Modelling of Plasma Confinement State Transitions using Hybrid Physics-based and Data Driven Approaches

C. Venturini[1], A. Pau[1], Y. Poels[1,2], M. Carpita[1], C. Donner[3], V. Menkovski[2], G. Romanelli[3], O. Sauter[1], E. Solano[4], TCV team[a] and JET contributors[b] and the EUROfusion tokamak exploitation team[c]

[1] *EPFL-SPC, CH-1015 Lausanne, Switzerland*

[2] *Eindhoven University of Technology, Eindhoven, The Netherlands*

[3] *Swiss Data Science Center, ETH Zurich, 8006 Zurich, Switzerland*

[4] *Laboratorio nacional de Fusión, CIEMAT, Madrid, Spain*

[a] *See the author list of H. Reimerdes et al 2022 Nucl. Fusion 62 042018*

[b] *See the author list of C. F. Maggi et al 2024 Nucl. Fusion in press*

[c] *See the author list of E. Joffrin et al 2024 Nucl. Fusion in press*

Introduction

To operate routinely a tokamak in high performance regimes advanced plasma control and exception handling are essential components. In the context of high performance regimes, a relevant factor is the so-called confinement state of the plasma, which influences stored energy and leads to characteristic physics behaviours in plasma dynamics, as well as affecting the transport properties. Moreover, information on transitions between confinement states is key to aid in the context of disruption avoidance, where such transitions are recognised as potential precursors to disruptions [1]. The plasma's confinement state can be classified in various ways, and amongst the main ones we can recognize two different modes: low confinement (L) and high confinement (H). Additionally, we can also consider a third confinement state called dithering (D) which represents an intermediate oscillating phase. Different kinds of H mode scenarios, commonly known as small ELMs or ELMs free regimes are also of great interest [2]. Being able to automatically label different confinement states is of great help both for post shot analysis and for real time control of the discharge. Previous work for automated labeling of plasma confinement state has been carried out, resulting in the development of ML models that have relatively good accuracy and are currently integrated in the control system of TCV [3][4]. Nonetheless, an interesting point in the context of this task is the uncertainty related to the prediction and the interpretability of the models decision of assigning a certain label. This is important as to not have the control system trusting a wrong output (real time setting), or so that the off-line automated labeling can be augmented with a quantification of the uncertainty. The ideal situation would be to employ a model that is both highly expressive and completely interpretable. In practice, we usually make a compromise between the two. A good example of such compromise is represented by some latent variable models, such as generative topographic mapping (GTM) [5], that allows to reach a good level of both expressiveness and interpretability. For this model, we can harness the interpretability given by the latent space representation, which can be easily visualisable since it has a lower dimension. This opens a vast array of possibilities such as exploring how the distribution of the different features looks like in this reduced informative space, as well as investigating how a discharge trajectory evolves. Once the model has been fit to data, any arbitrary discharge can be projected in the latent space and examined. In this report, we introduce preliminary results on the study of QCE [6] discharges from TCV, and on what the model can tell us about it. More in general, the GTM block is part of a more advanced hybrid modeling scheme for plasma confinement state classification [7] that combines in a single pipeline a deep neural network (NN), a GTM and rule based (RB) decision thresholds, as to harness the positive characteristics of each.

Dataset

The model is fit using data from the Tokamak à Configuration Variable (TCV) [8]. The full dataset is composed of 270 discharges, which have been hand labeled by domain experts. For each time slice, a label corresponding to the plasma confinement state (L, D, H) is added. Of all the discharges, 240 are used to construct the embedding of the GTM and 30 are kept aside for testing purposes. We moreover employ a set of QCE discharges, also fully labeled by a domain expert. In this case considered labels are (L, H, QCE). The features employed for the fitting of the model can be found in Table 1. Data access is carried out through the DEFUSE tool, a framework supporting the development of the EUROfusion Disruption Database (EuDDB) [9].

Feature	Source	Description
FIR_core	FIR	Averaged line-integrated density at the core, considering channels between $0.87m$ and $0.91m$
WP	LIUQE	Total plasma stored energy
DML	DML	Plasma toroidal flux
PD_fft_4ms	PD	FFT of channel 13 of PD, then compute windowed standard deviation, stride 1, window of $4ms$
PD_fft_80ms	PD	Same as above, window of $80ms$
BETAP	LIUQE	β_p, poloidal β parameter
LI	LIUQE	Internal inductance

Table 1: Features employed for GTM model fitting

Preliminary results

After fitting the model on the dataset of ELMy H mode TCV shots, we perform a projection of the QCE shots on the latent space of the GTM (Figure 1). Since these shots have not been part of the training process, the model has no notion of QCE and as such it is interesting to investigate if by looking at the projection of these shots in latent space we can infer