ML4Science X *SPC PROJECTS*: QCE PLASMA REGIME INVESTIGATION

“Simple” explanation of the physical meaning of the dataset parameters

For each experiment, a table is provided.

Inside these tables, for each “instant”, a row of data with different parameters is provided.

The dataset is made of 60 experiments, with around 1s of analyses each, sampled at 10kHz.
So the order of magnitude for the data samples number is 60*1*10.000=600.000.

	1 shotnumber	2 time	3 H98y2calc	4 TS_Te_87_93	5 TS_Ne_87_93	6 TS_Te_93_99	7 TS_Ne_93_99	8 isbaffled	9 ip
1	78068	0.6500	0.9765	63.8451	9.8976e+18	5.5008	5.4440e+17	1	-1.6844e.
2	78068	0.6501	0.9765	63.8451	9.8976e+18	5.5008	5.4440e+17	1	-1.6844e.
3	78068	0.6502	0.9765	63.8451	9.8976e+18	5.5008	5.4440e+17	1	-1.6844e.
4	78068	0.6503	0.9765	63.8451	9.8976e+18	5.5008	5.4440e+17	1	-1.6844e.
5	78068	0.6504	0.9765	63.8451	9.8976e+18	5.5008	5.4440e+17	1	-1.6845e.
6	78068	0.6505	0.9765	63.8451	9.8976e+18	5.5008	5.4440e+17	1	-1.6845e.
7	78068	0.6506	0.9765	63.8451	9.8976e+18	5.5008	5.4440e+17	1	-1.6845e.
8	78068	0.6507	0.9765	63.8451	9.8976e+18	5.5008	5.4440e+17	1	-1.6845e.
9	78068	0.6508	0.9765	108.7578	1.7254e+19	51.7184	9.8330e+18	1	-1.6845e.
10	78068	0.6509	0.9765	108.7578	1.7254e+19	51.7184	9.8330e+18	1	-1.6845e.
11	78068	0.6510	0.9765	108.7578	1.7254e+19	51.7184	9.8330e+18	1	-1.6845e.
12	78068	0.6511	0.9765	108.7578	1.7254e+19	51.7184	9.8330e+18	1	-1.6845e.
13	78068	0.6512	0.9765	108.7578	1.7254e+19	51.7184	9.8330e+18	1	-1.6845e.
14	78068	0.6513	0.9765	108.7578	1.7254e+19	51.7184	9.8330e+18	1	-1.6845e.
15	78068	0.6514	0.9765	108.7578	1.7254e+19	51.7184	9.8330e+18	1	-1.6845e.

shotnumber

Integer.
Reference number for each experiment.

time

Measured in seconds (from the start of the experiment).
The time identifying the sampled data points. It is in between 0.650s, so that we’re sure our plasma is in a diverted configuration[1], and 1.800s (or earlier, if the plasma disrupts[2] before)

H98y2calc

Adimensional.
A metric to evaluate the “quality” of the confinement in the plasma. The higher, the better.

TS_Ne_xx_yy & TS_Te_xx_yy

Measured in (meters)^-3 and electronvolt.
Density and temperature of the plasma at specific locations, identified by flux coordinates[3], in the range [0.xx, 0.yy], as evaluated from the Thomson scattering diagnostic.

isbaffled

Boolean.
If, during the experiment, baffle structures[4] were present (=1) or not (=0).

ip

Measured in Ampere.
The current driven inside the plasma, forming the poloidal magnetic field[5].

b0

Measured in Tesla.
The toroidal magnetic field[5] intensity (on axis).

nel

Measured in (meters)^-3.
The plasma density, as evaluated from a linear average along the line of sight of the Thomson scattering diagnostic[6].

ptot

Measured in Watt.
The total power supplied to the plasma, as the sum of the Ohmic power (generated by the current inside the plasma through the Joule effect) and neutral-beam heating systems[7].

pdiv & pmid

Measured in Pascal.
The neutral pressure inside the vacuum vessel, as measured by pressure gauges, in the divertor region and in the main chamber[4].

q95 & betan & kappa & deltaavg & deltaupp & deltalow

Adimensional.
Adimensional parameters which define the geometry (the shape) of the magnetic “cage”[8] in which we confine our plasma.

