Advances in Deep-learning-based Prediction and Control of Plasma Instabilities and Disruptions in Tokamaks

2021 IAEA-PPPL Workshop

G. Dong 1, W. Tang 1, X. Wei 2, Z. Lin 2, K. Felker 3
1 Princeton Plasma Physics Laboratory, Princeton, NJ, U.S.A.
2 University of California, Irvine, Irvine, CA, U.S.A.
3 Argonne National Laboratory, Lemont, IL, U.S.A.

PPPL PRINCETON PLASMA PHYSICS LABORATORY

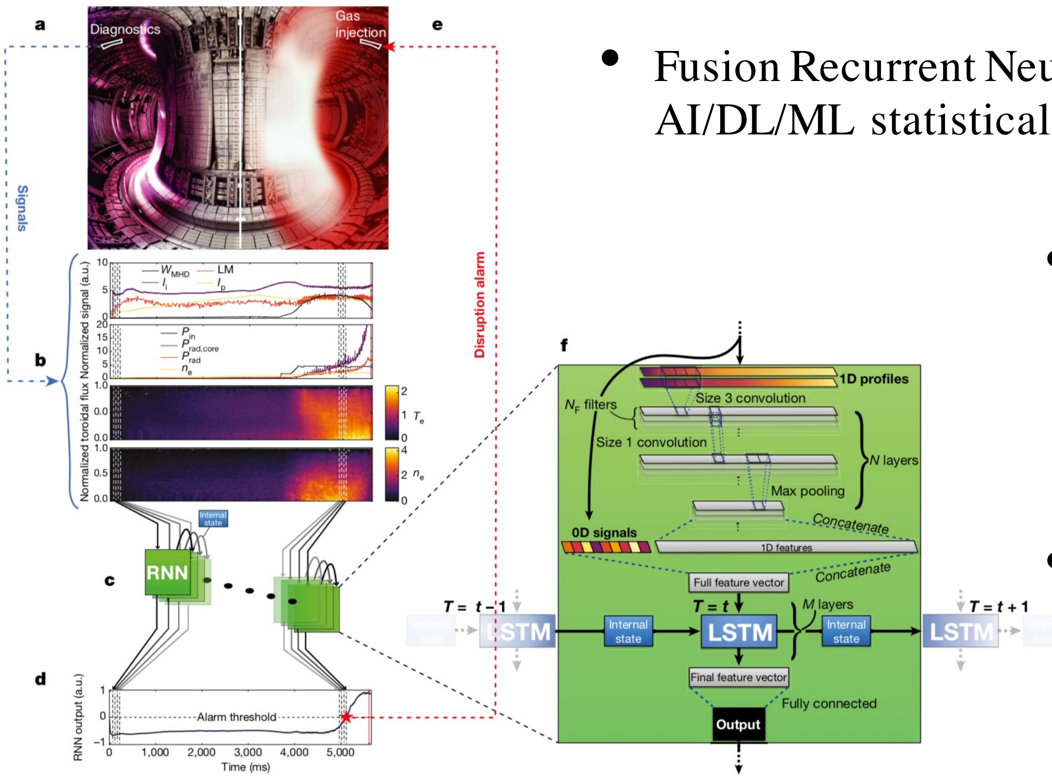
UCI

Argonne NATIONAL LABORATORY

Outline

- **Advances in deep-learning-based disruption predictor**
- **Disruption relevant instability forecasts**
- **Design of real-time control schemes**
- **Physics-informed model for real-time plasma instability analysis**

Introduction: FRNN disruption predictor



- Fusion Recurrent Neural Network (FRNN) is software based on AI/DL/ML statistical methods designed for disruption predictions.

- It is not just a RNN-based model, but synthetic software with **multiple built-in architectures (Deep: LSTM, TCN, TTLSTM, Classical: /ML random forest, xgboost,…)**

- **Less feature engineering ->** Reliable performance for **cross-machine predictions**

[Kates-Harbeck, J., Svyatkovskiy, A., and Tang, W., Predicting disruptive instabilities in controlled fusion plasmas through deep learning, Nature, vol 568, pp. 526–31, 2019]

Cross machine predictive capability

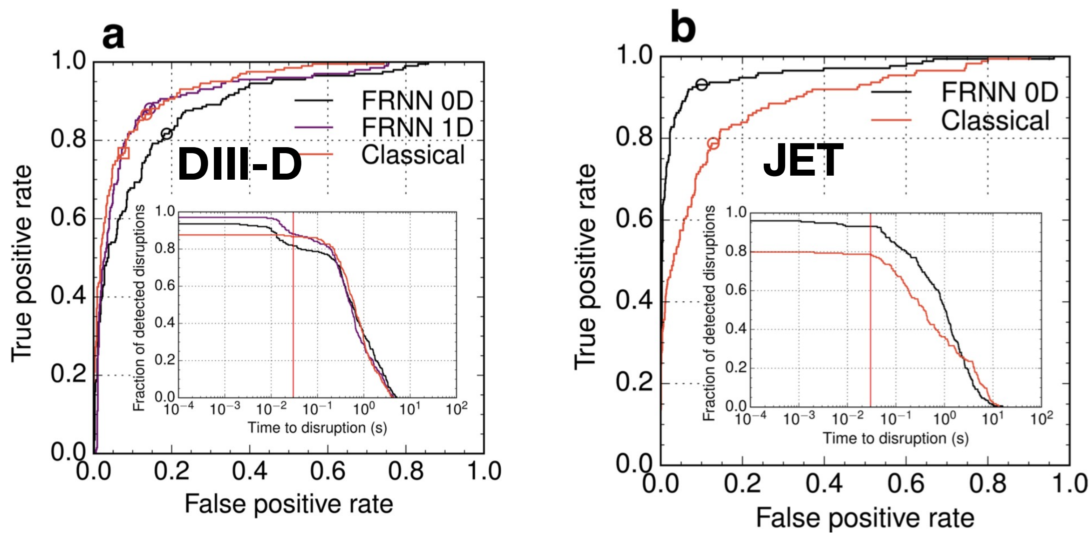
Table 1 | Prediction results

	Single machine		Cross-machine		Cross-machine with 'glimpse'
Training set	DIII-D	JET (CW)	JET (CW)	DIII-D	DIII-D + δ
Testing set	DIII-D	JET (ILW)	DIII-D	JET (ILW)	JET (ILW) − δ
Best classical model	**0.937**	0.893	0.636	0.616	0.851
FRNN 0D	0.890	**0.952**	**0.761**	0.817	0.879
FRNN 1D	0.922	–	–	**0.836**	**0.911**

Performance of the best models on test datasets, measured as AUCs at 30 ms before a disruption. We compare FRNN with (1D) and without (0D) profile information and the best classical approach. The best model for each data set is shown in bold. The last column shows results for cross-machine testing with a small amount (a 'glimpse') of data, δ, from the testing machine added to the training set (see text). A score of 1.0 represents perfect performance and 0.5 is equivalent to random guessing. Because the relevant diagnostic for 1D profiles was not available on most JET shots from the carbon wall dataset, 1D profiles are not included when training on JET data. CW, carbon wall.

Different wall condition

Cross Machine



Distinguishing disruptive and non-disruptive tearing modes

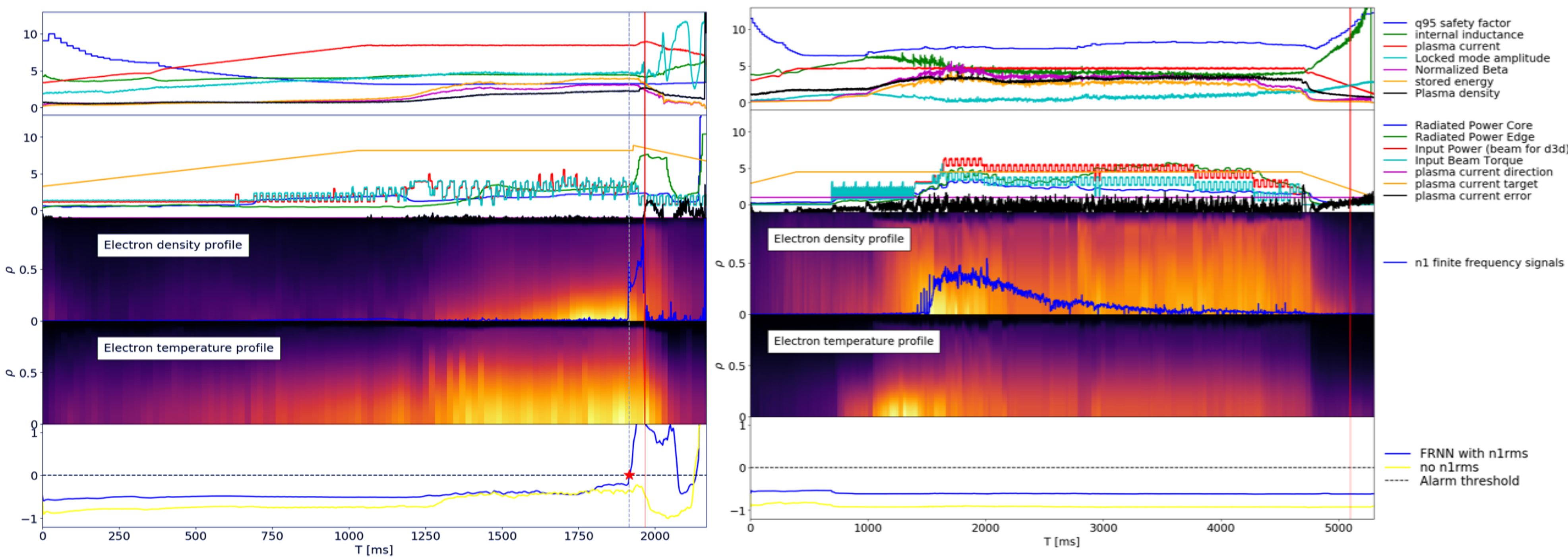
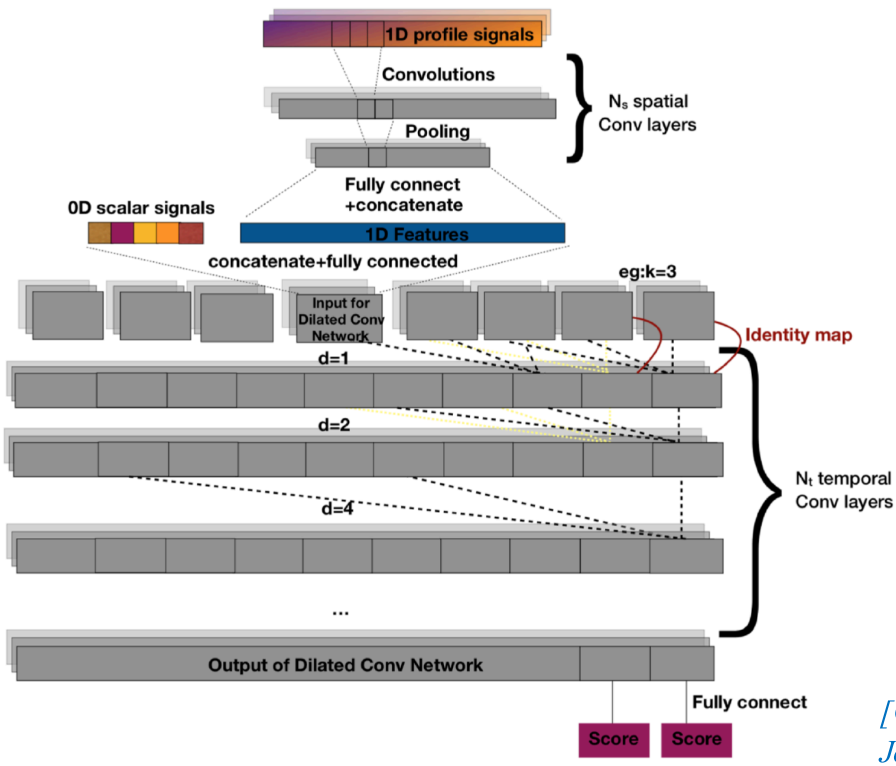
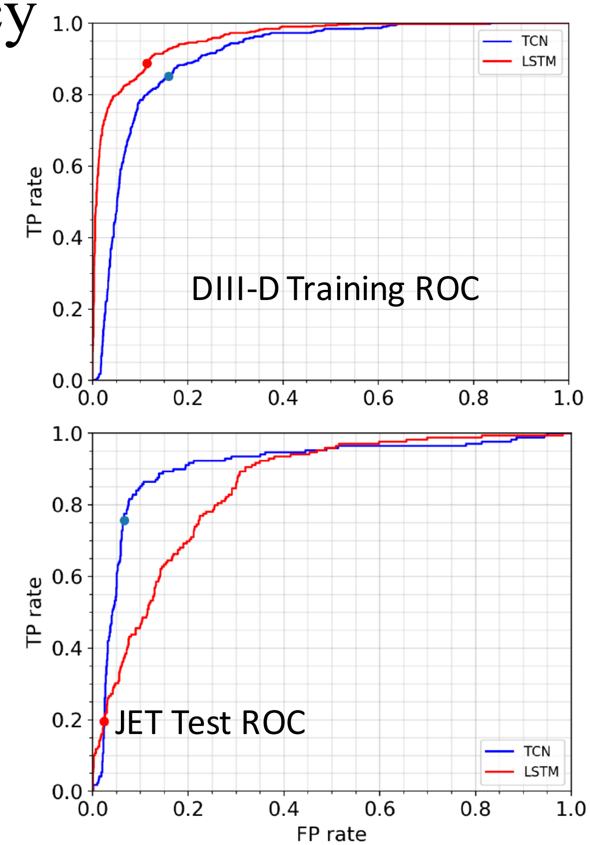


