Augmenting the extrapolation of disruption prediction to extended parameter regimes by predict-first neural network

Z. Y. Yang[1] , W. L. Zhong[1, *], J. Y. Li[1], Y. H. Chen[1], D. Li[1], J. Z. Zhang[1], X. Fan[1], B. Li[1], Y. P. Zhang[1], Y. B. Dong[1], J. Artola[2], T. F. Sun[1], Z. H. Xu[1], R. S. Qiu[1], X. Sun[1] and J. R. Wen[1]

1. Southwestern Institute of Physics, Chengdu 610041, China

2. ITER Organization, route de Vinon sur Verdon, 13067 St Paul Lez Durance, France

E-mail: zy-yang@swip.ac.cn

ID: 3178

Background

- Plasma disruption is a critical challenge to the safe operation of tokamak devices.
- The current primary approach to managing disruptions follows the paradigm of prediction–avoidance–mitigation.
- However, for future fusion reactors, it remains an underexplored issue how to develop prediction algorithms based on low-parameter operational data from early stages and subsequently apply them for disruption warning in high-parameter operational regimes.
- This study utilizes the HL-3 tokamak to test the extrapolation capability of disruption prediction algorithms across different parameter regimes and conducts targeted optimizations.

HL-3 Dataset

- Since its initial plasma discharge in 2020, HL-3 has continuously elevated its operational parameters, establishing an broad dataset.
- Up to now, the dataset comprises 3819 validated discharges. Figure 1 illustrates the distribution relationships among the maximum plasma current (I_p), average toroidal field strength (B_t), and maximum normalized beta (β_N).

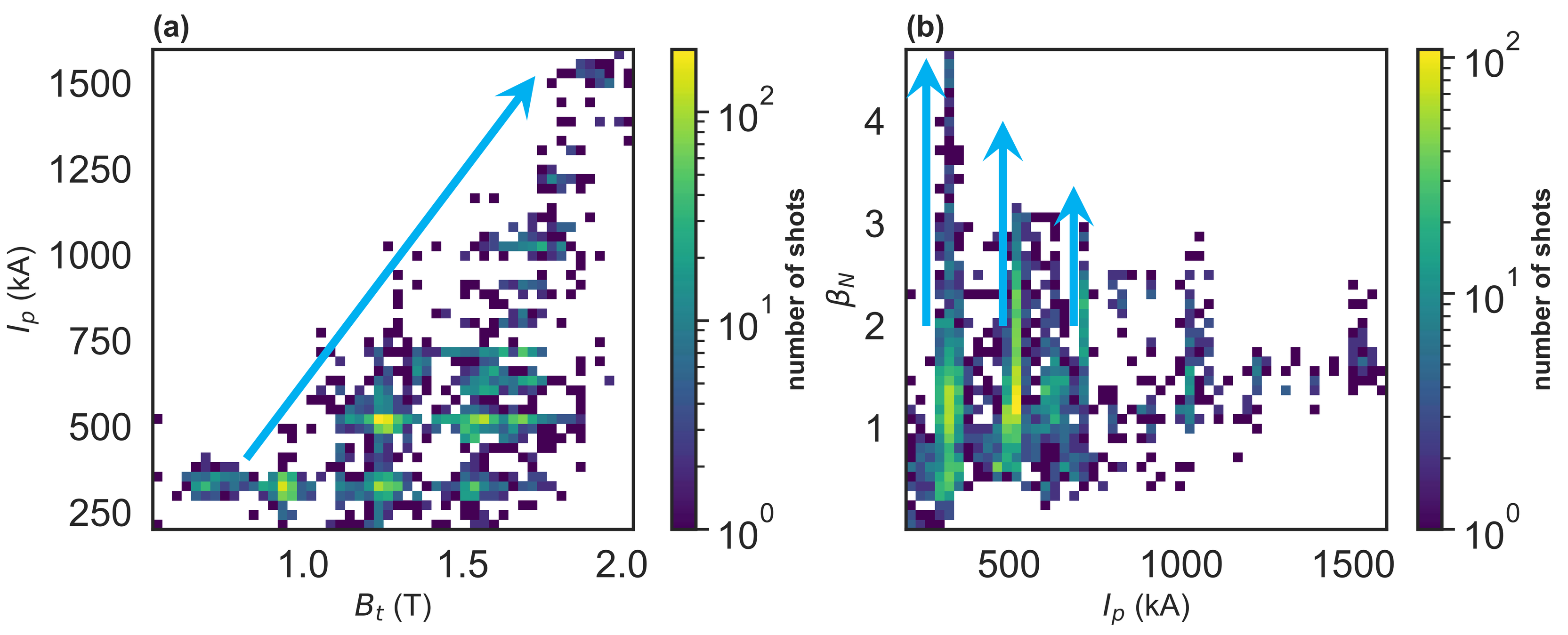

Figure 1. Distribution of plasma current, toroidal field, and beta normal in the HL-3 disruption database.

