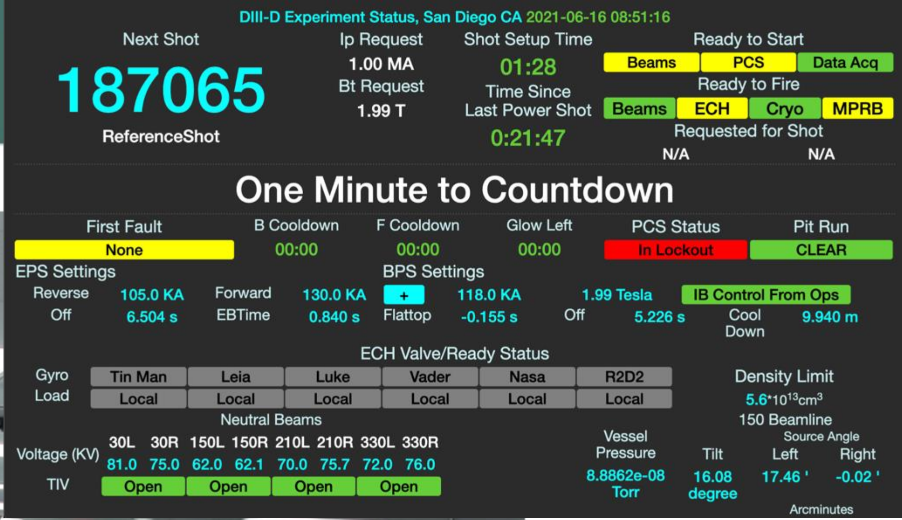
Real-time tearing mode control on DIII-D:

Leveraging ML models for high accuracy and interpreting the ML model predictions

A. Rothstein[1], H.J. Farre-Kaga[1], S.K. Kim[2], A. Jalalvand[1], J. Seo[3], R. Shousha[2], J. Josephy-Zack[2], C.S. Byun[1], J. Abbate[1], R. Conlin[1], and E. Kolemen[1,2]

[1]Princeton University, [2]Princeton Plasma Physics Lab, [3]Chung-Ang University



Plasma Control

Improving plasmas by trial-and-error

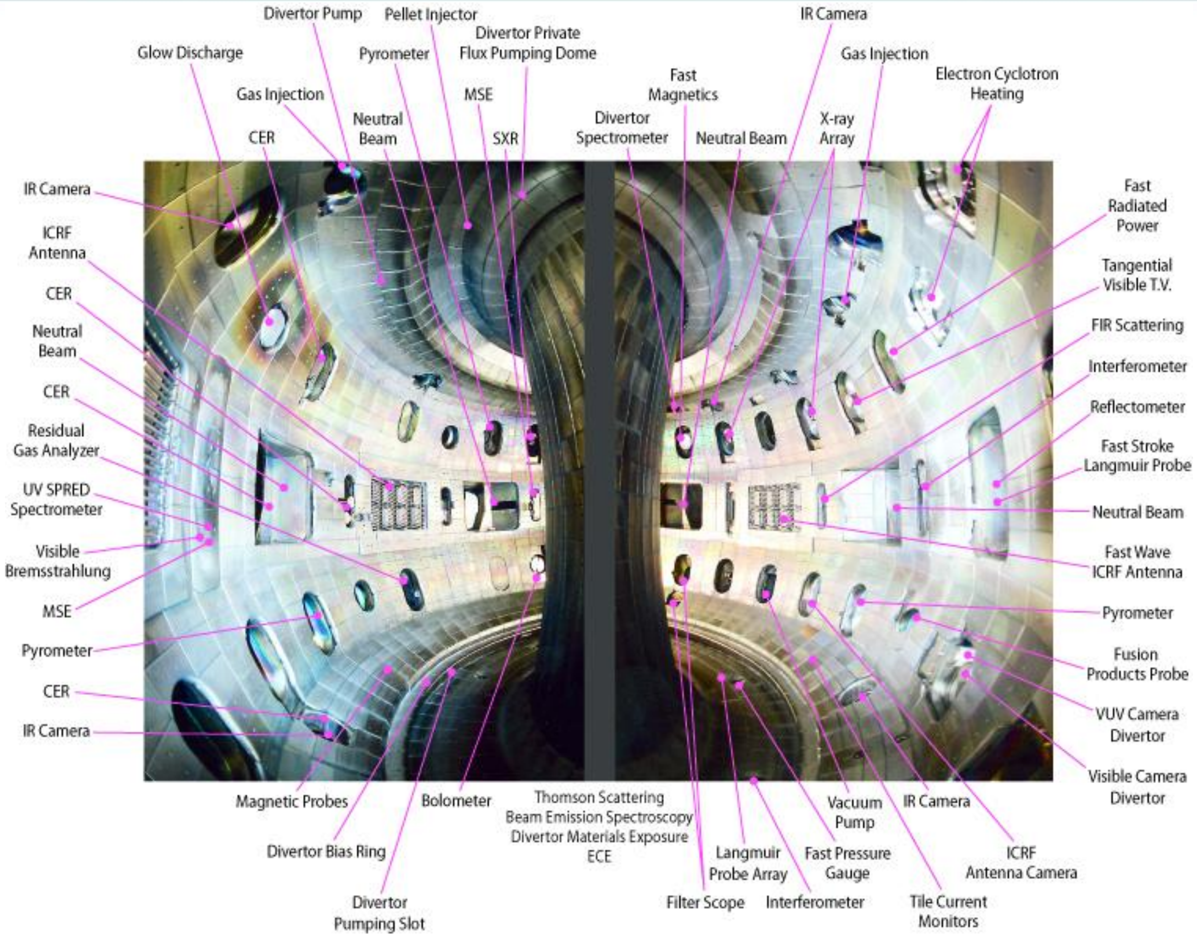
- “[after a previous full-day experiment we were] able to reproduce 133103 in 180636, 180643 and 180644.
- Many shots had MHD modes at 3 s… to try to improve that we changed **Electron Cyclotron Heating** deposition (180639-180642), and go to lower (180643-180646) and higher (180647) **plasma current**… none of which were successful.
- We also **tried lowering the voltage on the off-axis beams** (180645) to get rid of the bursty modes and **moving the BetaN ramp earlier** (180646.)”
- Ultimately, got “good reproduction of 133103, but no significant improvement”

Human operators combine simulations, heuristics, and experience to achieve desired state by trial-and-error

Outline

- **Crash course to tokamak experiments**
- Long term TM prediction and preemptive ECCD TM suppression and understanding ML predictions
- Surrogate models for real-time control

Observing the plasma state



Plasma Control

Reconstructing the plasma state

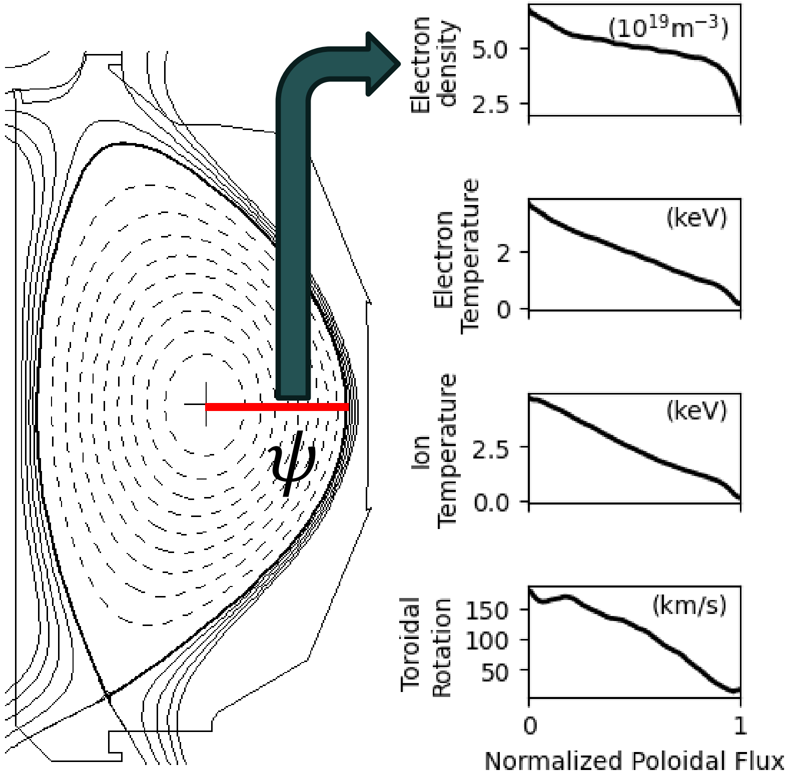
- Use diagnostic data to reconstruct plasma equilibrium