gapin & gapout

Measured in meters.
Gaps between the plasma and the inner / outer wall.

zmag & rmag

Measured in meters.
Vertical (measured from the horizontal midplane of the machine) and radial position (measured from the axis of the torus) of the center of the plasma.

rmin

Measured in meters.
Characteristic radius (measured from its center, to the end of the confined region) of the plasma.

lpar_ot & lpol_io

Measured in meters.
Characteristic lengths connecting the outer midplane of the plasma to the outer target in the direction parallel to the total magnetic field (*lpar_ot*), and the inner midplane to the outer midplane in the direction along the poloidal component of the magnetic field (*lpol_io*)[9].

prad

Measured in Watt.
The total power radiated away from the plasma.

zeff

Adimensional.
Effective charge of the plasma[10].

nu_sol

Adimensional.
It measures the collisionality (in the fluid sense) of plasma at the edge of the confined region.

alpha

Adimensional.
A parameter which tries to capture which turbulence regime mechanism is dominating on the edge of the confined region (*i.e.* which physical phenomena is regulating the plasma "escape" from the confined region). This is properly defined in A. Stagni *et al*, Nucl. Fusion 62 (2022) 096031.

LHD_label
* Still TBD if this will be provided in the same table or in an additional table.

Integer.
A label, identifying in which regime[11] the plasma is (1 = L-mode, 2 = QCE H-mode, 3 = ELMy H-mode).

Useful definitions

1. *Diverted configuration*
   A plasma in a tokamak can be either in a limited (*i.e.* “squashed on the wall”) or diverted (*i.e.* “far from the wall”) configuration. The diverted configuration is called in this way as it presents a point in which the magnetic field lines “divert”, effectively producing two “legs” which are connecting the confined plasma and the wall. Usually, all the analyses of interest (and most of the routines used to fetch the data above) are done for diverted configurations. However, at the beginning of each experiment, there is a phase in which the plasma is in the limited configuration, for technical reasons.

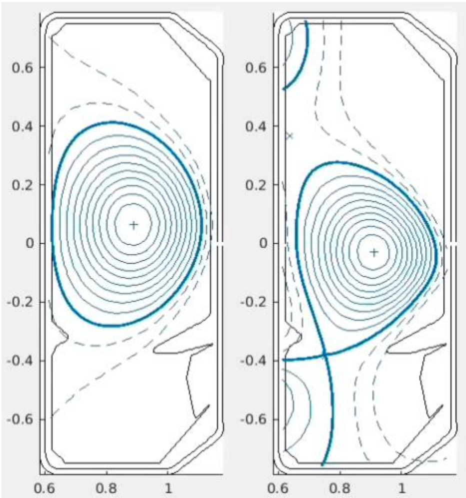


Figure 1 - Limited (left) and diverted (right) plasma configurations from the same experiment

2. *Disruptions*
   Sometimes, an experiment ends abruptly due to a very fast and strong instability in the plasma, leading to its disruption and complete loss of confinement.

3. *Flux coordinates*
   In plasma physics, it is often used the flux (ψ) coordinate system to identify positions inside the machine. Flux coordinates are coordinates that assign a label to each magnetic flux surface, starting from 0 at the center of the magnetic “cage”, up to 1 at the interface between the confined (closed flux surfaces) and unconfined (open flux surfaces, *i.e.* not closed on themselves) regions, and beyond in unconfined regions.

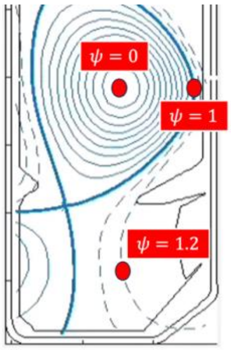


Figure 2 - Labels for different closed (solid line) and open (dotted line) magnetic flux surfaces

4. <u>*Baffle structures*</u>
   These are material structures we might (or might not) have in our experiments. They might come in different sizes, but overall they provide a better separation of the main chamber (region above the baffles, where the confined plasma is) and the divertor chamber (region below the baffles, where the plasma "legs" are).