- The dataset is divided into 5 × 4 sub datasets according to the parameter regimes of plasma, as shown in Table 1.
- The sub datasets with ID:1 are used to train and validate the disruption prediction models. And sub datasets with ID:2~4 are used to test the extrapolation capability.

Table 1 The parameter scope and number of shots within each sub datasets.

Parameter	Range	Number of shots	Dataset ID
I_p (kA)	0~520	2570	1
	520~700	748	2
	700~1000	281	3
	1000~1700	220	4
B_t (T)	0~1.55	2652	1
	1.55~1.7	769	2
	1.7~1.78	253	3
	1.78~2.2	141	4
q_{95}	4.5~10	2495	1
	4~4.5	870	2
	3.7~4	292	3
	1.8~3.7	159	4
W_e (kJ)	0~160	2545	1
	160~220	871	2
	220~300	252	3
	300~700	148	4
β_N	0~1.5	2485	1
	1.5~2.1	800	2
	2.1~2.5	324	3
	2.5~5	207	4

PFNN: Predict-First Neural Network

- As HL-3's operational parameters continue to rise, disruption predictors must adapt to avoid frequent failures. Our Parameter Preview Neural Network (PPNN) addresses this by first forecasting plasma evolution, then using the prediction-reality discrepancy as a disruption criterion. This logic shift enables superior extrapolation across parameter regimes.