Scalar Parameters

- Plasma shape and boundary ($\kappa, \delta_{u,l}$ etc)
- Normalized pressure (β_N)
- Plasma current (I_P)
- Magnetic field (B_T)

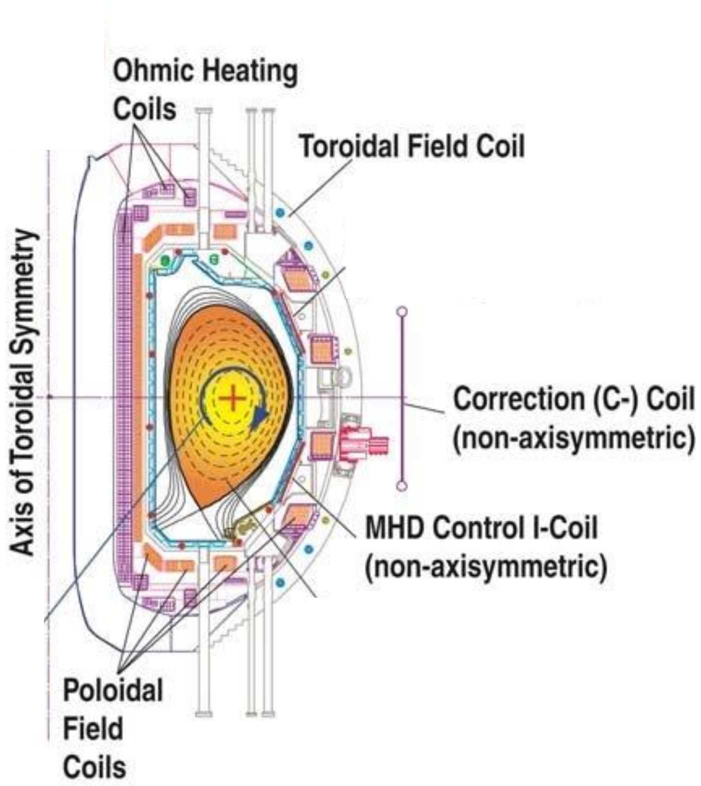
1D Profiles

- Pressure (P)
- Safety factor (q)
- Electron temperature and density (T_e, n_e)
- Ion temperature and density ($\mathrm{T_i}, \mathrm{n_i}$)
- Rotation (Ω)



Plasma Control

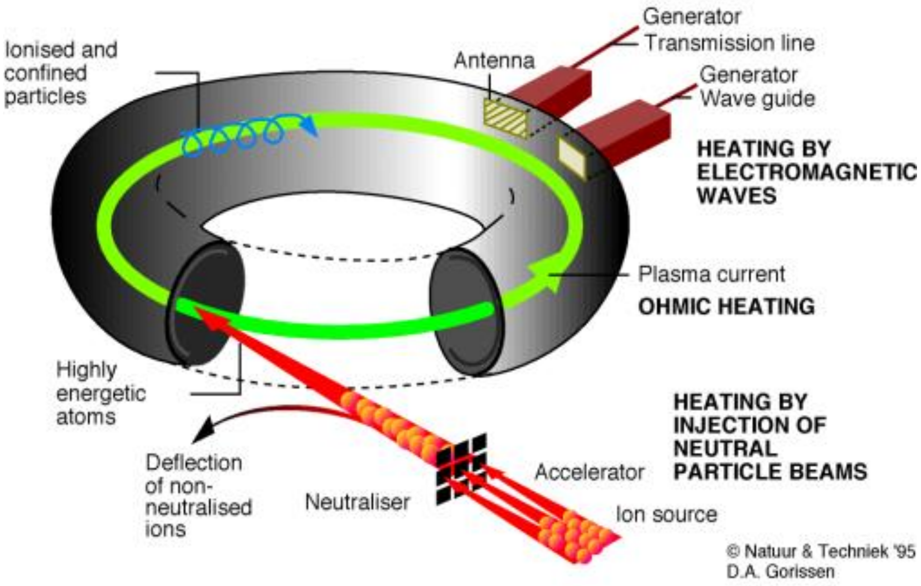
Actuators that affect plasma state



Magnetic Coils

- Central solenoid ramp rate
- Toroidal field coils
- Poloidal field coils
- 3D field coils to perturb toroidal symmetry

Actuators that affect plasma state

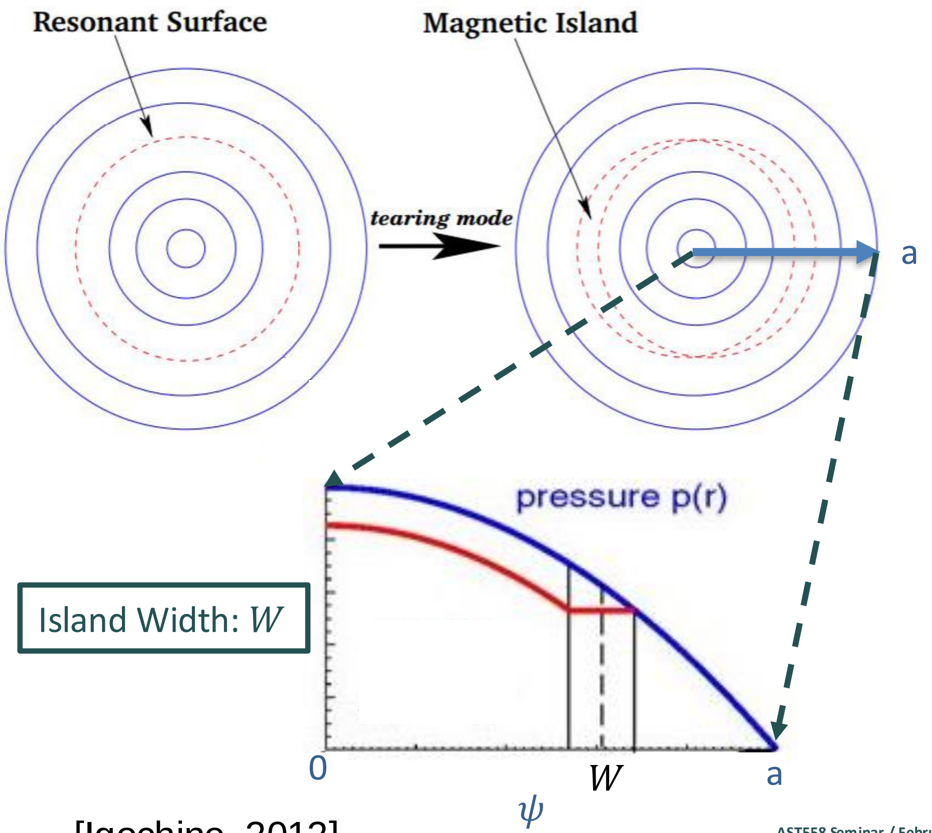


Magnetic Coils
- Central solenoid ramp rate
- Toroidal field coils
- Poloidal field coils
- 3D field coils to perturb toroidal symmetry

Heating Sources
- Neutral Beams
- Electron Cyclotron Heating
- Other RF Waves (Helicon + Lower Hybrid)

Gases
- Gas valves
- Pellet injection

What are tearing modes?



Island Width: W

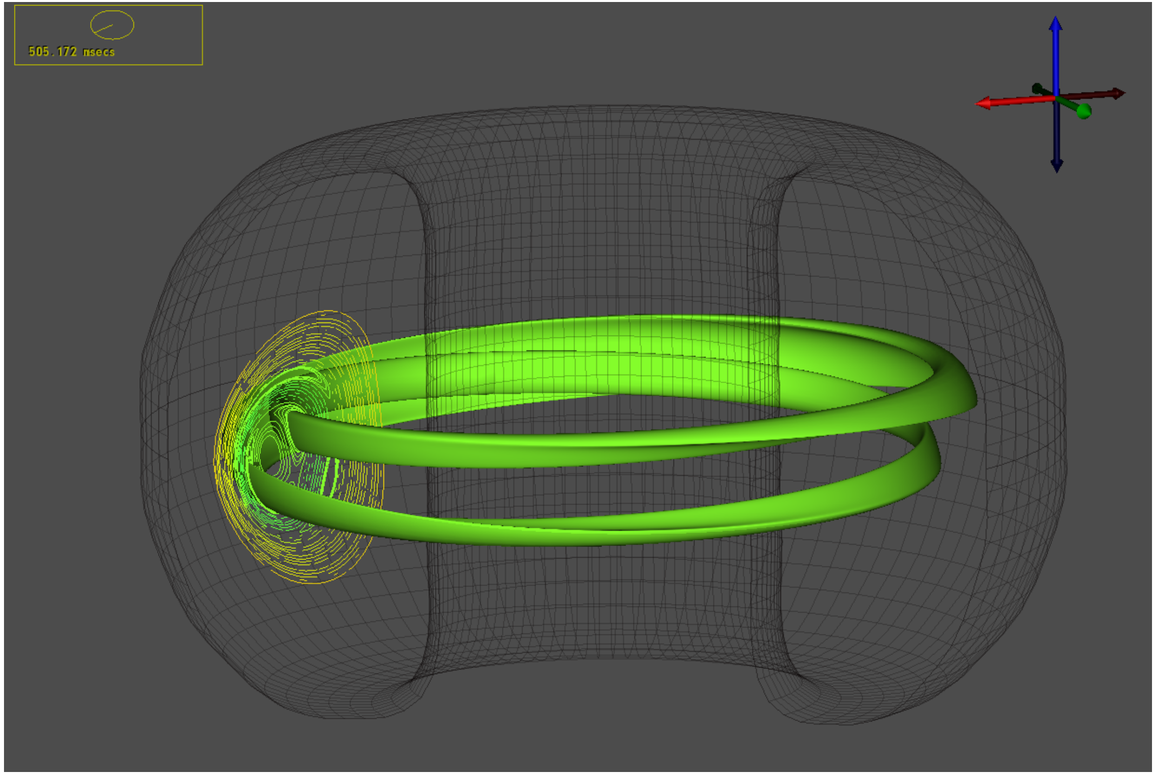
- Magnetic field reconfigures to lowest energy
- Occurs at rational surfaces
- Breaks nicely nested flux surfaces

