern recognition. *Proceedings of the International Conference on Machine Learning and Applications*, pages 159–163, 2002b.
- PATTON, R. J., FRANK, P. M. and CLARKE, R. N. *Fault diagnosis in dynamic systems: theory and application*. Prentice-Hall, Inc., 1989. ISBN 0-13-308263-6.
- PAUTASSO, G., EGOROV, S., TICHMANN, C., FUCHS, J. C., HERRMANN, A., MARASCHEK, M., MAST, F., MERTENS, V., PERCHERMEIER, I., WINDSOR, C. G. and ZEHETBAUER, T. Prediction and mitigation of disruptions in asdex upgrade. *Journal of Nuclear Materials*, vol. 290-293(0), pages 1045–1051, 2001. ISSN 0022-3115.

- PAUTASSO, G., TICHMANN, C., EGOROV, S., ZEHETBAUER, T., GRUBER, O., MARASCHEK, M., MAST, K., MERTENS, V., PERCHERMEIER, I., RAUPP, G., TREUTTERER, W. and WINDSOR, C. On-line prediction and mitigation of disruptions in asdex upgrade. *Nuclear Fusion*, vol. 42(1), page 100, 2002. ISSN 0029-5515.
- PEREIRA, A., VEGA, J., MORENO, R., DORMIDO-CANTO, S., RATTÁ, G. A. and PAVÓN, F. Feature selection for disruption prediction from scratch in jet by using genetic algorithms and probabilistic predictors. *Fusion Engineering and Design*, vol. 96-97, pages 907–911, 2015. ISSN 0920-3796.
- POULARIKAS, A. *Transforms and Applications Handbook*. CRC Press, third edition, 2010. ISBN 9781420066531.
- RATTÁ, G., VEGA, J., MURARI, A., VAGLIASINDI, G., JOHNSON, M. and DE VRIES, P. An advanced disruption predictor for jet tested in a simulated real-time environment. *Nuclear Fusion*, vol. 50(2), page 025005, 2010. ISSN 0029-5515.
- RATTÁ, G. A., VEGA, J. and MURARI, A. Improved feature selection based on genetic algorithms for real time disruption prediction on jet. *Fusion Engineering and Design*, vol. 87(9), pages 1670–1678, 2012. ISSN 0920-3796.
- RATTÁ, G. A., VEGA, J., MURARI, A., DORMIDO-CANTO, S. and MORENO, R. Global optimization driven by genetic algorithms for disruption predictors based on apodis architecture. *Fusion Engineering and Design*, vol. Accepted for publication(0), 2015.
- RATTÁ, G. A., VEGA, J., MURARI, A. and JOHNSON, M. Feature extraction for improved disruption prediction analysis at jeta). *Review of Scientific Instruments*, vol. 79(10), page 10F328, 2008.
- REBUT, P. and HUGON, M. *Thermal instability and disruptions in a tokamak*. International Atomic Energy Agency (IAEA), Plasma Physics and Controlled Fusion Research, 1985.
- RUTHERFORD, P. H. Nonlinear growth of the tearing mode. *Physics of Fluids (1958-1988)*, vol. 16(11), pages 1903–1908, 1973.
- SAUNDERS, C., GAMMERMAN, A. and VOVK, V. *Transduction with confidence and credibility*, vol. 1-2 of *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*. Morgan Kaufmann Pub Inc, San Francisco, 1999. ISBN 1-55860-613-0.
- SAVTCHKOV, A. *Mitigation of Disruptions in a Tokamak by Means of Large Gas Injection*. Forschungszentrum, Zentralbibliothek, 2004.

- SCHÖLKOPF, B. and SMOLA, A. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, 2002. ISBN 9780262194754.
- SCHULLER, F. Disruptions in tokamaks. *Plasma Physics and Controlled Fusion*, vol. 37(11A), page A135, 1995. ISSN 0741-3335.