Figure: DIII-D shot number 161362 in the left panel and DIII- shot number 170239 in the right panel. Solid red line: minimum warning time: 200 ms

New architectures improve prediction accuracy

- Temporal Convolutional Neural Network (TCN)

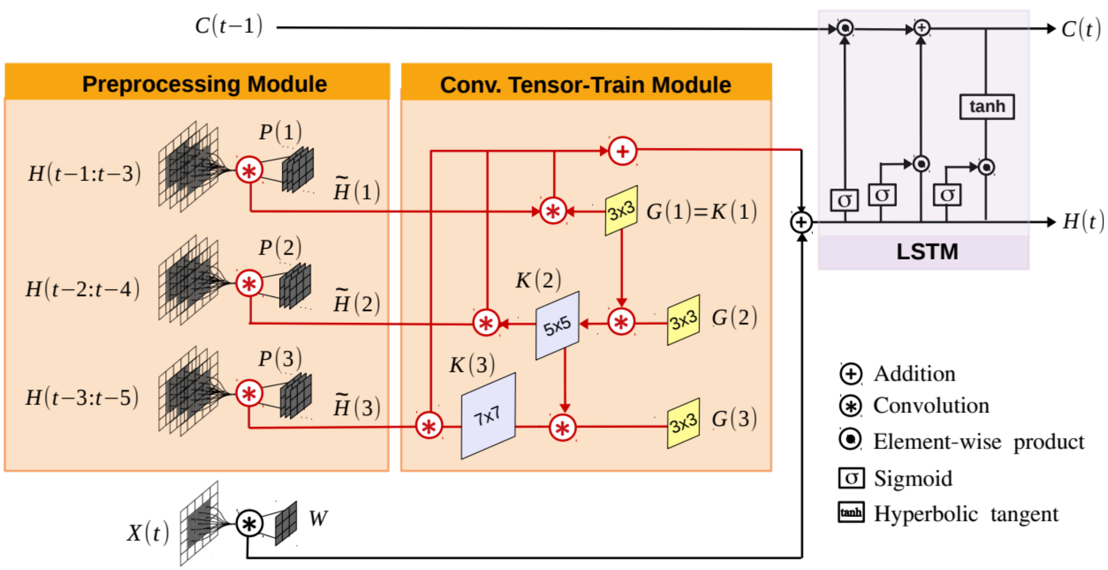




[G Dong, KG Felker, A Svyatkovskiy, W Tang, J Kates-Harbeck Journal of Machine Learning for Modeling and Computing 2 1 (2021)]

New architectures can improve early warning capability

- Convolutional Tensor-Train LSTM (TTLSTM)



Groups previous hidden states

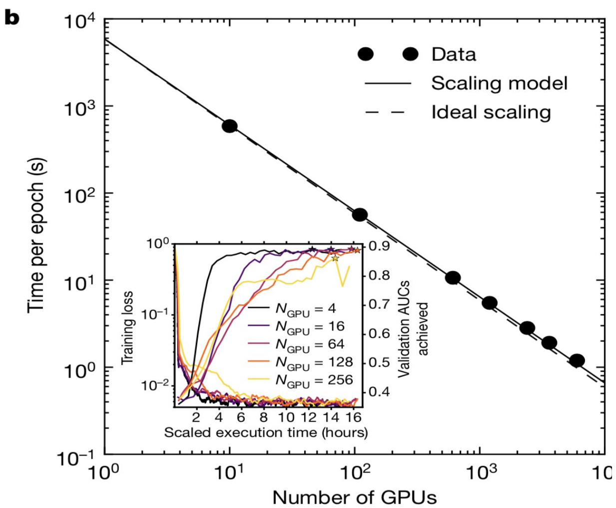
Aggregates Spatio-temporal info

Hyperparameters

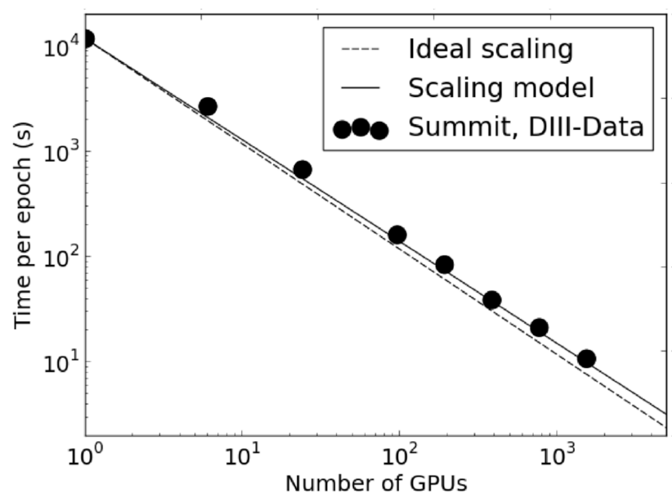
Hyperparameter	Explanation	Tuning range
Lr	Learning rate	1e-3 – 1e-5
Lr_{decay}	Learning rate decay factor	1.1—10
Lr_decay_patience	# no improve epoch before learning rate decay	3-10
N_{batch}	Training batch size	8 –128
$T_{warning}$	Warning time for target function, which becomes positive at $T_{warning}$	5-50
Target	Type of target function	ttdinv, ttd, hinge
N	**Order of the LSTM**	**1-4**
K	**Rank of the decomposition**	**1-4**
M	**Cell steps**	**1-4 (>=N)**
n_{cell}	**TT-LSTM hidden channels**	**8-200**
Lambda	Weighing factor for positive examples	1-16
Extra_Dense	Extra dense layer before output layer	True/False
Dropout	Dropout probability	0.1—1

Effective utilization of HPC resources

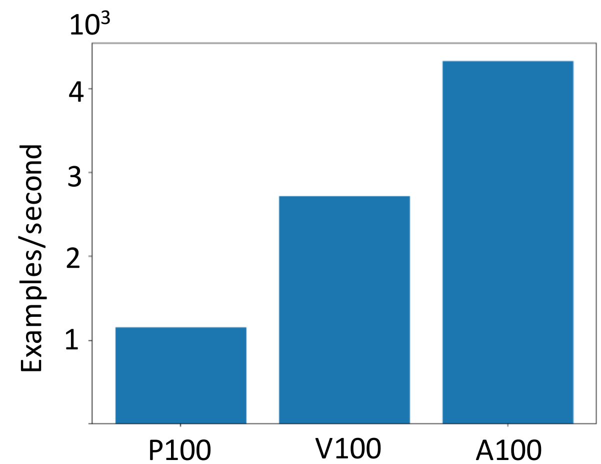
- Engaging **supercomputing resources** effectively **->** fast automatic hyperparameter tuning



LSTM architecture on
Titan (OLCF)