So why do we care?

- "Short circuits" transport
- Modes can lock to wall → disrupts plasma

[Igochine, 2012]

Tearing Modes are 3D Structures

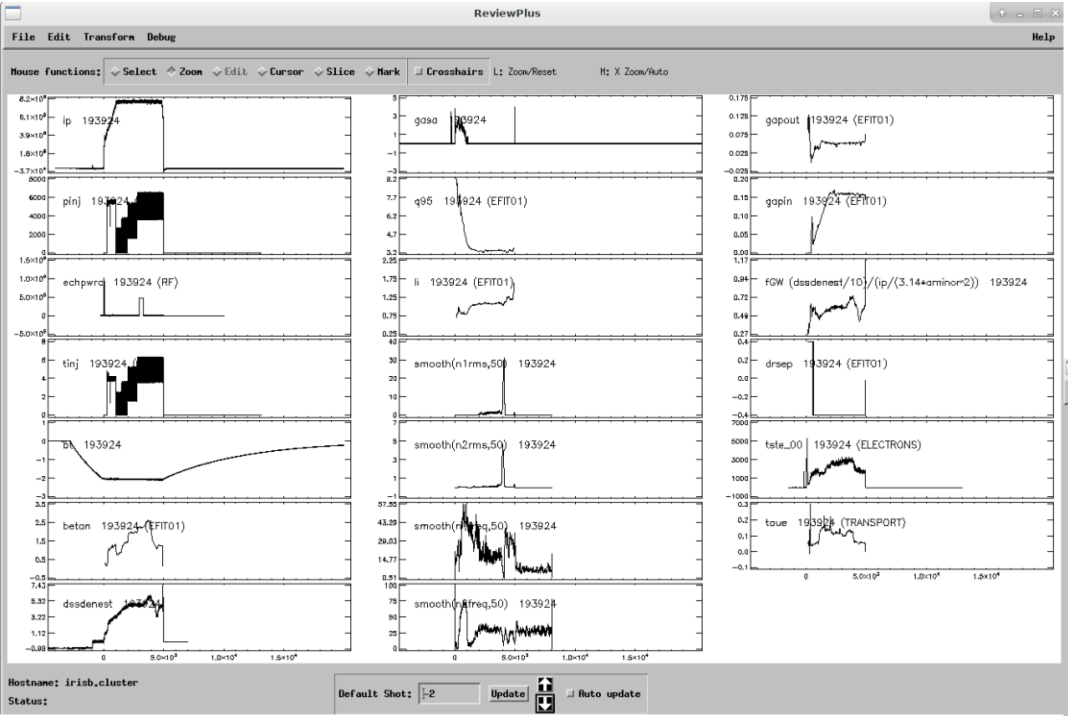


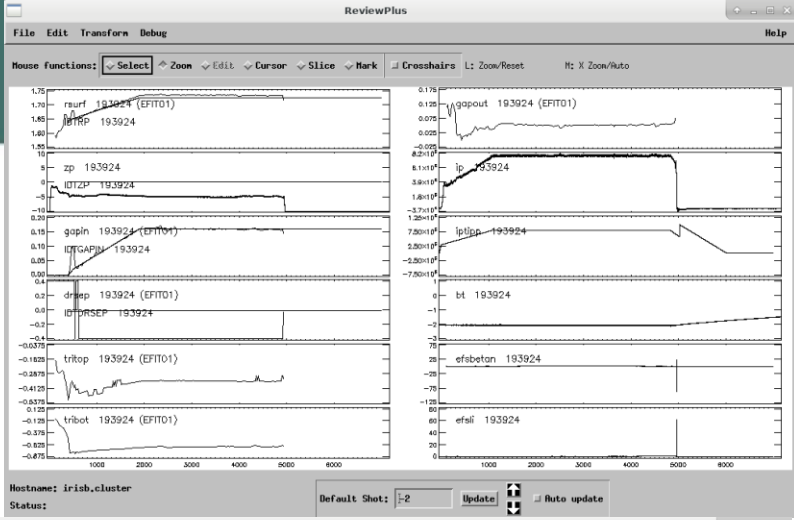
[VACET]

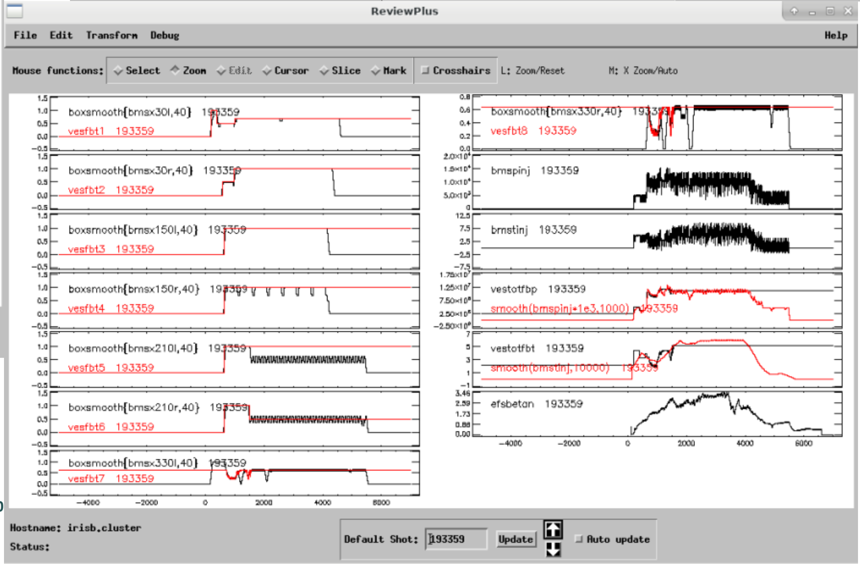
Why machine learning?

- Lots of data

Too much experimental data

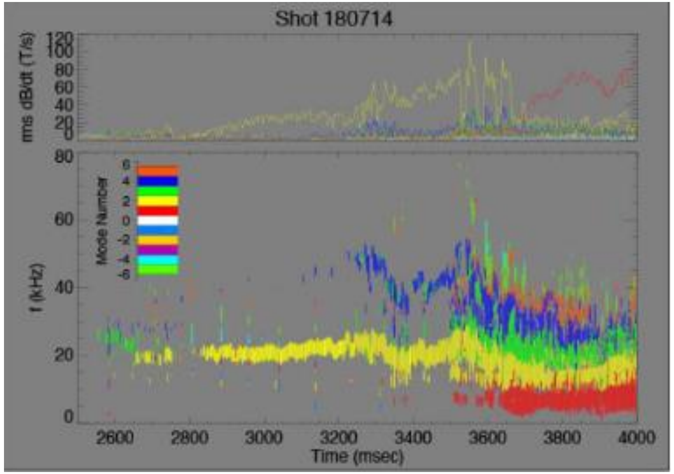


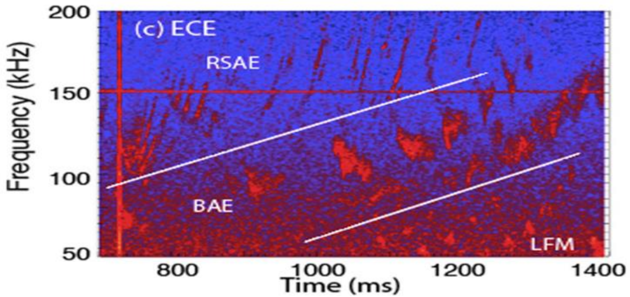




Why machine learning?