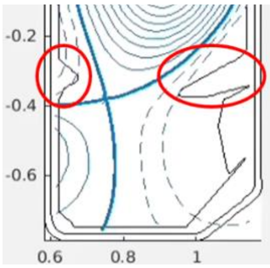


Figure 3 - Highlight of baffle structures

5. <u>*Magnetic field components*</u>
   The magnetic field in a tokamak looks like a "twisted" helix. This is ultimately the combination of a magnetic field generated in the toroidal direction (by external currents flowing in coils) and a magnetic field generated in the perpendicular (*i.e.* poloidal) direction (by the current carried by the plasma itself).

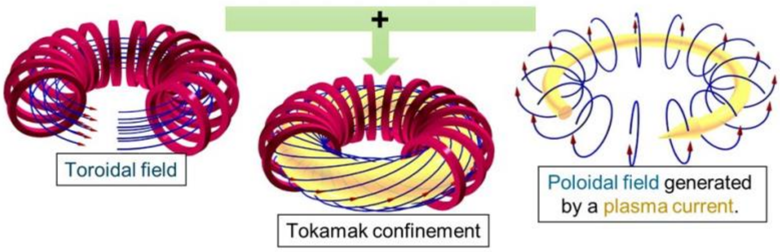


Figure 4 - Magnetic fields in a tokamak

6. <u>*Thomson scattering diagnostic*</u>
   A fundamental diagnostic to measure, with high spatial resolution, density and temperature in the plasma. A laser is shot vertically inside the plasma (line of sight direction) and its photons are scattered and captured by an array of cameras. This acquisition is then used to reconstruct the plasma temperature and density.

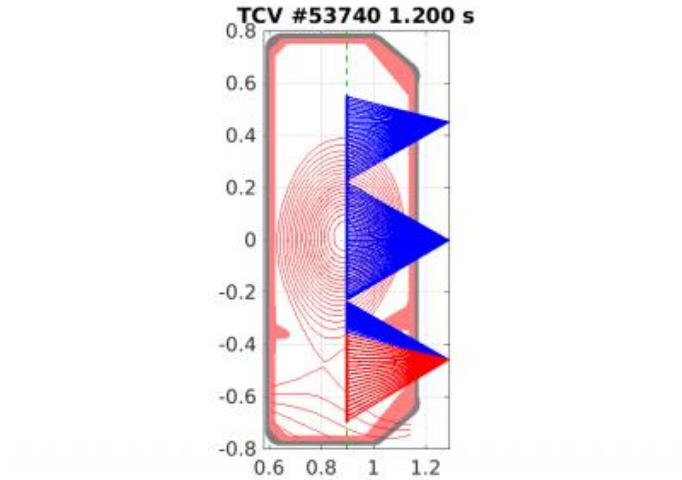


Figure 5 - Thomson scattering system in TCV

7. *Heating systems*
   3 main heating systems are used in a tokamak: Ohmic heating (the plasma is heated by the current passing through it via the Joule effect), resonant wave absorption (an electromagnetic wave is sent through the plasma and absorbed), neutral-beam injectors (very fast – high kinetic energy – neutral particles are shot in the plasma, where they ionize and become part of the plasma itself, increasing its energy).

8. *Plasma shapes*
   The magnetic "cage" confining the plasma (and therefore, the plasma itself) can be shaped in different ways, in order for it to be more or less elongated, larger or smaller, more or less triangular, … This has a influence on the plasma performance. We use different adimensional parameters to quantify this.