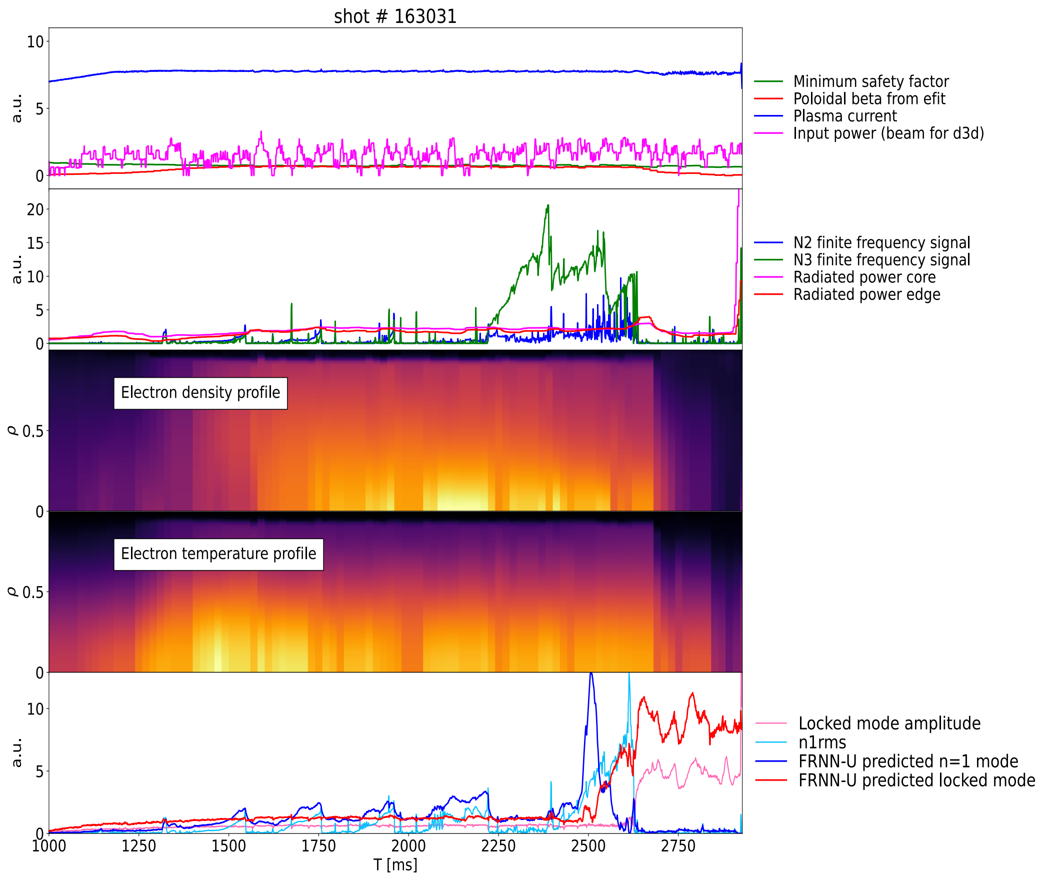


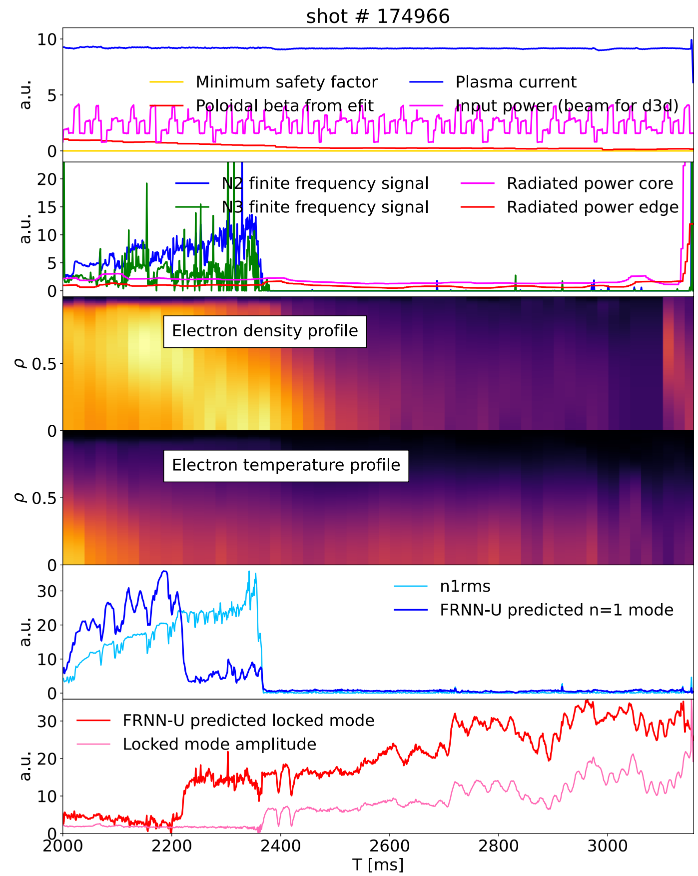
TCN architecture on
Summit (OLCF)



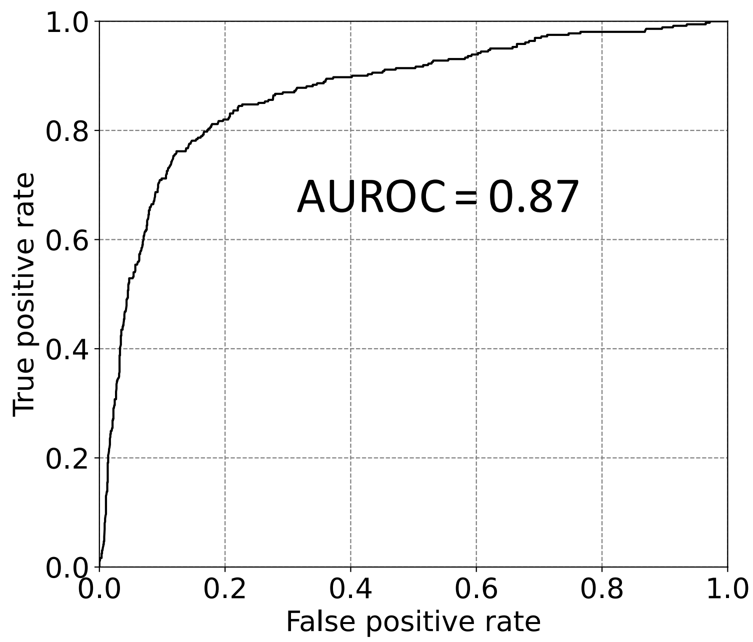
Comparison of performance across 3 generations of NVDIA hardware
(LSTM, batch_size = 256)

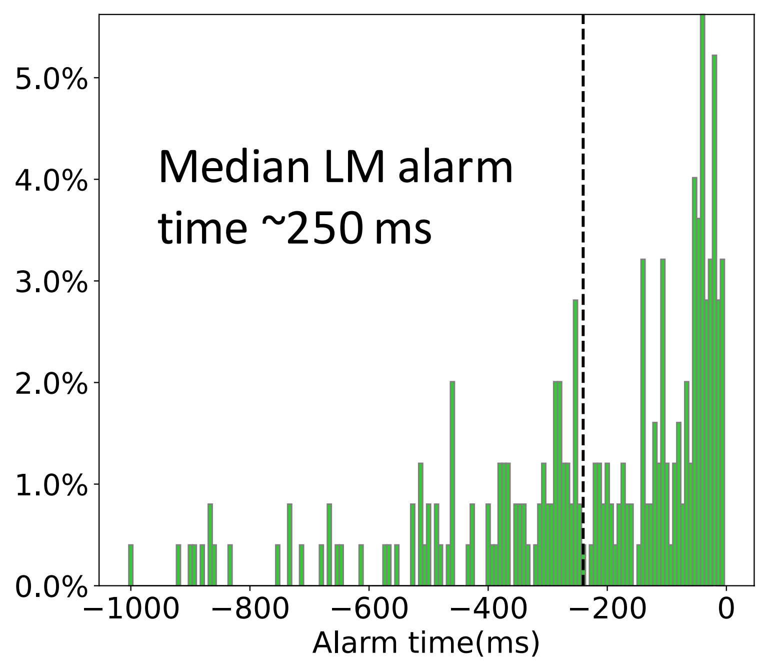
Forecast of n=1 mode and locked mode





Capability of forecasting mode locking



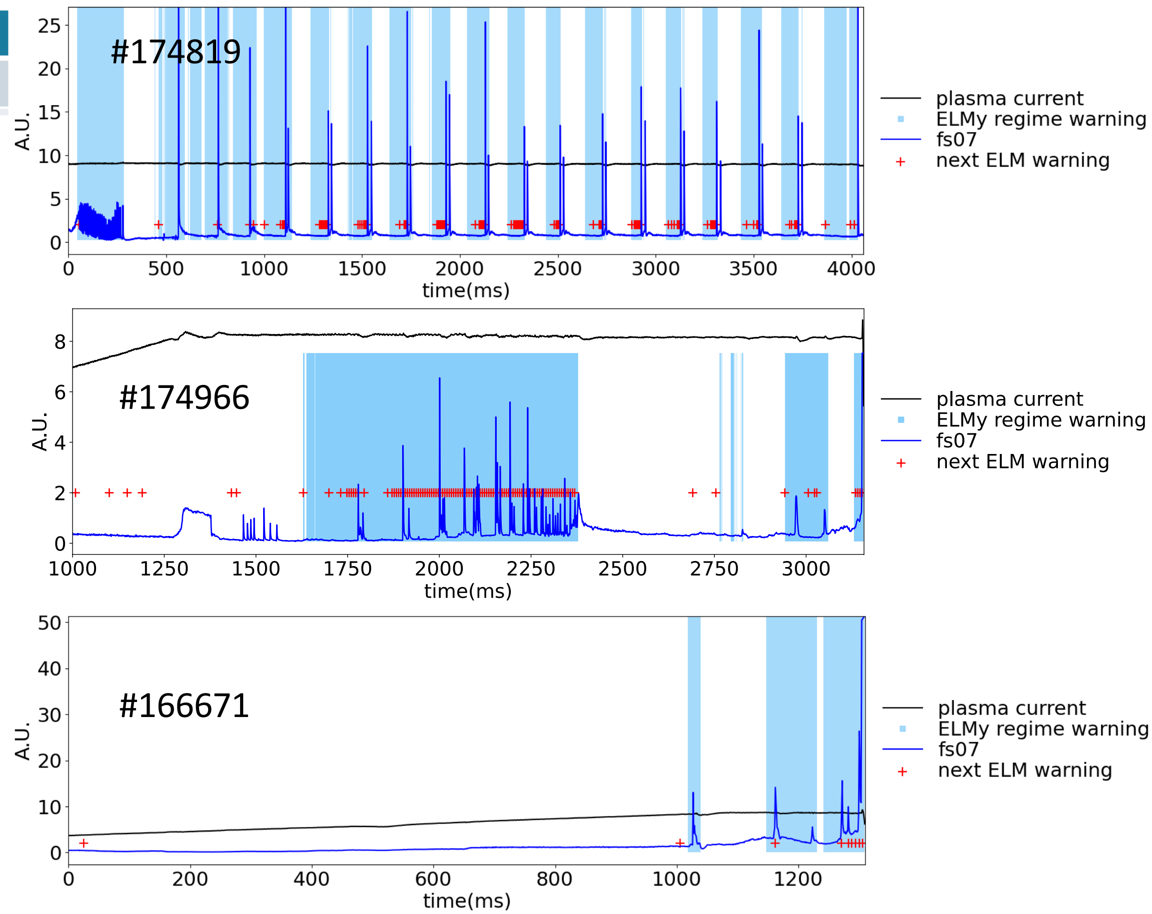


ELM forecasting capabilities

DIII-D shots	#Valid shots	# ELMs
#143861-#184800	8505	~300,000

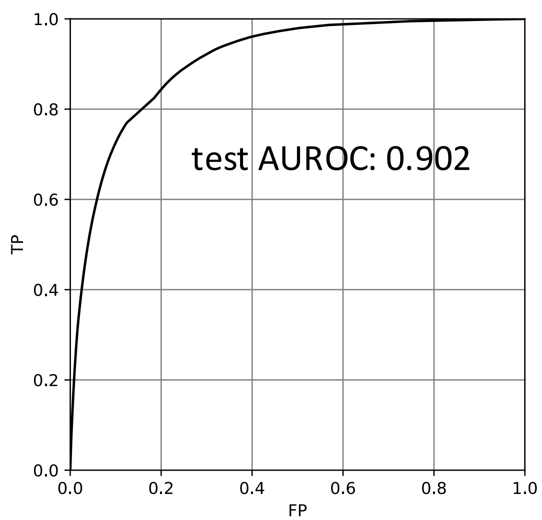
Additional database:

- **Pedestal parameters:** peped, newid, neped, teped, tewid (Important)
- **Spectrogram** (No effect on general performance)
- **Rational surface positions** (No effect on general performance, but affects specific ELM predictions)

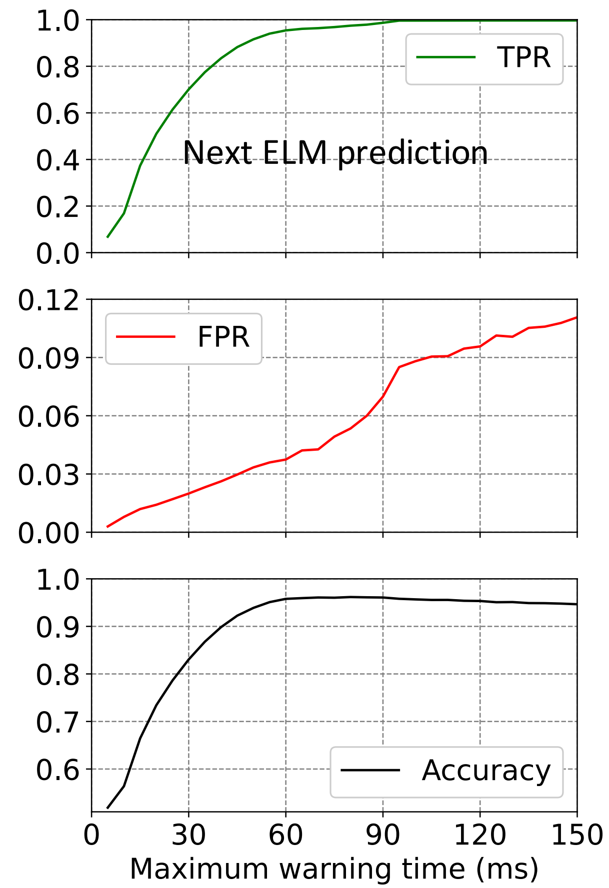


Strong ELM predictive capability

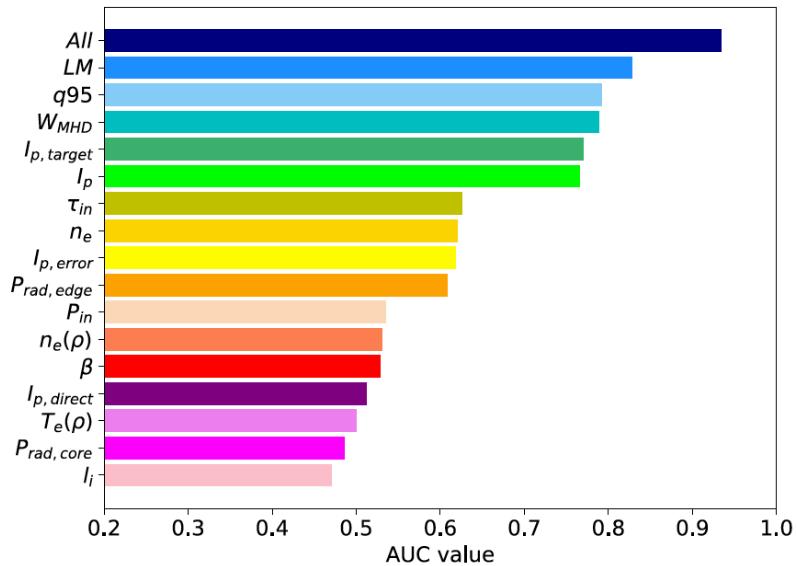
- ELM regime prediction: a time step is considered ELMy regime if it is 1. during ELM, 2. going to ELM in the next 100ms
- If a time step is ELMy and predicted to be ELMy, it is considered TP



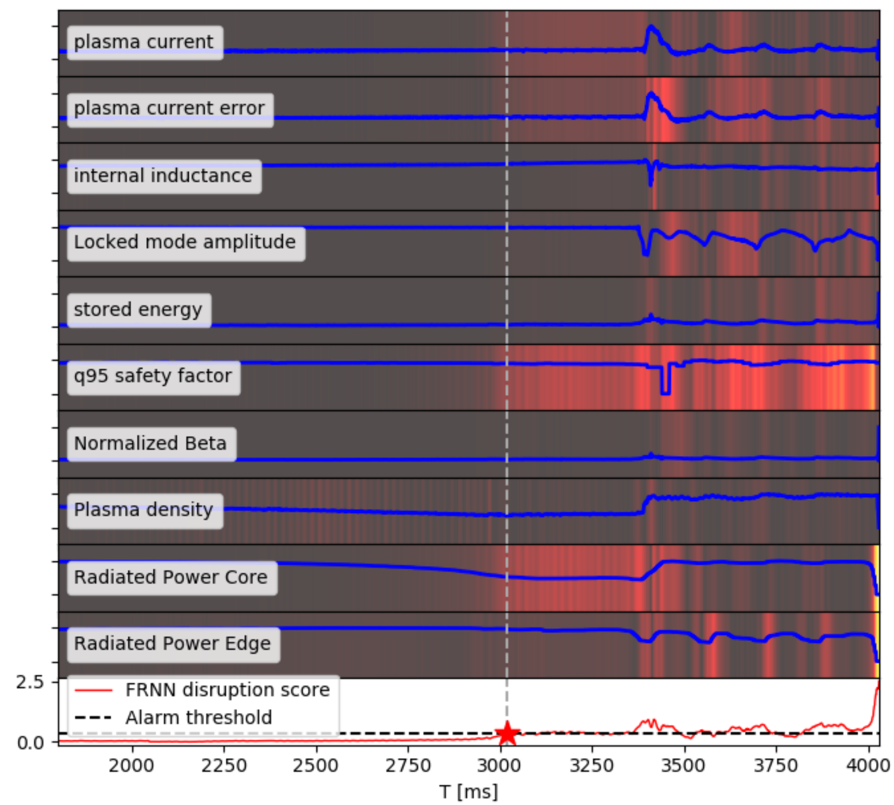
- Next ELM prediction: for an ELM event if a warning is present within 'Maximum warning time' **before** it happens, it is a TP.
- For randomly selected non-ELMy time chunk of length 'maximum warning time', if warning exist within this chunk or within 'Maximum warning time' before it, it is a FP.
- **Outputs can be connected to the PCS proximity controller for targeted instability control.**



Contribution of signals to disruption score in individual shot

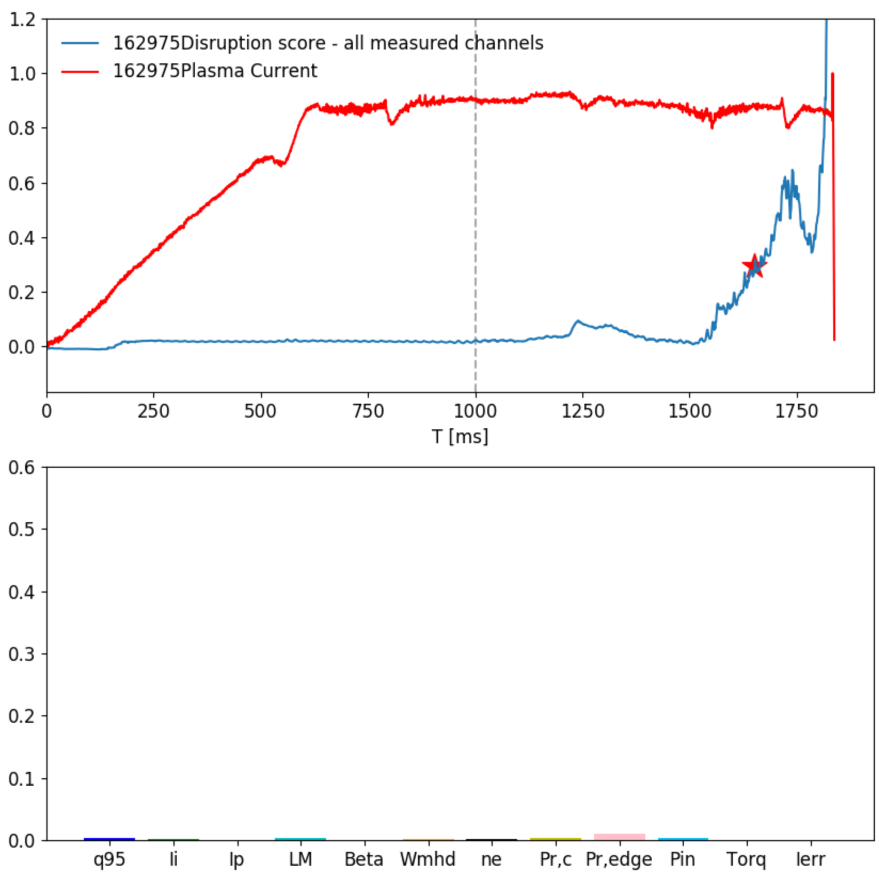


Signal importance study: Test AUROC for models trained on single signal



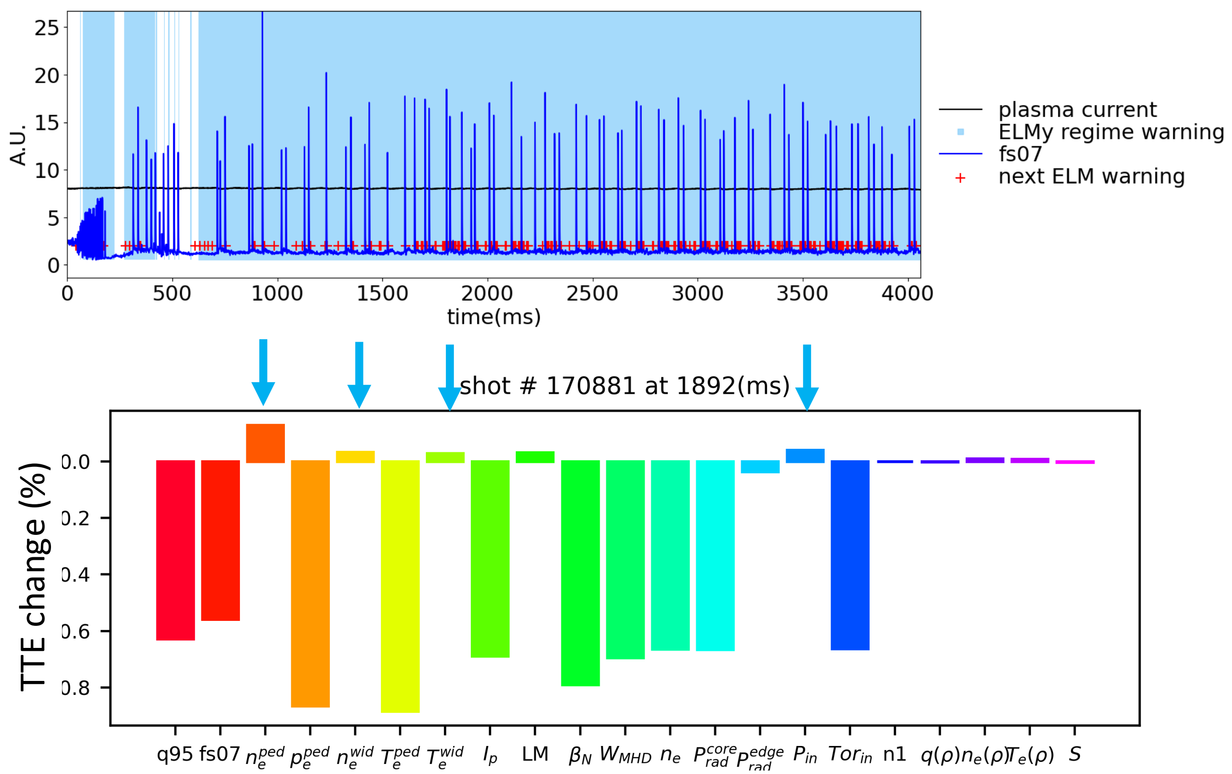
Sensitivity score heat-map of the shot #164582

Evolution of channel sensitivity



DIII-D shot #162975. The upper panel shows the evolution of the plasma current (red line) and the FRNN output of the disruption score (blue line -- with the lower panel showing the sensitivity scores (for associated signals such as q95, etc.) at the time of the disruption alarm (red star in the upper panel).