- Lots of data
- Models can be run real-time (ms time-scale)
- ML can find patterns to predict instabilities
- Can learn multi-actuator effects on the plasma





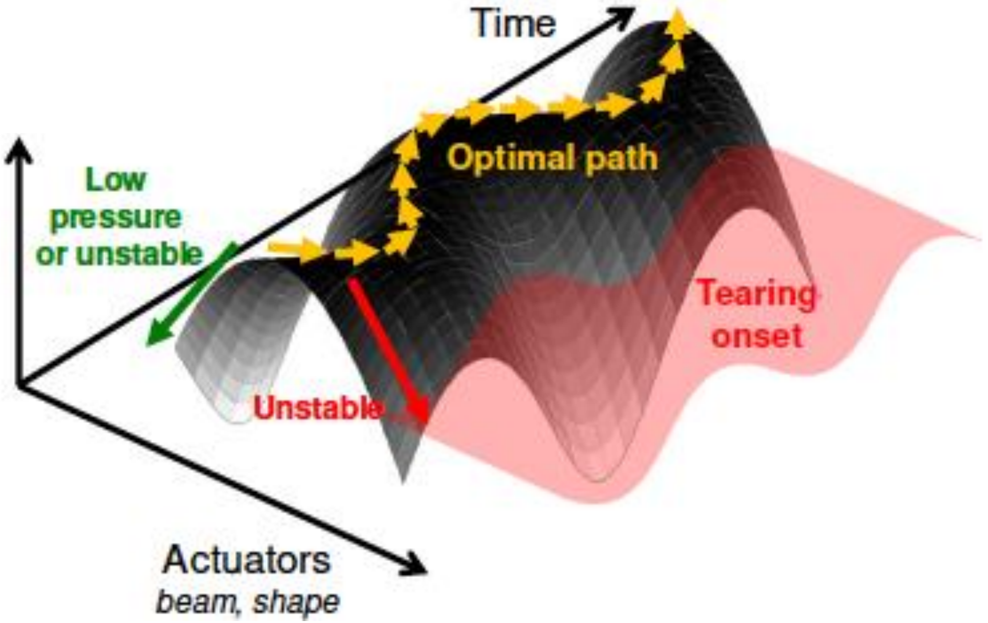
[Victor IAEA 2020]
[Heidbrink NF 2021]

Plasma Control

What should a good ML controller do?

What ML is not good for:
- Shot design - developing new scenarios
- Extrapolating to new regimes

What ML is good for:
- Maintaining stability in previously explored spaces
- Recovering from small deviations to optimized scenario



[Seo Nature 2024]

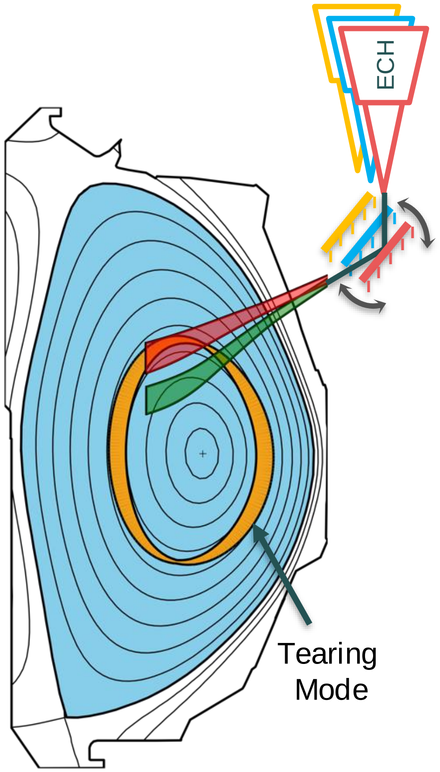
Plasma Control

Outline

- Crash course to tokamak experiments
- **Long term TM prediction and preemptive ECCD TM suppression and understanding ML predictions**
- Surrogate models for real-time control

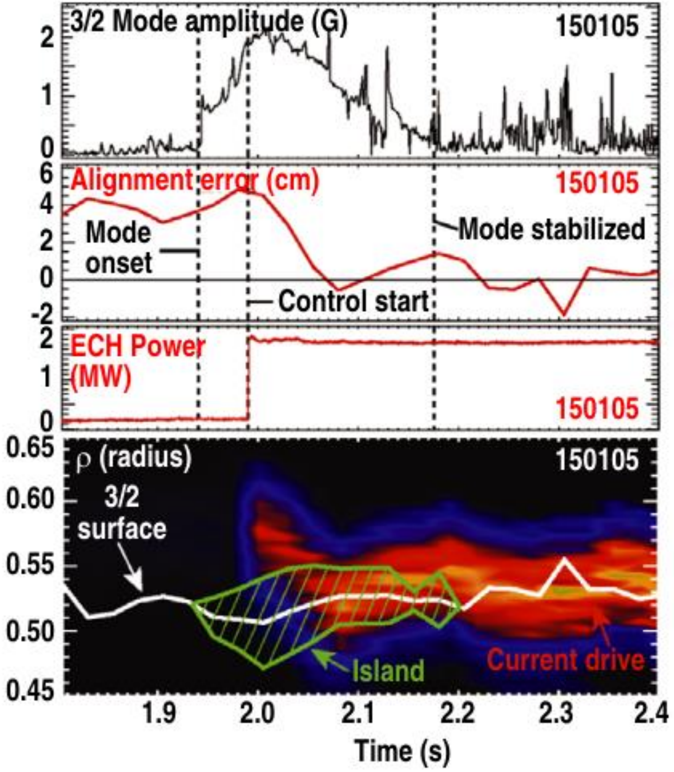
ECCD as the main actuator for TM suppression

- Electron cyclotron current drive (ECCD) has been shown to suppress TMs
  - Replaces missing bootstrap current to heal island
- Full preemption:
  - Feedforward ECH for TM control is wasteful and cannot handle dynamic scenario



A first step beyond full preemption

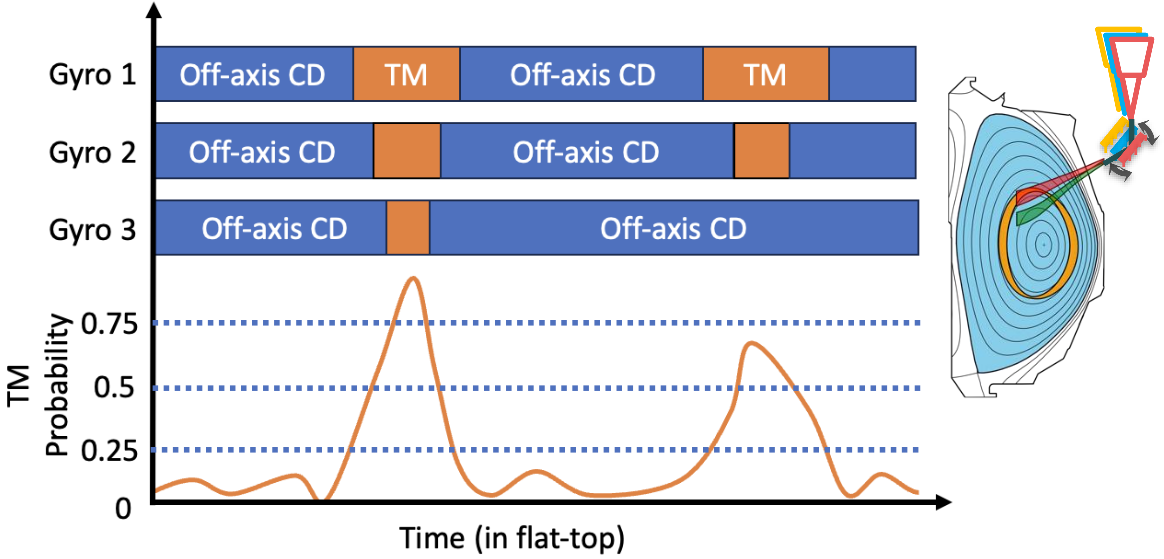
- Catch and Suppress
  - Must wait for large mode amplitude to not confuse with other magnetic activity
  - Lacks multi-tasking capabilities because all ECH mirrors must continuously follow rational surface