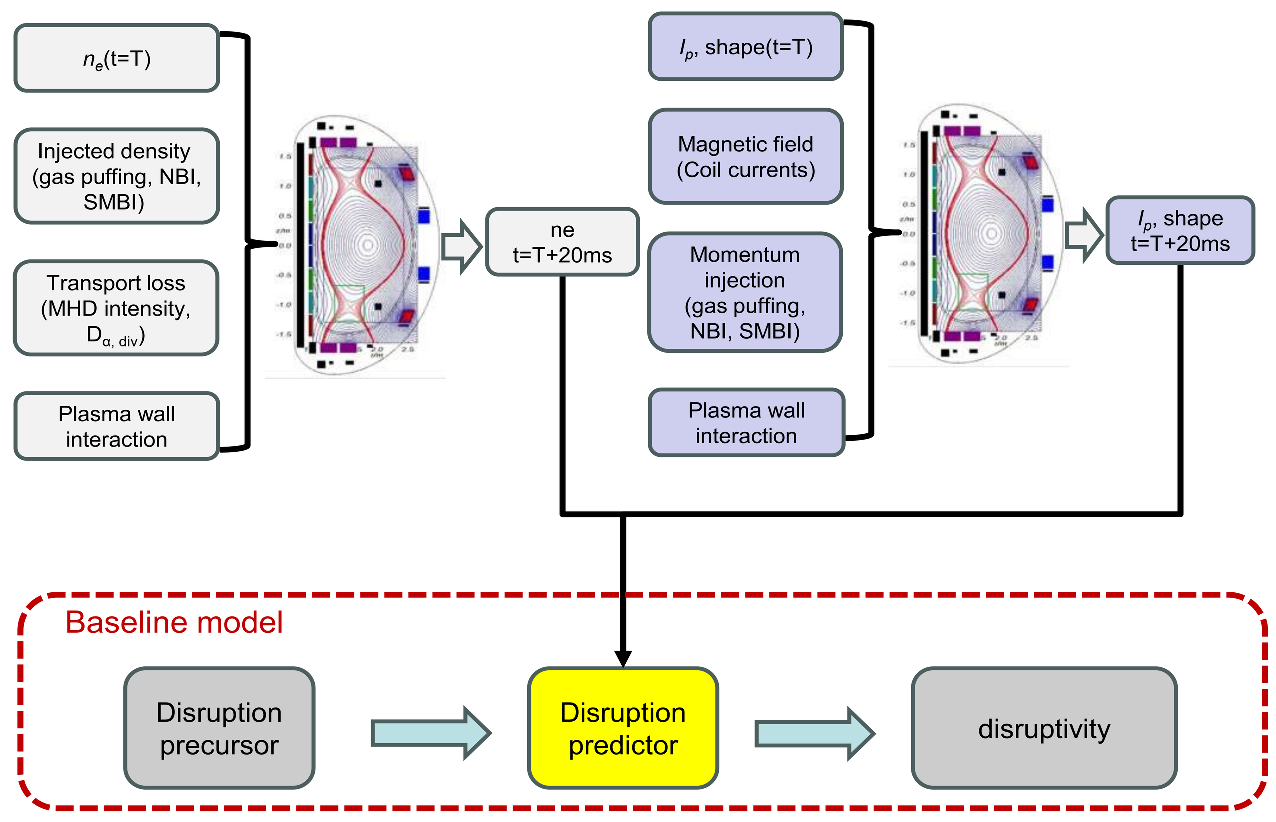

Fig 2 the overall architecture of the PFNN model.

- Figure 3 shows the cross-regime testing results across key parameter intervals. The PFNN substantially outperforms the baseline in prediction accuracy.