Active control scheme

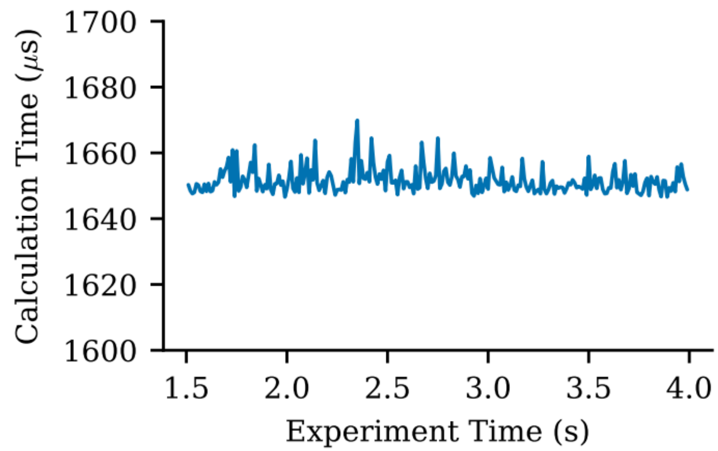


Change in time to ELM when signal amplitude is **decreased**

- Active control scheme can be developed based on real-time sensitivity scheme results
- Problems: 1. Is the signal heatmap 'subjective' to model? 2. Is the correlation causal?

Real-time implementation

- FRNN has been implemented in DIII-D PCS with Keras2C API to translate inference model into C language.
- Recent NVDIA product **TensorRT** can automate the process. Enabling easy and fast model maintenance and update.
- Figure: FRNN disruption prediction computing time in DIII-D PCS



[R Conlin et al. Keras2c: A library for converting Keras neural networks to real-time compatible C Engineering Applications of Artificial Intelligence 100, 104182 (2021)]

Motivation for physics-informed instability simulators

- First-principles based global plasma instability simulation such as the gyrokinetic toroidal code (GTC) provides physics-based insights, for example, kink-like modes tearing modes, energetic particle (EP) modes, and detailed mode structure, to guide targeted real-time plasma control.
- GTC can serve as an **accurate plasma simulator** for important instabilities such as various MHD modes, which can eventually lead to disruptions.
- To make GTC available in real-time, we need a surrogate model to reduce simulation time (~hours) to model inference time (~ms).

Fishbones & kinks-tearing- LM->disruption

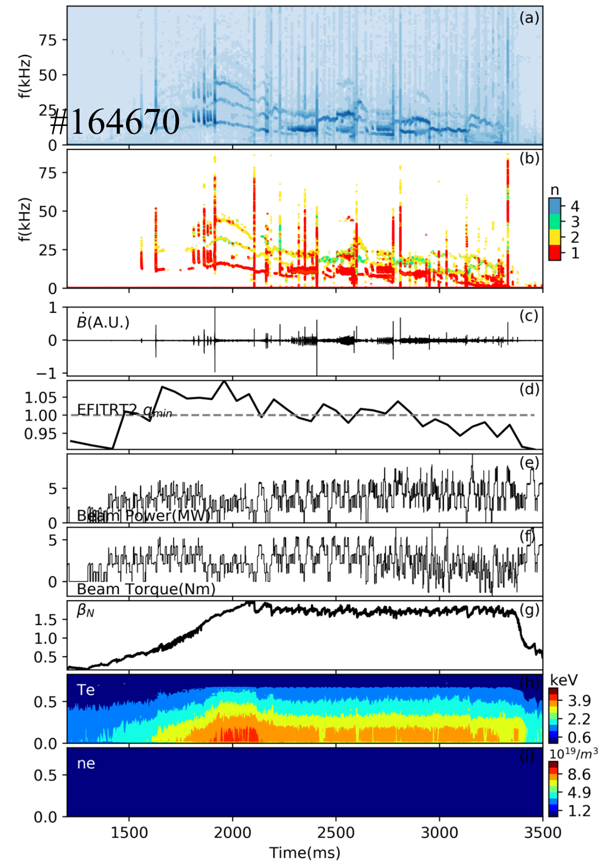
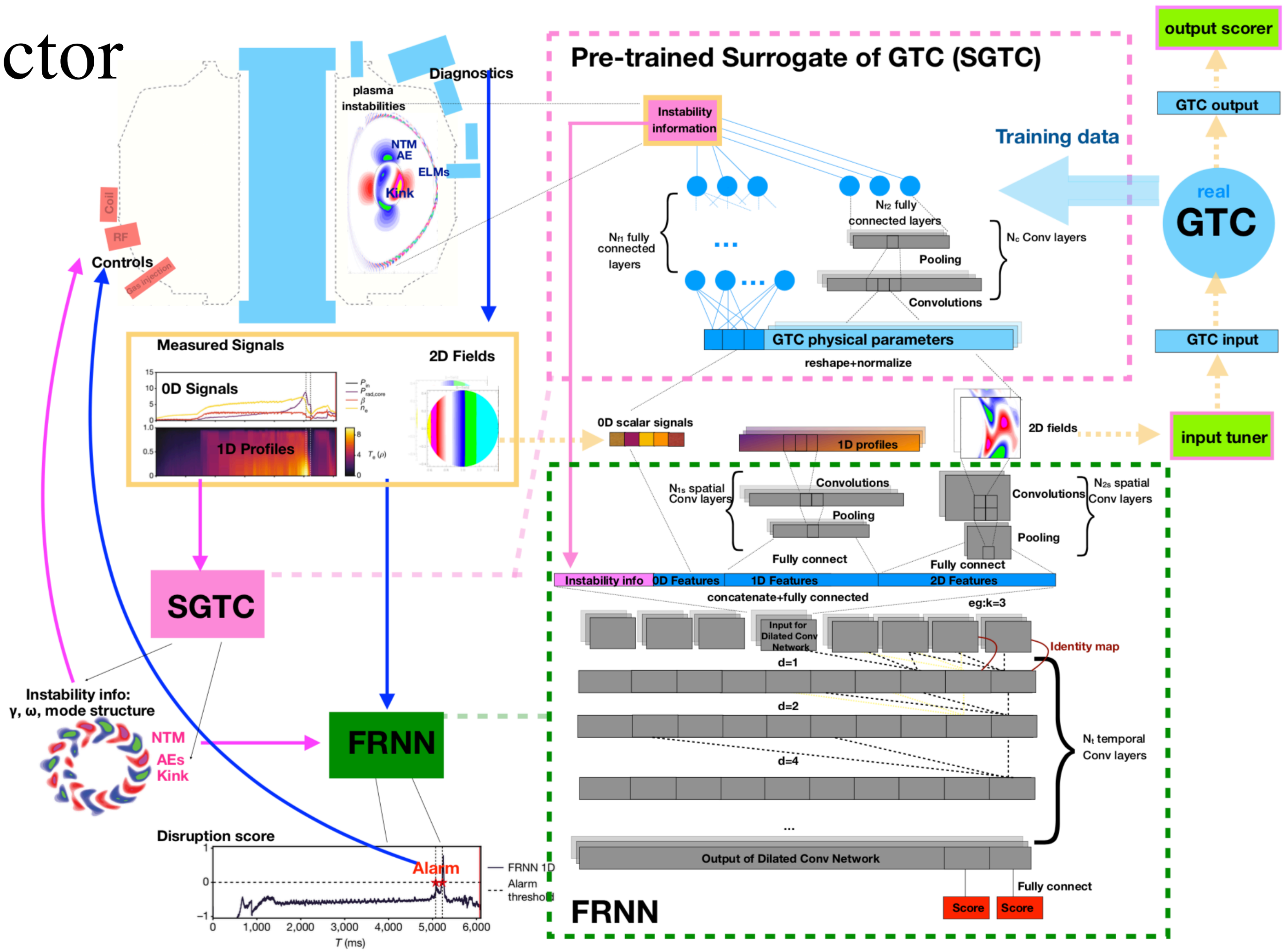