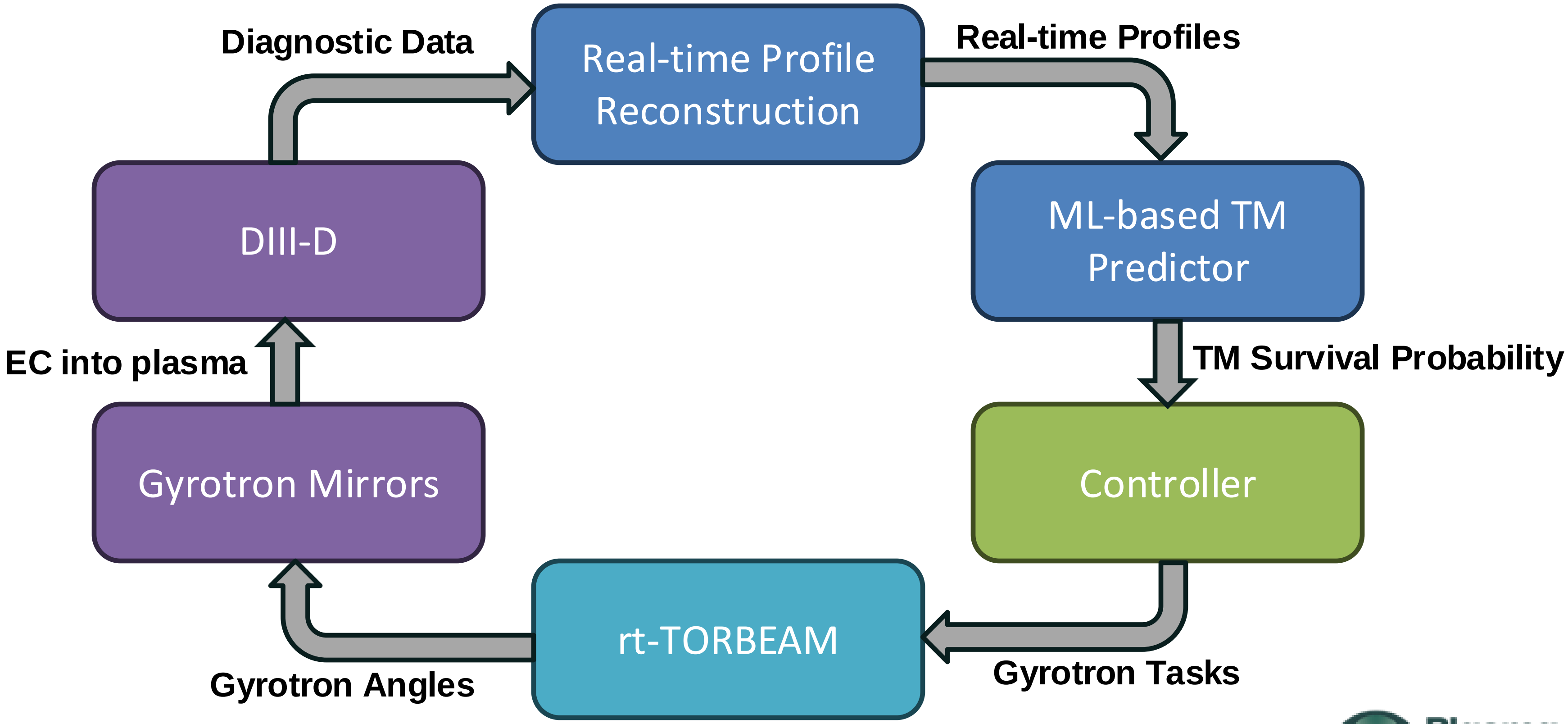
[Kolemen, NF 2014]

New control scheme capable of active TM control and EC multi-tasking

- With sufficient TM prediction time, can steer to rational surfaces when needed
- Need accurate aiming for active TM suppression task
- Allows for multi-task gyrotrons to reach other scenario goals



Closed control loop uses RTCAKENN profiles for feedback control



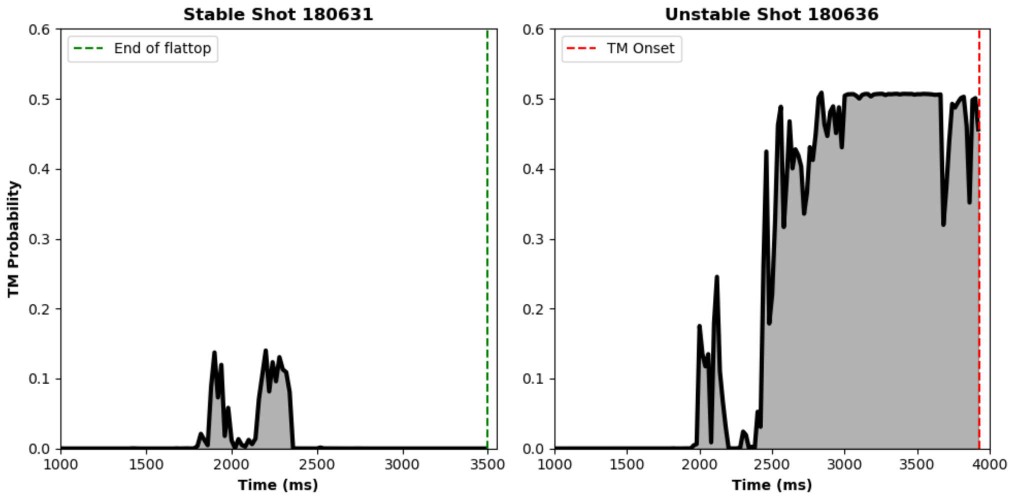
ML Model predicts TM survival probability using RTCAKENN profiles

Profile Inputs

- Electron temperature (T_e)
- Electron density (n_e)
- Ion temperature (T_i)
- Rotation (v_{tor})
- Safety Factor (q)
- Pressure (p)
- Current density (J)

Scalar Inputs

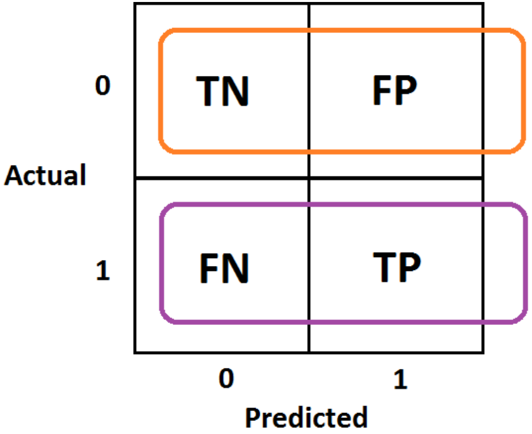
- Shape parameters: δ_{top}, δ_{bot}, κ, R_{axis}, a_{minor}
- RTEFIT scalars: q_{min}, β_N, l_i, V_{plas}
- P_{NBI}, T_{NBI}, P_{ECH}, I_P, B_T,



TM probability: probability of TM occurring over $t_{horizon}$=1000ms

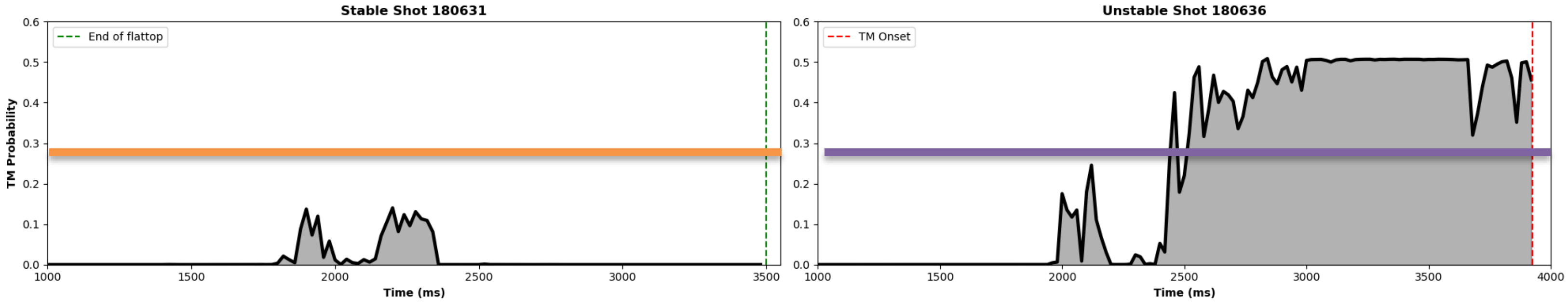
Survival models: event prediction models applied from medical fields