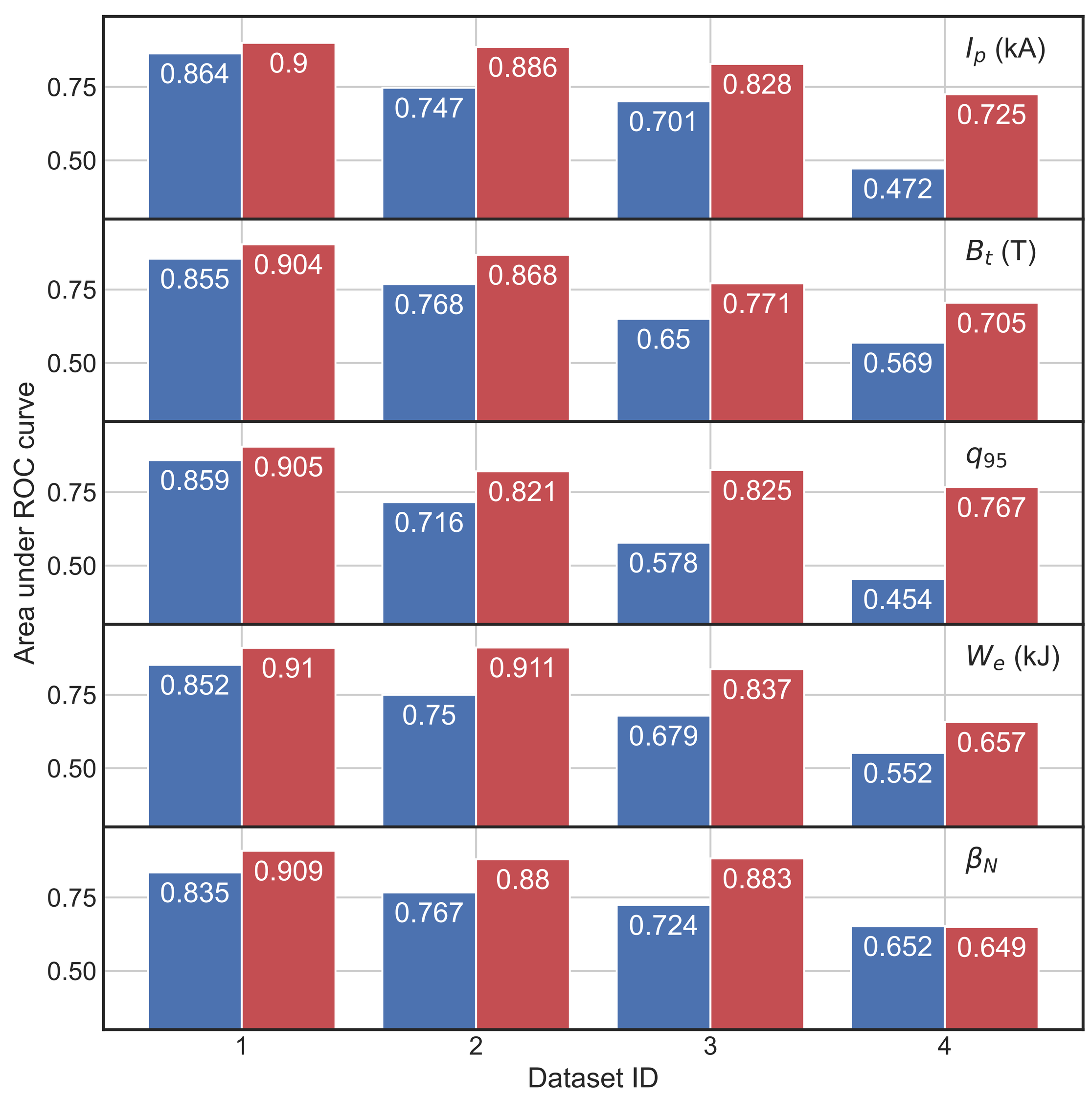

Fig 3 When trained on low parameter regimes and tested on high regimes, the baseline model's accuracy (blue) declines rapidly with extrapolation distance, while the PPNN (red) generalizes effectively.

Real-Time Prediction and Mitigation

- Supported by PFNN, HL-3 has achieved real-time disruption prediction and mitigation. The algorithm trained on dataset of which the maximum β_N is 3.3, maintained stability even when the β_N was pushed beyond 4. Across a series of 22 consecutive high β_N discharges, the system demonstrated a reliability rate of 95.5%.

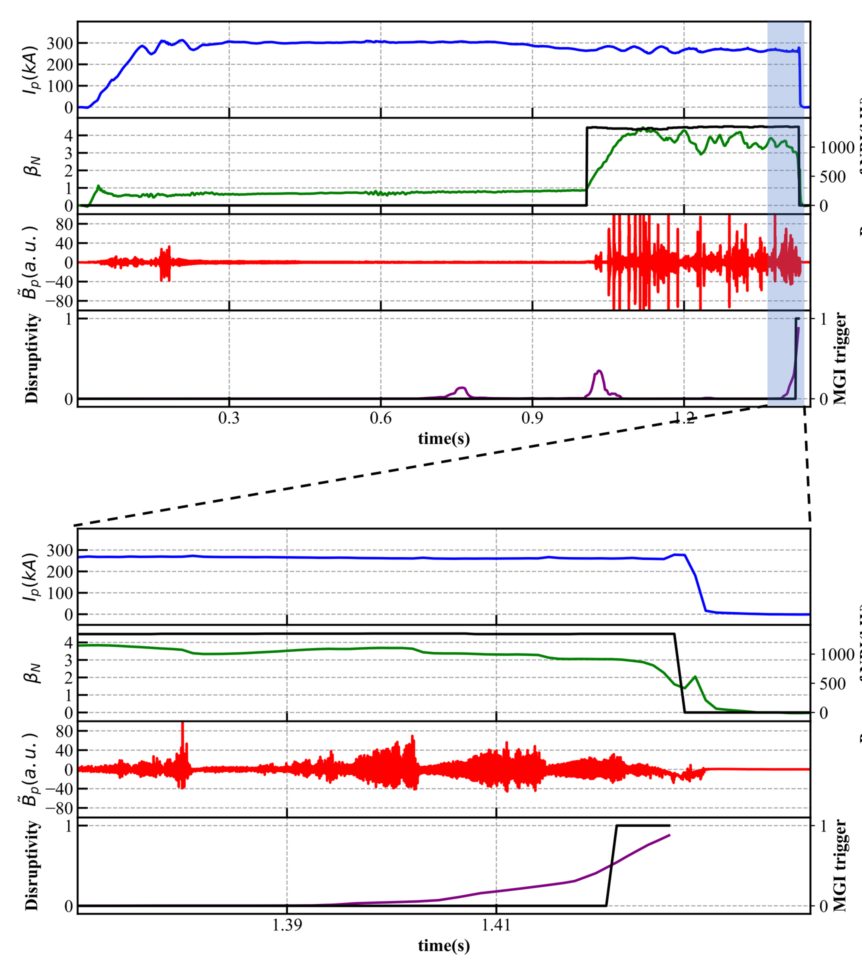

Fig4 During Shot 12478, the plasma β_N increased with NBI heating, approaching 4. Concurrently, the B_p signal showed growing MHD instabilities. After t=1.4s, β_N decayed and confinement degraded, prompting the AI to trigger MGI for mitigation.