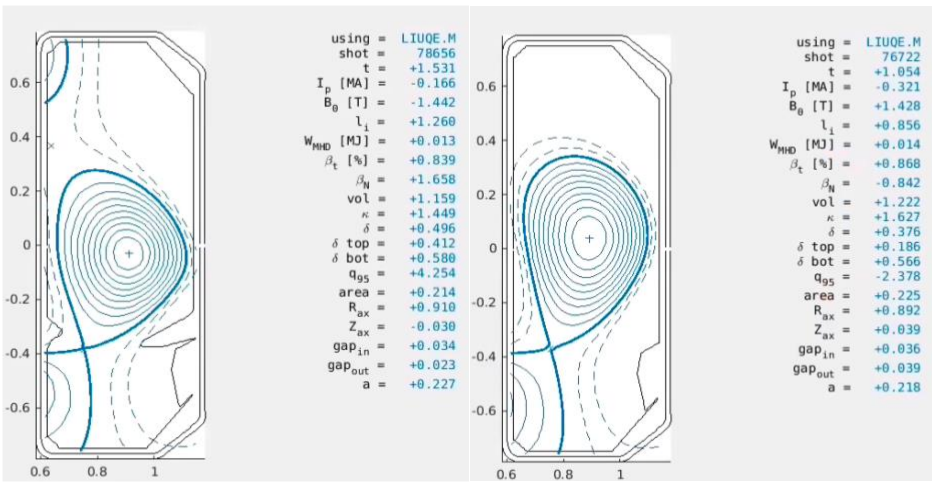
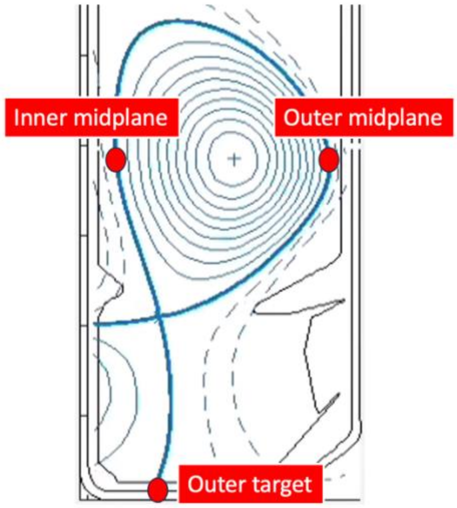


Figure 6 - Two different plasma shapes

9. *Some points of interest in the plasma configuration*



10. *Effective charge of the plasma*

    Our plasmas are often made of hydrogen isotopes, which have 1 and only electron. Hence their net charge, when ionized, is always just +1.

    However, when impurities are present in the plasma (*e.g.* Carbon, coming from the wall of the machine), they might have also higher ionization states (*e.g.* Carbon has 6 electrons). The effective charge of the plasma is then a sort of weighted average of the ionization state of the plasma: if it's close to 1, it means we have a very low impurity content; if it's higher than 1, it means we have a significant content of impurities.

11. *Confinement regimes*

    Different types of confinement can be achieved in a plasma. These can be of low efficiency (L-mode) or high efficiency (H-mode). H-mode plasmas can be characterized by the presence of strong and violent particle and energy bursts called ELMs (ELMy H-mode) or by their absence* (QCE H-mode).

    * It is more correct to say that QCE still has ELMs, but at a much higher time frequency with much lower amplitude, as if it was a quasi-continuous release.

Goal of the project

The goal is to investigate the conditions which favor the access to the QCE regime in plasmas during experiments in the TCV tokamak.
Two following approaches are proposed (depending on the learning method assigned):

Proposed approaches		
Supervised learning We train a model to try to predict, given a set of inputs, in which of the three possible confinement regimes our plasma will be (L-mode, ELMy H-mode, QCE H-mode). The **labels** are provided in the column named **LHD_label**.	*Specifics*: use **machine** inputs only (see below). Corentin Genton Timo Michoud Claire Koechlin	*Specifics*: use **physical** inputs only (see below). Johannes Grafen Paul Hofmeyer Raphael Hellmann
Unsupervised learning We train a model to try to highlight some patterns in the dataset, and later we try to look for correlations to the access to the QCE regime (using the LHD_label data column).	*Specifics*: use **machine** inputs only (see below). Hugo Majerczyk Eric Saikali Aymeric de Chillaz	*Specifics*: use **physical** inputs only (see below). Lucien Chapuisat Adrien Bouquet Nathan Felber

Division of the dataset		
Machine inputs The type of quantities we can control a priori when designing an experiment	isbaffled ip b0 nel ptot pdiv q95 betan kappa deltaavg	deltaupp deltalow gapin gapout zmag rmag rmin lpar_ot zeff
Physical inputs A mix of quantities we can and can't control a priori when designing an experiment, but all with significant meaning in terms of the physics of the QCE regime	H98y2calc TS_Te_87_93 TS_Ne_87_93 TS_Te_93_99 TS_Ne_93_99 nel ptot pmid	q95 betan lpar_ot lpar_io prad zeff nu_sol alpha

Please note that these proposed approaches are flexible. Your input (as well as the input of the SPC ML team) on the matter is very welcome, both on the method to apply and the dataset to analyze. Don't be afraid to provide ideas on your field of expertise.