Assessing Event Prediction Models



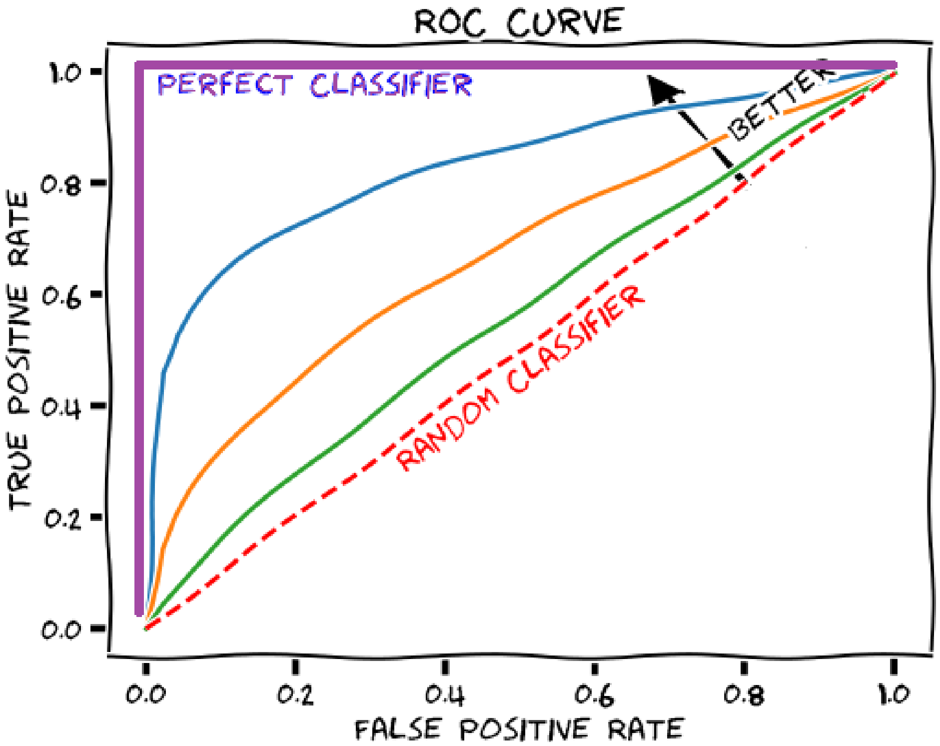
False Positive Rate (FPR) = $\dfrac{FP}{(FP + TN)}$

True Positive Rate(TPR) = $\dfrac{TP}{(TP + FN)}$



AUC Metric

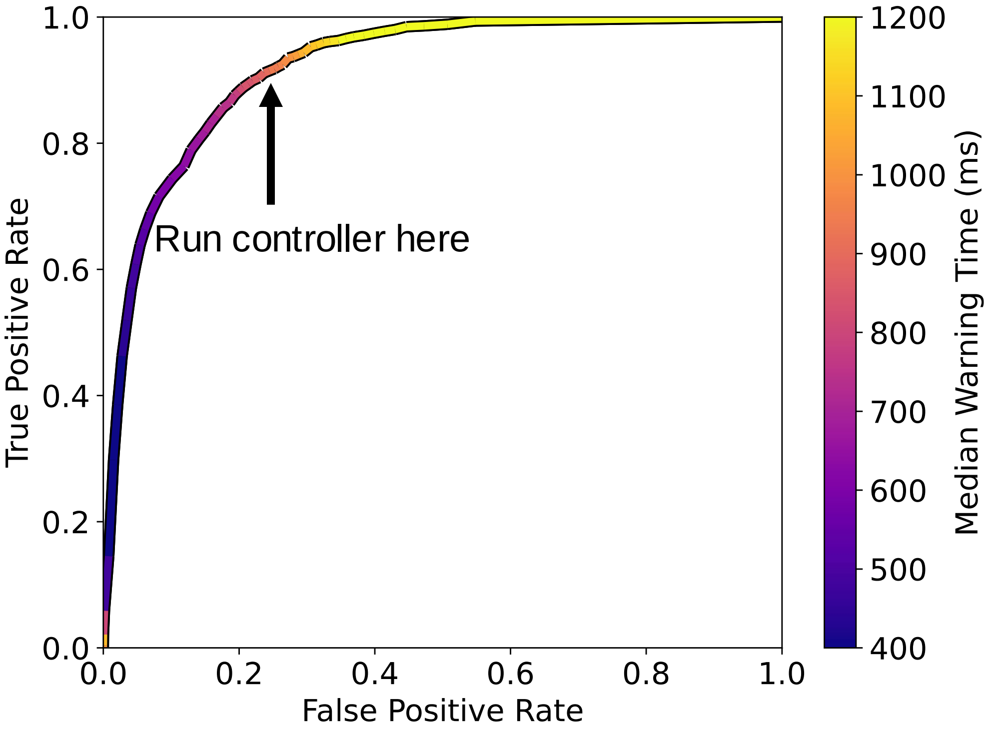
- AUC metric integrates TPR by sweeping threshold from $1 \to 0$
  - FPR sweeps from $0 \to 1$
- AUC values:
  - Perfect classifier = 1
  - Random classifier = 0.5



TM predictor performance further demonstrated with event prediction metrics

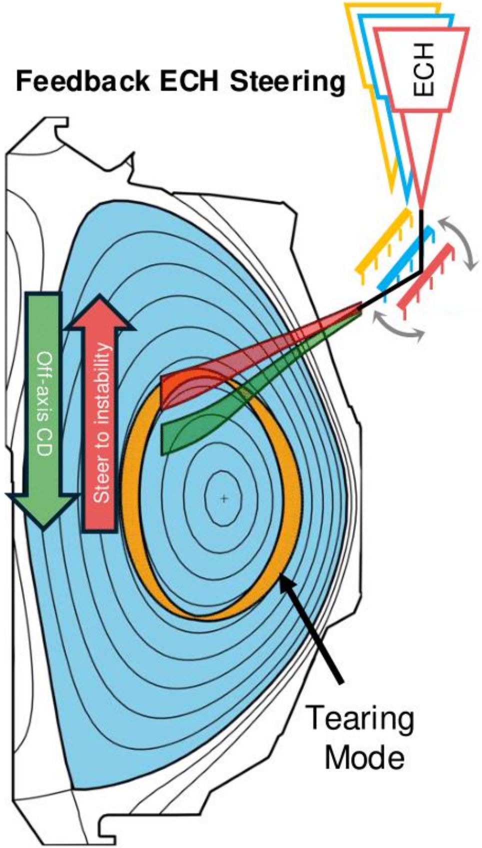
- AUROC Curve:
  - False positives are not bad because moving EC is not a costly action
  - False negatives are very costly (miss a TM) so pick high TPR
- Higher TPR and FPR also typically gives larger warning times

$$TPR = \frac{TP}{TP + FN} \qquad FPR = \frac{TN}{FP + TN}$$



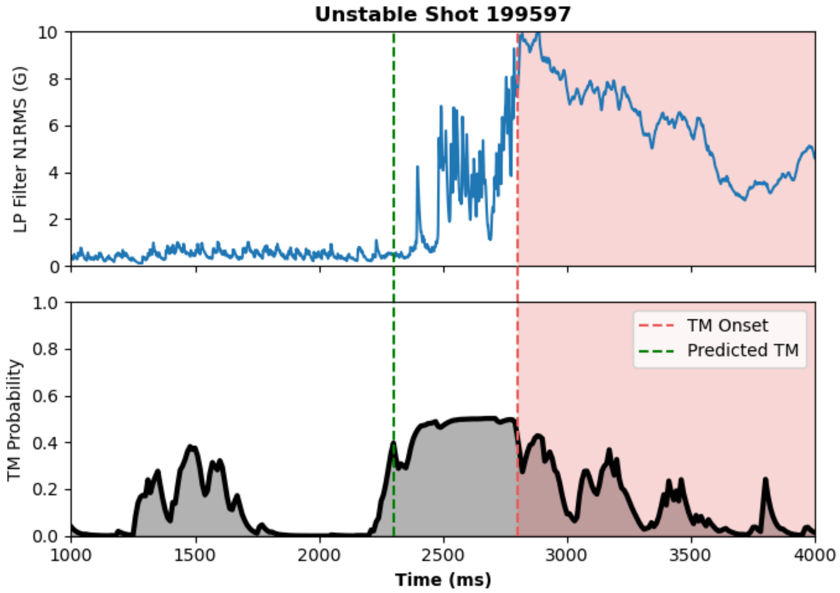
Control demonstrated on elevated q_{min} scenario which requires multi-tasking EC control

- Scenario performance requires reversed shear q-profile
  - Needs broad off-axis ECCD
- High performance generates highly unstable TMs
  - Needs feedback-controlled rational surface tracking ECCD
- EC multi-tasking is required for high performance, stable scenario operation