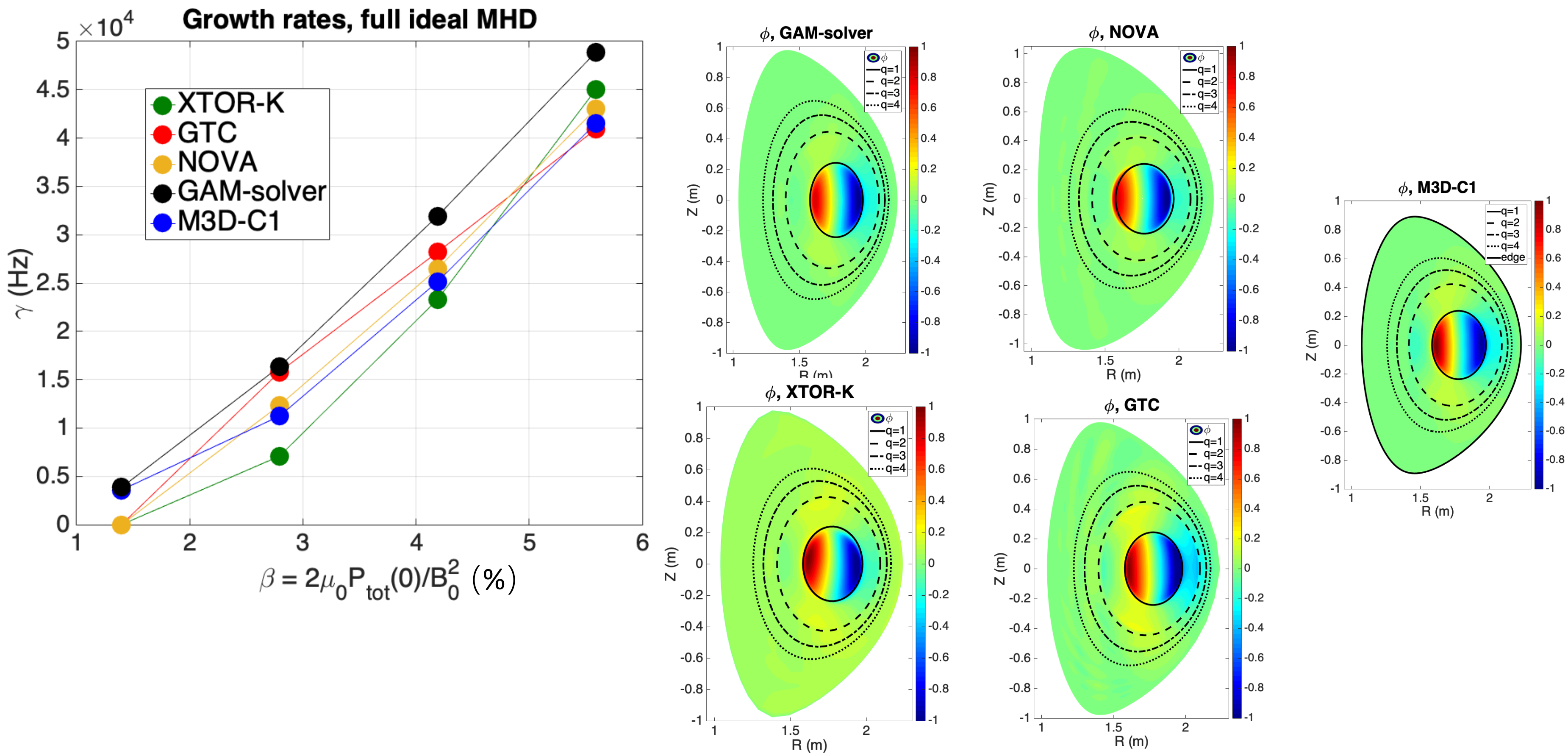
Figure: Example DIII-D shot #164670
MHD instabilities finally lead to mode locking and disruption. Multiple modes can be coupling with complex dynamics.

Instability predictor workflow

- SGTC output
  → PCS for real-time MHD instability control
- SGTC outputs
  → other predictive algorithms such as FRNN models as an input, or simulator for disruption prediction and control.



Benchmark of linear internal kink simulations



DIII-D equilibrium Data preparation

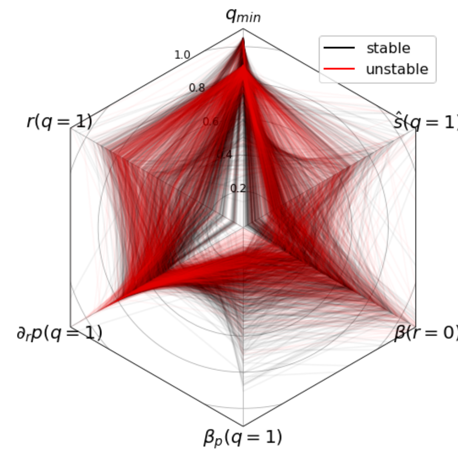
- Time sliced data are selected randomly from shot # 139520 to shot # 180844
- 5758 DIII-D equilibria simulated
- 2872 are based on EFIT01 output, and 2886 are based on EFIT02 output.
- These simulations have been carried out in 12 GTC runs, each simulating 500 DIII-D experiments in parallel using 2000 nodes of the Summit supercomputer at ORNL for about 30 minutes (Summit has about 4700 nodes).
- **'input tuner'** :ORBIT code [White 1984] to convert equilibrium data to Boozer coordinates

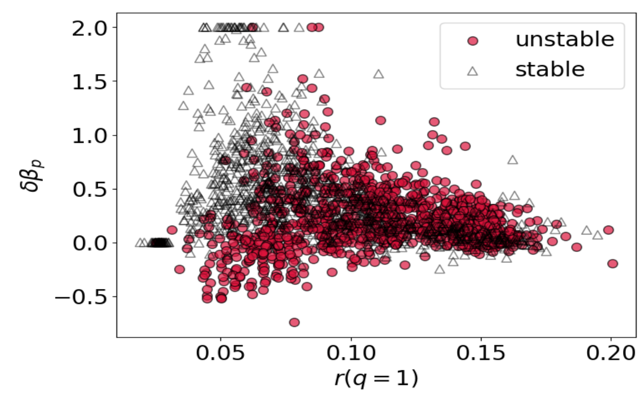
shot number	data
5758 equilibriums from shot # 139520-180844	EFIT01 gfile
	EFIT02 gfile
	ZIPFIT01 electron temperature profile
	ZIPFIT01 electron density profile
	ZIPFIT01 ion temperature profile
	Magnetic perturbation (mpi66M307D)
	Magnetic perturbation (mpi66M322D)

GTC kink simulations for 5000+ DIII-D equilibirums

- Global electromagnetic simulations in the MHD limit are run for each input for 3000 time steps. time step size $0.01\frac{R_0}{C_s}$.
- Filter: n=1.
- Typical physical duration :0.1 ms.
- **‘output analyzer’** : examine the output data, exclude numerical instabilities; prepare proper target data such as the mode growth rate, and poloidal eigenmode structure for SGTC.
- n=1 mode growth rate is calculated with a linear fit of the perturbed parallel vector potential from the last 1000 time steps of the simulations.
- 1972 equilibria have unstable n=1 kink modes.

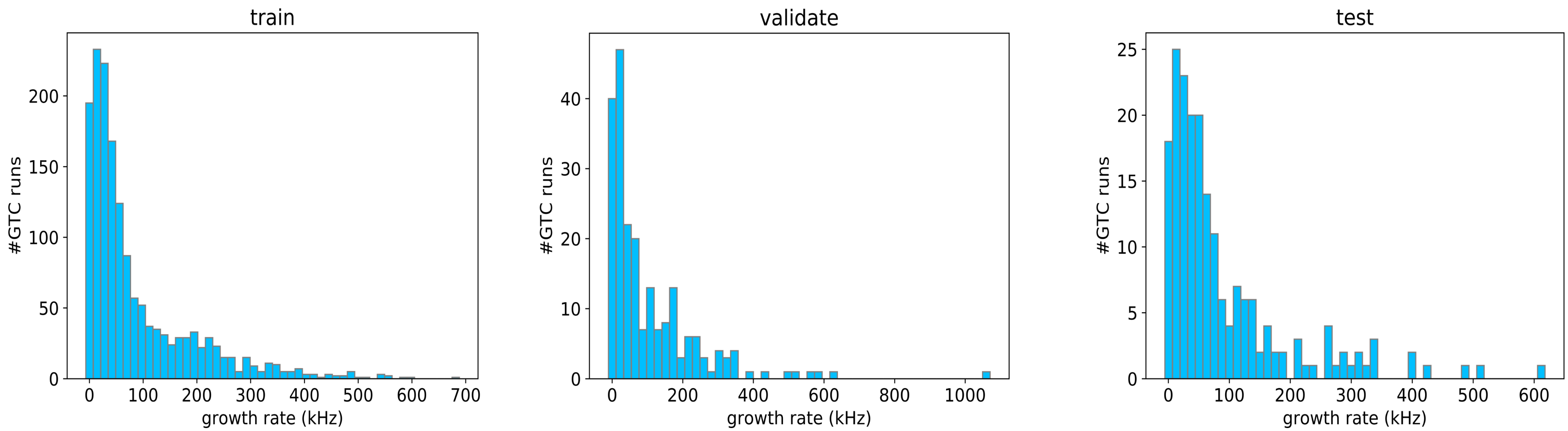
$$\delta\beta_p = -\frac{{R_0}^2 \int^{r1} p' r^2\, dr}{{B_0}^2 {r_1}^4}$$





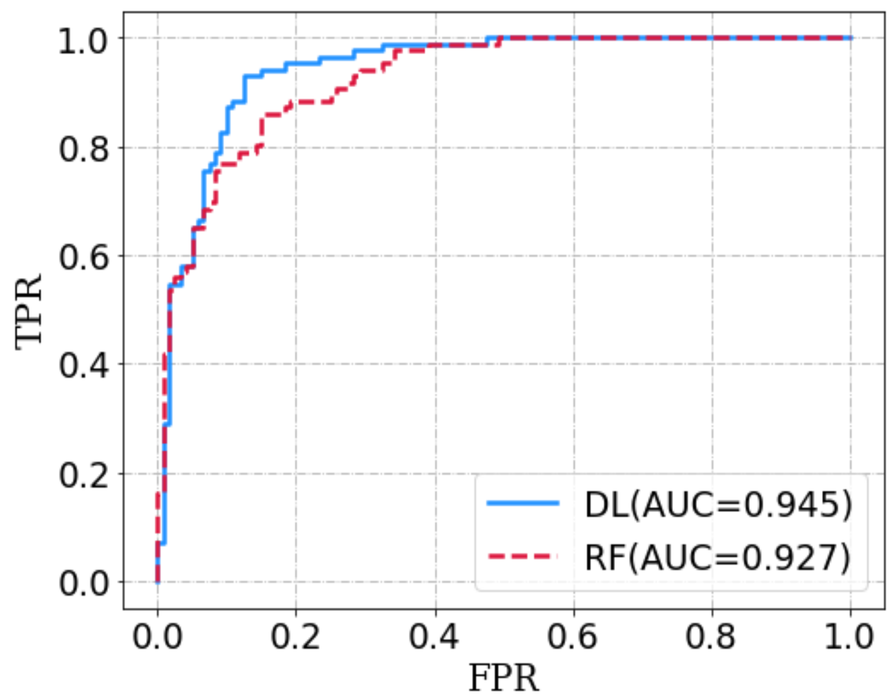
Growth rate distribution

GTC simulation results of all the 1972 unstable cases and 2531 stable cases are used for training (80% of the data), validation (10% of the data), and testing (10% of the data) of SGTC.