- SENGUPTA, A. and RANJAN, P. Prediction of density limit disruption boundaries from diagnostic signals using neural networks. *Nuclear Fusion*, vol. 41(5), page 487, 2001. ISSN 0029-5515.
- SHAFER, G. and VOVK, V. A tutorial on conformal prediction. *J. Mach. Learn. Res.*, vol. 9, pages 371–421, 2008. ISSN 1532-4435.
- SMOLA, A. and SCHÖLKOPF, B. A tutorial on support vector regression. *Statistics and Computing*, vol. 14(3), pages 199–222, 2004. ISSN 0960-3174.
- SPITZER, L. The stellarator concept. *Physics of Fluids (1958-1988)*, vol. 1(4), pages 253–264, 1958.
- SPITZER, L. *Physics of fully ionized gases*. Interscience Publishers, New York, 1962.
- SPITZER, L. . A proposed stellarator. 1951.
- STEELE, J. *Stochastic Calculus and Financial Applications*. Springer New York, 2001. ISBN 9781468493054.
- STRANG, G. *Introduction to Applied Mathematics*. Wellesley-Cambridge Press, 1986. ISBN 9780961408800.
- TALEBZADEH, S., MURARI, A., GAUDIO, P., GELFUSA, M., MORENO, R., LUNGARONI, M., PELUSO, E. and VEGA, J. From machine learning tools to mathematical models: The case of disruption prediction and avoidance. In *1st IAEA Technical Meeting on Fusion Data Processing, Validation and Analysis*. 2015.
- TEAM, J. *The JET Project (design Proposal)*. Commission of European Communities, 1976.
- THEODORIDIS, S. and KOUTROUMBAS, K. *Pattern Recognition, Fourth Edition*. Academic Press, 2008. ISBN 1597492728, 9781597492720.
- TONKS, L. and LANGMUIR, I. Oscillations in ionized gases. *Physical Review*, vol. 33(2), pages 195–210, 1929.
- VAPNIK, V. *Statistical learning theory*. Wiley, 1998. ISBN 9780471030034.

- VAPNIK, V. *The nature of statistical learning theory*. Springer-Verlag New York, Inc., 2000. ISBN 0-387-94559-8.
- VAPNIK, V. and KOTZ, S. *Estimation of Dependences Based on Empirical Data*. Springer New York, 2006. ISBN 9780387342399.
- VAPNIK, V. N. and CHERVONENKIS, A. Y. *Theory of Pattern Recognition [in Russian]*. Nauka, USSR, 1974.
- VEGA, J., DORMIDO-CANTO, S., LÓPEZ, J., MURARI, A., RAMÍREZ, J., MORENO, R., RUIZ, M., ALVES, D. and FELTON, R. Results of the jet real-time disruption predictor in the iter-like wall campaigns. *Fusion Engineering and Design*, vol. 88(6-8), pages 1228–1231, 2013a. ISSN 0920-3796.
- VEGA, J., MORENO, R., PEREIRA, A., MURARI, A. and DORMIDO-CANTO, S. Investigation of plasma dynamics to detect the approach to the disruption boundaries. In *1st IAEA Technical Meeting on Fusion Data Processing, Validation and Analysis*. 2015a.
- VEGA, J., MURARI, A., DORMIDO-CANTO, S., ALVES, D., FARIAS, G., LÓPEZ, J. M., MORENO, R., PEREIRA, A., RAMÍREZ, J. and RATTÁ, G. A. Overview of real-time disruption prediction in jet: applicability to iter. In *41st EPS Conference on Plasma Physics*. 2014a.
- VEGA, J., MURARI, A., DORMIDO-CANTO, S. and CRUZ, T. Simulations of nuclear fusion diagnostics based on projections with venn predictors and context drift detection. *Annals of Mathematics and Artificial Intelligence*, vol. 74(1-2), pages 223–247, 2015b. ISSN 1012-2443.
- VEGA, J., MURARI, A., DORMIDO-CANTO, S., MORENO, R., PEREIRA, A. and ACERO, A. Adaptive high learning rate probabilistic disruption predictors from scratch for the next generation of tokamaks. *Nuclear Fusion*, vol. 54(12), page 123001, 2014b. ISSN 0029-5515.