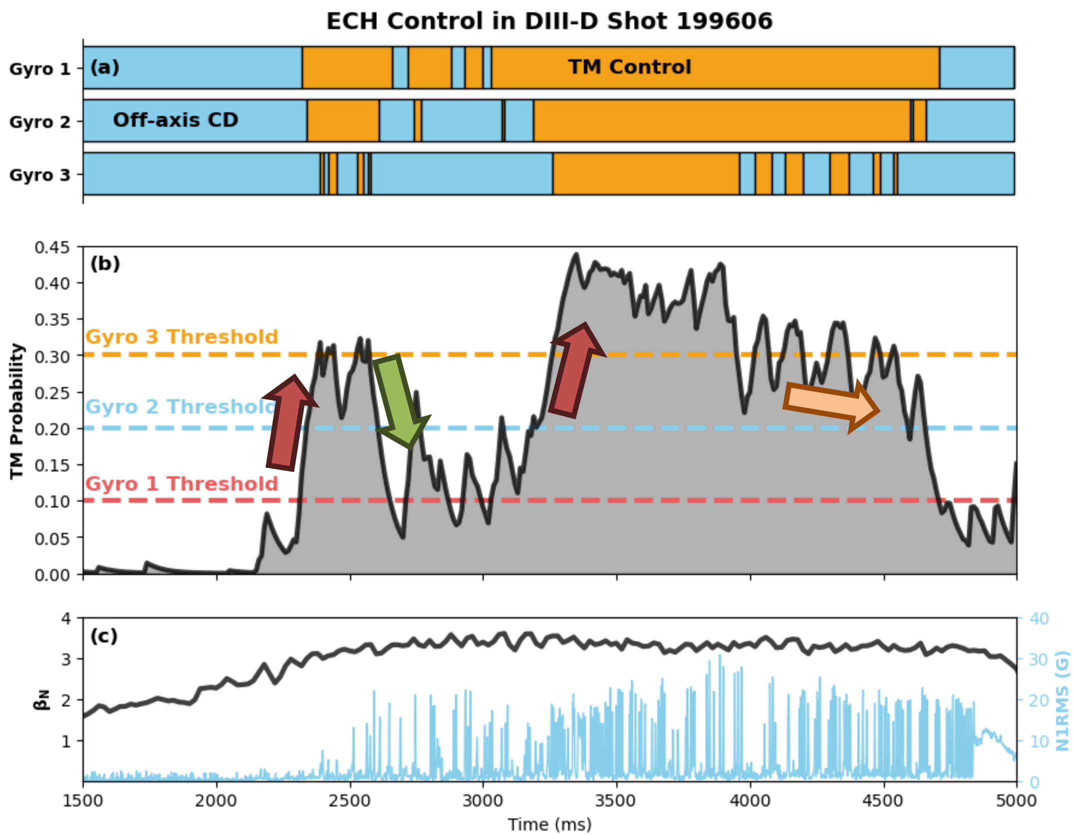
Unstable reference in elevated q_{min} scenario

- Reference shot with no active control, constant gyrotron location
- As expected, TM occurs in flattop phase
- Model correctly predicts with sufficient time to steer (>200ms)



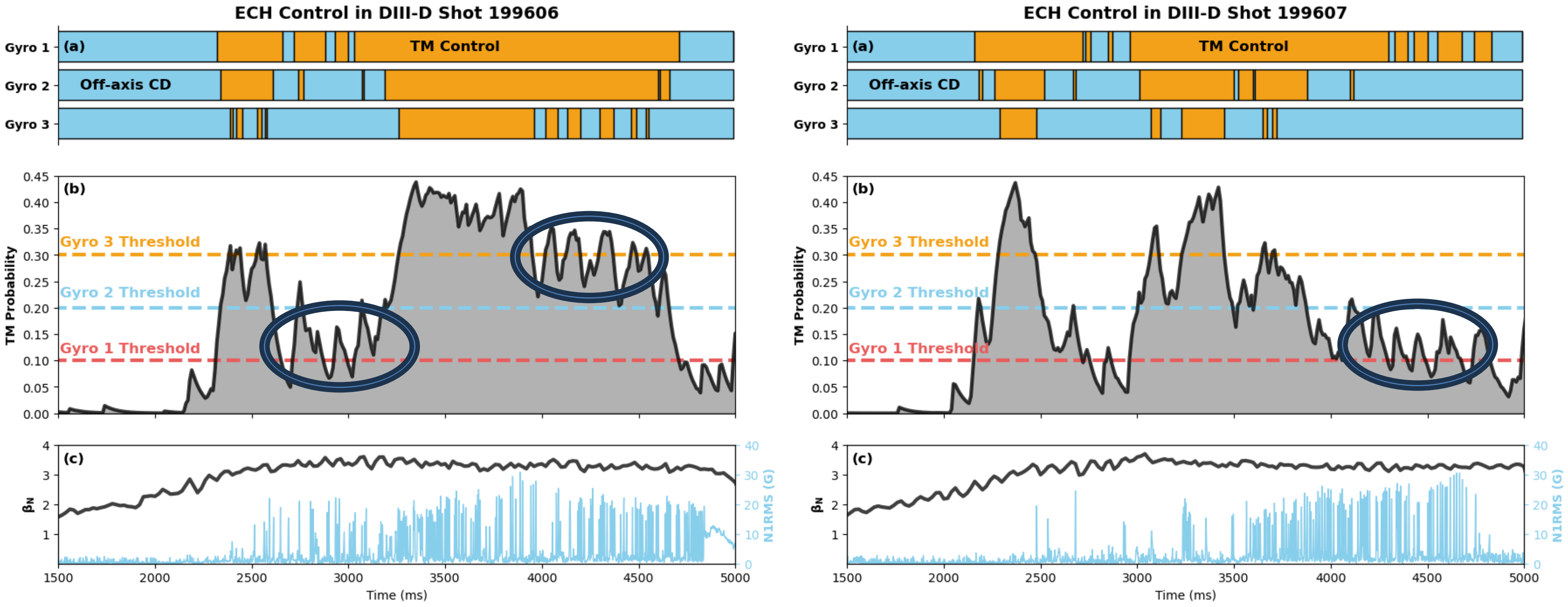
Plasma Control

Preemptive ECCD successfully avoid TMs



Plasma Control

Robustness: repeated successful control



Plasma Control

Experimental results are promising but more analysis is needed

- Physics analysis underway to better understand profile evolution, bootstrap currents and q=2 ECCD

- Control appears to have changed scenario stability rather than suppressing seed islands when the form
  - Mode spectra show no signs of seed islands, however very noisy due to fishbone activity

- How can we understand the ML TM stability predictions?

How to understand ML predictions: Game Theory

- Example problem: Predicting football team win rates

Team	Team 1	Team 2
Cost (millions)	\$25	\$24
Age (years)	22	23
Injury count	8	5
Win rate	0.58	0.57

Team 1

Team 2

$$\text{winRate} = \frac{1}{\text{norm}}\left[2 \cdot (\text{Cost}_{\text{squad}})^2 - 10 \cdot (\text{Age}_{\text{avg}} - 24)^2 - 20 \cdot (\text{Num}_{\text{injury}})\right]$$

Plasma Control

How to understand ML predictions: Game Theory

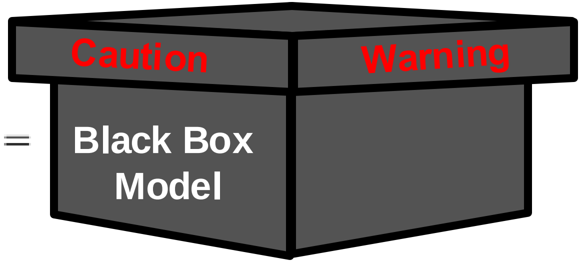
- Example problem: Predicting football team win rates

Team	Team 1	Team 2
Cost (millions)	$25	$24
Age (years)	22	23
Injury count	8	5
Win rate	0.58	0.57

Team 1

Team 2

$\text{winRate} =$

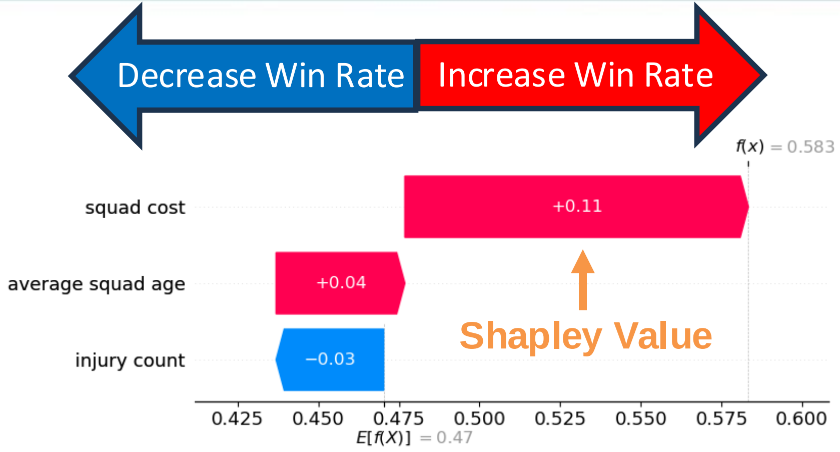


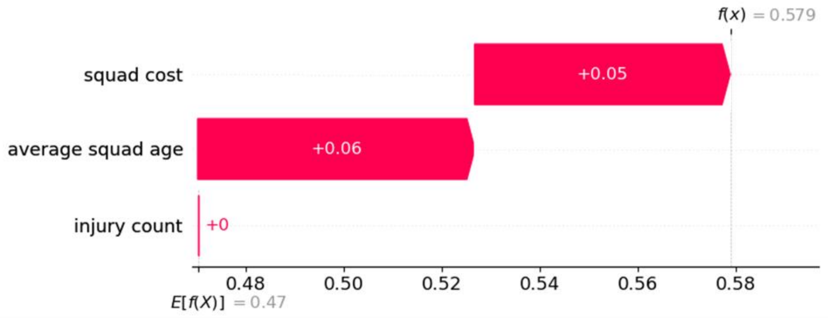
Assigning importance values to inputs: Shapley values