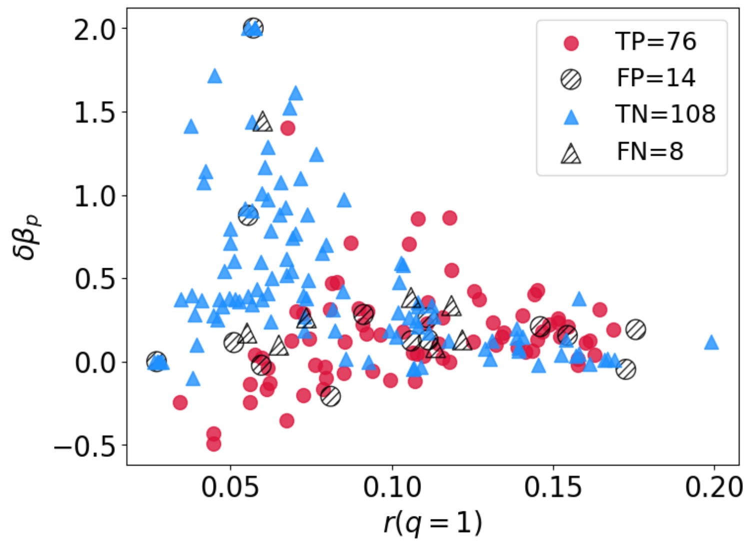


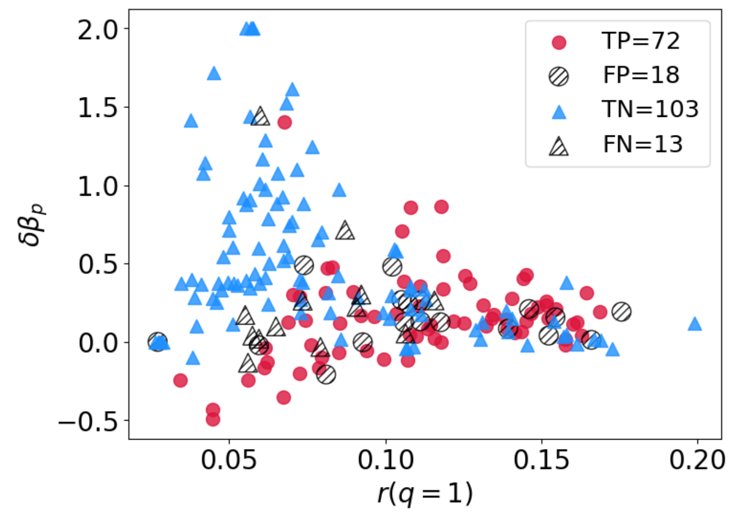
SGTC trained to predict→ 1. Instability; 2. Growth rate; 3. Mode structure

Training results: 1. predict for **instability**



- **classical algorithms vs deep neural network**
  - Trained multiple classical models in the sklearn package [sklearn]→ the best performing one is the random forest model.
  - Hand-tuned both the random forest model and the neural network for the instability prediction on the validation dataset, and reported the ROC curve on the test dataset. T
  - he AUC for the deep learning based model (DL) and the random forest (RF) model on the test dataset are 0.945 and 0.927 respectively

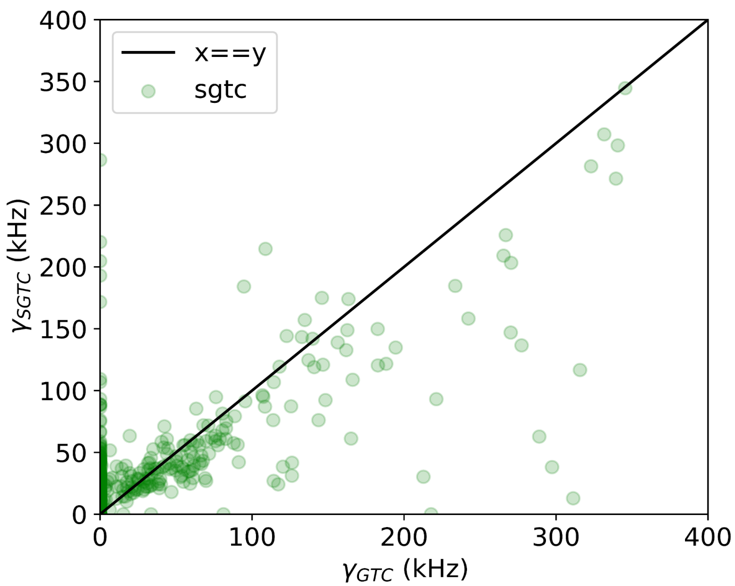


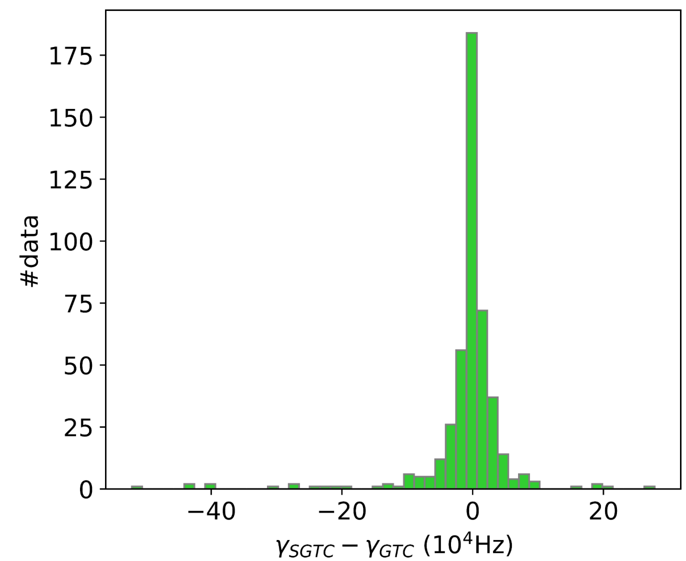


Prediction of kink instability on test dataset from deep learning method in the left panel and random forest method in the right panel. Solid red dot represents the true positive (TP), solid blue triangle represents the true negative (TN), shaded circles represent the false positive (FP), and shaded triangles represent the false negative (FN).

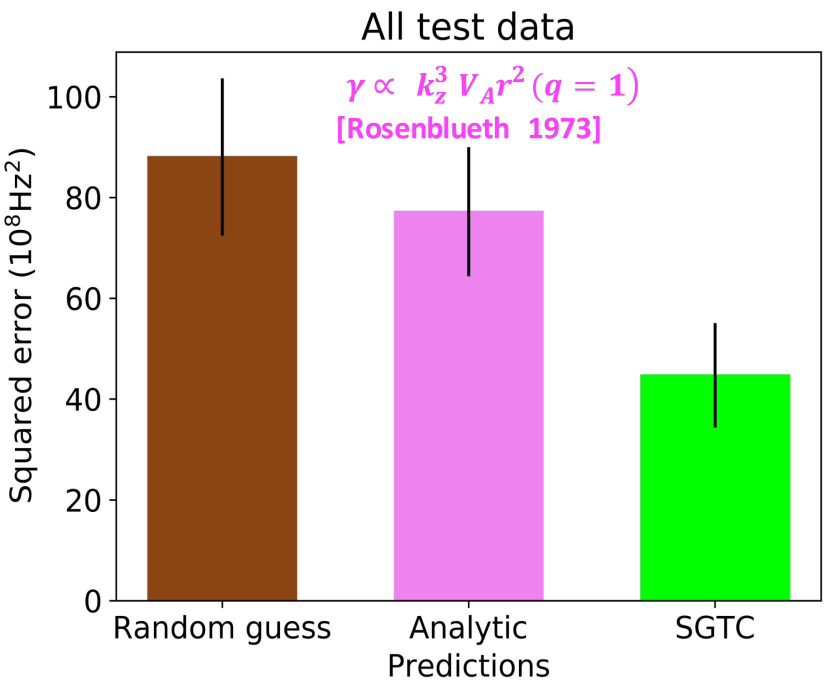
Training results: 2. predict for **growth rate**

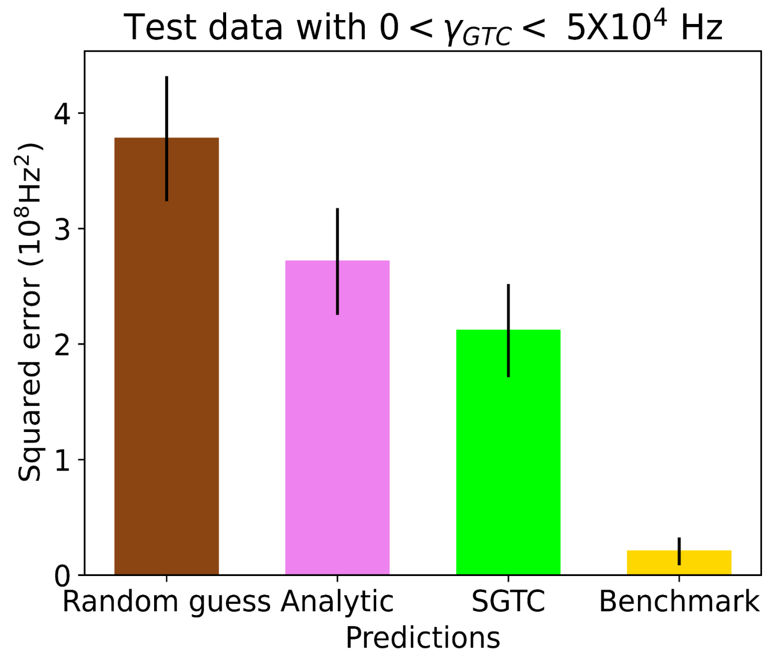
- Random hyperparameter tuning on the validation set.
- Inference time <1ms on Nvidia V100 GPUs (Princeton Traverse).





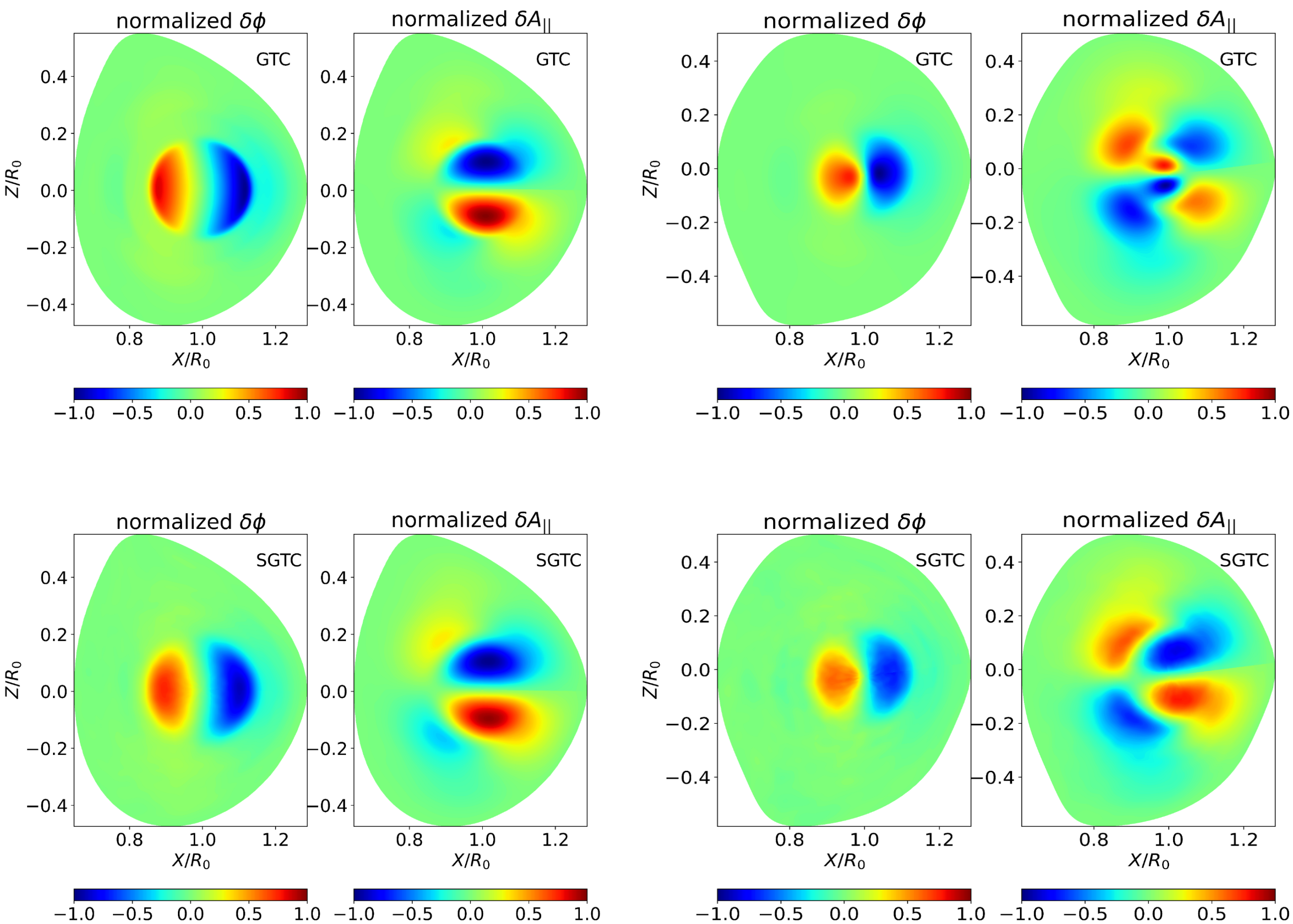
Prediction results of the kink growth rate for entire test dataset. The left panel visualizes the true value of the growth rate vs the predicted value of the growth rate. The solid black line indicates x == y where perfect predictions occur. The right panel shows a histogram of prediction error.