- VEGA, J., MURARI, A., DORMIDO-CANTO, S., MORENO, R., RAMÍREZ, J., LÓPEZ, J. M., ALVES, D., RATTÁ, G. A. and PEREIRA, A. Real-time prediction of disruptions: results in jet and research lines for iter. In *8th Workshop on Fusion Data Processing, Validation and Analysis*. 2013b.
- VEGA, J., MURARI, A. and GONZÁLEZ, S. A universal support vector machines based method for automatic event location in waveforms and video-movies: Applications to massive nuclear fusion databases. *Review of Scientific Instruments*, vol. 81(2), page 023505, 2010.
- VEGA, J., MURARI, A., GONZÁLEZ, S., PEREIRA, A. and PASTOR, I. Overview of statistically hedged prediction methods: From off-line to real-

- time data analysis. *Fusion Engineering and Design*, vol. 87(12), pages 2072–2075, 2012. ISSN 0920-3796.
- VEGA, J., MURARI, A., MORENO, R., GONZÁLEZ, S., PEREIRA, A., DORMIDO-CANTO, S., RAMÍREZ, J., LÓPEZ, J. M. and ALVES, D. Advanced data analysis techniques for event identification and prediction in plasma experiments. In *International Conference on Research and Application of Plasmas*. 2013c.
- VOVK, V. *On-line confidence machines are well-calibrated*, pages 187–196. Annual Ieee Symposium on Foundations of Computer Science. IEEE, 2002. ISBN 0272-5428 0-7695-1822-2. Times Cited: 5 43rd Annual IEEE Symposium on Foundations of Computer Science NOV 16-19, 2002 VANCOUVER, CANADA IEEE Comp Soc Tech Comm; Math Fdn Comp; ACM SIGACT 0 5.
- VOVK, V., GAMMERMAN, A. and SAUNDERS, C. Machine-learning applications of algorithmic randomness. In *Proceedings of the Sixteenth International Conference on Machine Learning (ICML-1999)*, pages 444–453. 1999.
- VOVK, V., GAMMERMAN, A. and SHAFER, G. *Algorithmic Learning in a Random World*. Springer, 2005. ISBN 9780387001524.
- VOVK, V., NOURETDINOV, I. and GAMMERMAN, A. Testing exchangeability on-line. In *20th International Conference Machine Learning*, pages 768–775. 2003.
- DE VRIES, B. J. M., VAN VUUREN, D. P. and HOOGWYJK, M. M. Renewable energy sources: Their global potential for the first-half of the 21st century at a global level: An integrated approach. *Energy Policy*, vol. 35(4), pages 2590–2610, 2007. ISSN 0301-4215.
- DE VRIES, P., JOHNSON, M. and SEGUI, I. Statistical analysis of disruptions in jet. *Nuclear Fusion*, vol. 49(5), page 055011, 2009. ISSN 0029-5515.
- DE VRIES, P., JOHNSON, M. F., ALPER, B., BURATTI, P., HENDER, T., KOSLOWSKI, H. and RICCARDO, V. Survey of disruption causes at jet. *Nuclear Fusion*, vol. 51(5), page 053018, 2011. ISSN 0029-5515.
- DE VRIES, P. C., ARNOUX, G., HUBER, A., FLANAGAN, J., LEHNEN, M., RICCARDO, V., REUX, C., JACHMICH, S., LOWRY, C., CALABRO, G., FRIGIONE, D., TSALAS, M., HARTMANN, N., BREZINSEK, S., CLEVER, M., DOUAL, D., GROTH, M., HENDER, T. C., HODILLE, E., JOFFRIN, E., KRUEZI, U., MATTHEWS, G. F., MORRIS, A. W., NEU, R., PHILIPPS, V., SERGIENKO, G. and SERTOLI, M. The impact of the iter-like wall at jet on disruptions. *Plasma Physics and Controlled Fusion*, vol. 54(12), page 124032, 2012. ISSN 0741-3335.