Comparison of SGTC performance for all test dataset with random guess and analytic formula in the left panel. Right panel shows the comparison of SGTC performance with random guess, analytic formula for test data with kink growth rate smaller than 50 kHz. Yellow bar represents the difference between GTC and four other MHD-based simulations [Brochard 2021] for DIII-D shot #141216 at 1750 ms

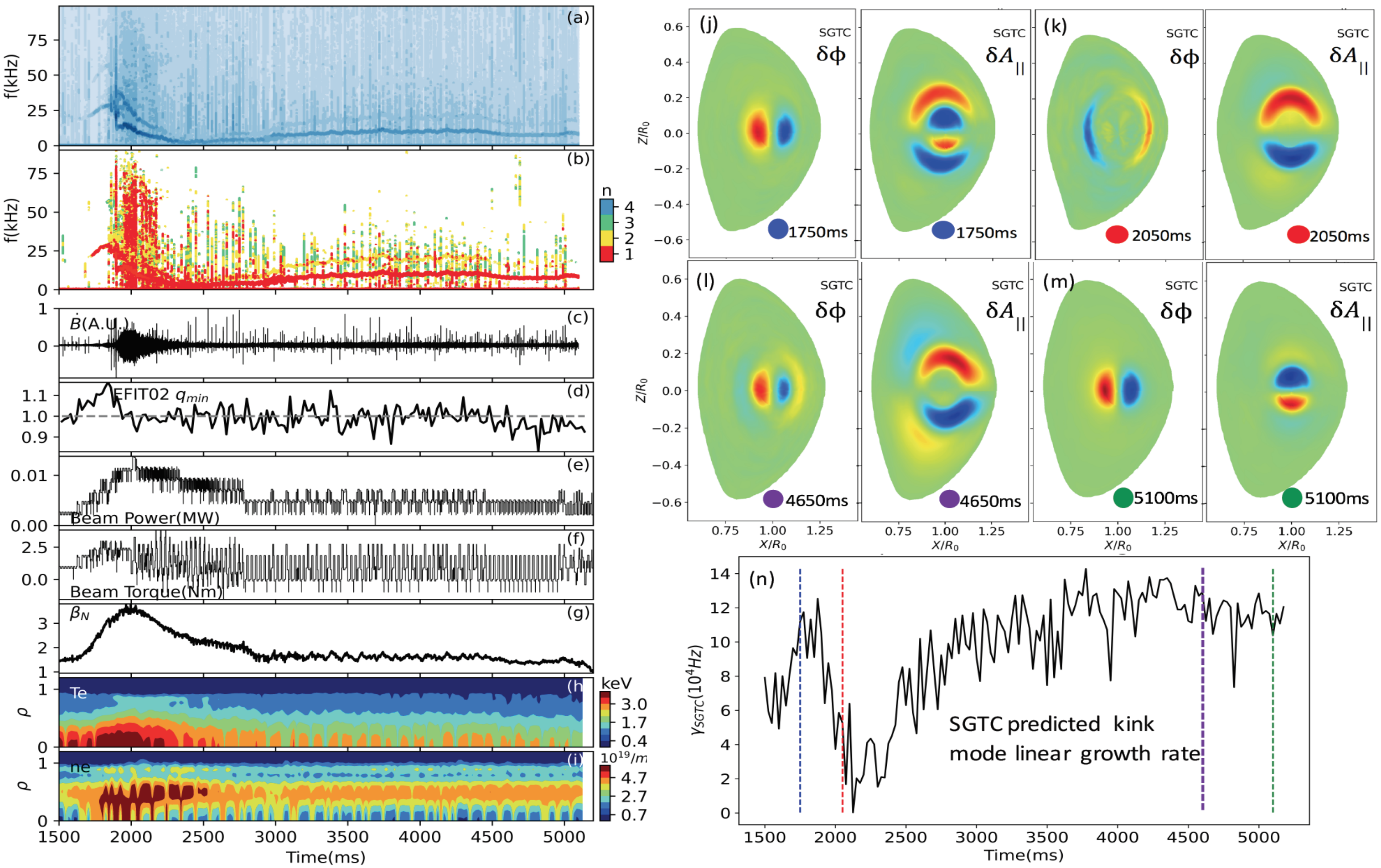
Training results: 3. predict for **mode structure**



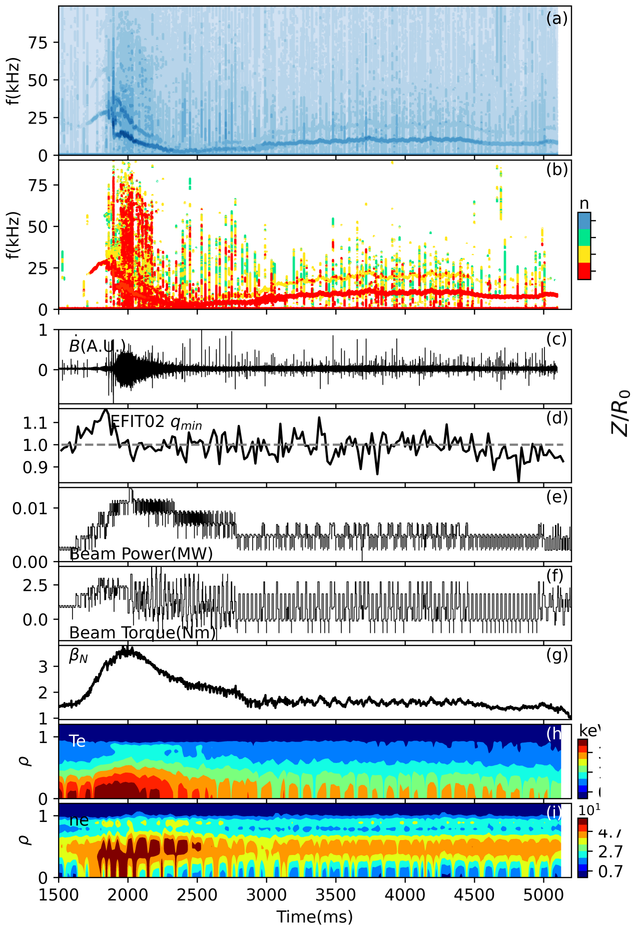
- Random hyperparameter tuning on the validation set.
- Around 85% test data output has qualitatively correct mode structure

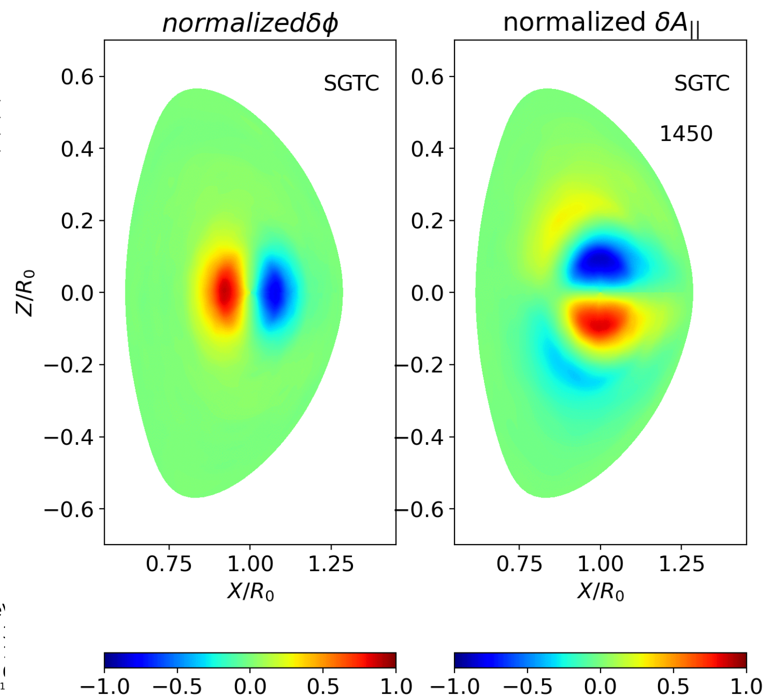
Figure: shot #162930 at 1820 ms using EFIT01 reconstruction & shot #140510 at 3145ms using EFIT02 equilibrium reconstruction.

SGTC outputs evolution of the internal kink mode



141216 Evolution of 2D mode structure





- Future & Ongoing → **fishbone, tearing mode, Alfven eigenmode, and microturbulence,** nonlinear dynamics, transport etc.
- The methodology of SGTC can also be applied to training emulators for other first-principles plasma simulations such as the MHD codes.

Summary

- We developed the FRNN framework of deep learning based models for plasma instability predictions, including disruptions, n=1 mode, and ELMs.
- Tested implementation in DIII-D PCS.
- Developed physics informed instability simulators for linear internal kink mode.
- SGTC shortens the simulation time by at least **six orders of magnitude** → presents for the first time the possibility of bringing physics-based instability information from the first-principles based massively parallel simulations into the PCS of modern tokamaks.
- The methodology of our deep learning plasma instability predictions and control can be easily applied to general plasma state and transient event predictions, showing promise for AI-based PCS capabilities in future plasma devices.

- **Thanks!**