- WAGNER, F., BECKER, G., BEHRINGER, K., CAMPBELL, D., EBERHAGEN, A., ENGELHARDT, W., FUSSMANN, G., GEHRE, O., GERNHARDT, J., GIERKE, G. v., HAAS, G., HUANG, M., KARGER, F., KEILHACKER, M., KLÄBER, O., KORNHERR, M., LACKNER, K., LISITANO, G., LISTER, G. G., MAYER, H. M., MEISEL, D., MÄLLER, E. R., MURMANN, H., NIEDERMEYER, H., POSCHENRIEDER, W., RAPP, H., ROHR, H., SCHNEIDER, F., SILLER, G., SPETH, E., STABLER, A., STEUER, K. H., VENUS, G., VOLLMER, O. and YU, Z. Regime of improved confinement and high beta in neutral-beam-heated divertor discharges of the asdex tokamak. *Physical Review Letters*, vol. 49(19), pages 1408–1412, 1982. PRL.
- WALKER, J. *A Primer on Wavelets and Their Scientific Applications*. Taylor & Francis, 1999. ISBN 9780849382765.
- WEISS, G. M. and PROVOST, F. Learning when training data are costly: the effect of class distribution on tree induction. *J. Artif. Int. Res.*, vol. 19(1), pages 315–354, 2003. ISSN 1076-9757.
- WESSON, J. *The Science of Jet*. Publications Officer, JET Joint Undertaking, Abingdon, Oxon, OX14 3EA, JET Publications Office, 1999.
- WESSON, J. and CAMPBELL, D. *Tokamaks*. OUP Oxford, 3-2011 edition, 1987. ISBN 9780199592234.
- WESSON, J., GILL, R., HUGON, M., SCHÄLLER, F., SNIPES, J., WARD, D., BARTLETT, D. V., CAMPBELL, D., DUPERREX, P. A., EDWARDS, A. W., GRANETZ, R. S., GOTTARDI, N. A. O., HENDER, T. C., LAZZARO, E., LOMAS, P. J., LOPES CARDOZO, N., MAST, K., NAVE, M. F. F., SALMON, N. A., SMEULDERS, P., THOMAS, P. R., TUBBING, B. J. D., TURNER, M. F. and WELLER, A. Disruptions in jet. *Nuclear Fusion*, vol. 29(4), page 641, 1989. ISSN 0029-5515.
- WESSON, J., SYKES, A. and TURNER, M. *Tokamak disruptions*. International Atomic Energy Agency (IAEA), Plasma Physics and Controlled Fusion Research, 1985.
- WINDSOR, C., PAUTASSO, G., TICHMANN, C., BUTTERY, R. and HENDER, T. A cross-tokamak neural network disruption predictor for the jet and asdex upgrade tokamaks. *Nuclear Fusion*, vol. 45(5), page 337, 2005. ISSN 0029-5515.
- WOLBERG, W. H. and MANGASARIAN, O. L. Multisurface method of pattern separation for medical diagnosis applied to breast cytology. *Proceedings of the National Academy of Sciences*, vol. 87(23), pages 9193–9196, 1990.

- WROBLEWSKI, D., JAHNS, G. and LEUER, J. Tokamak disruption alarm based on a neural network model of the high- beta limit. *Nuclear Fusion*, vol. 37(6), page 725, 1997. ISSN 0029-5515.
- YOSHINO, R. Neural-net disruption predictor in jt-60u. *Nuclear Fusion*, vol. 43(12), page 1771, 2003. ISSN 0029-5515.
- YOSHINO, R. Neural-net predictor for beta limit disruptions in jt-60u. *Nuclear Fusion*, vol. 45(11), page 1232, 2005. ISSN 0029-5515.
- ZOHM, H. Edge localized modes (elms). *Plasma Physics and Controlled Fusion*, vol. 38(2), page 105, 1996. ISSN 0741-3335.