Team	Team 1	Team 2
Cost	$25	$24
Age	22	23
Injury count	8	5
Win rate	0.58	0.57

Team 1

Team 2





Reference distributions matter for Shapley values

- Example problem:
  - Predicting football team win rates **<u>of professionals</u>**

Team 1

VS

Team 2

Plasma Control

Reference distributions matter for Shapley values

- Example problem:
  - Predicting football team win rates **<u>of professionals</u>**
- Example problem 2:
  - Predicting football team win rates **<u>compared to amateurs</u>**

Team 1

Team 2

VS

Comparing Team 1 to different distributions changes Shapley values

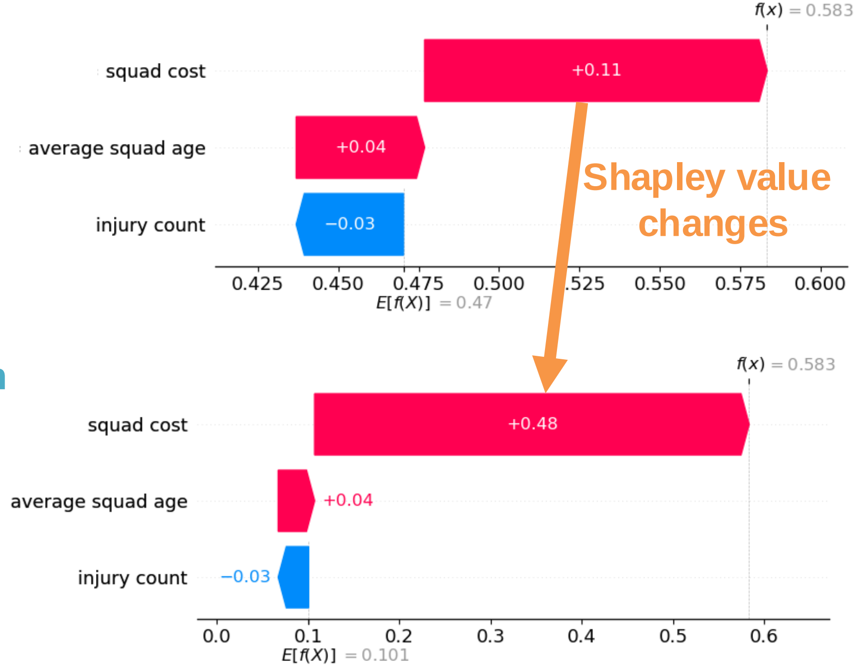
Background Distributions

League	Pro	Amateur
Cost	$15-30	$13-15
Age	20-30	20-30
Injury count	0-10	0-10

Team	Team 1
Cost	$25
Age	22
Injury count	8
Win rate	0.58

Team 1 vs Pro Team

Team 1 vs Amateur Team



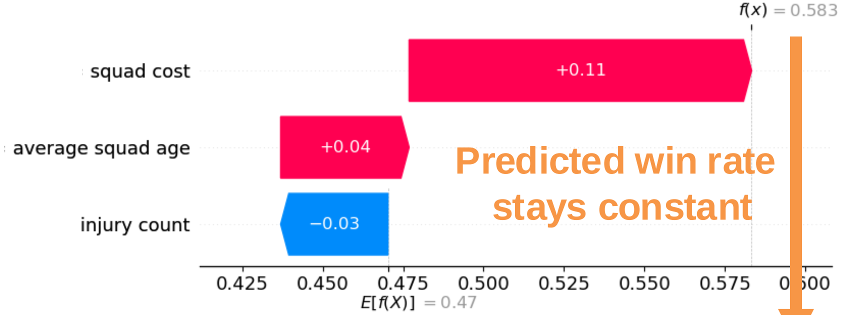
Comparing Team 1 to different distributions changes Shapley values

Background Distributions

League	Pro	Amateur
Cost	$15-30	$13-15
Age	20-30	20-30
Injury count	0-10	0-10

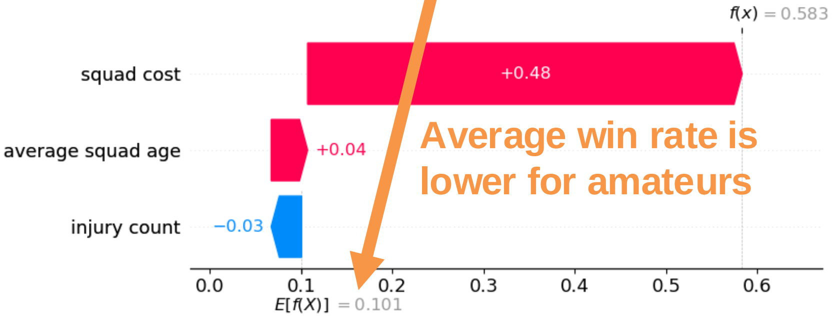
Team	Team 1
Cost	$25
Age	22
Injury count	8
Win rate	0.58

Team 1 vs Pro Team



Predicted win rate stays constant

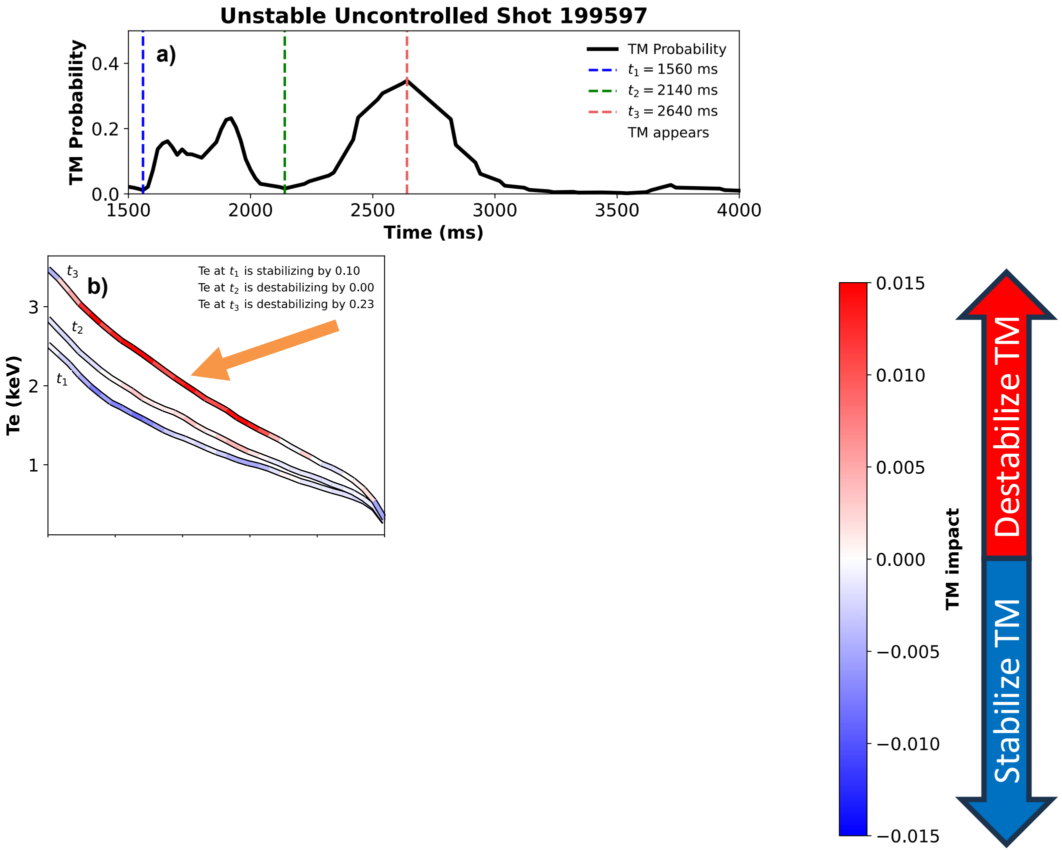
Team 1 vs Amateur Team



Average win rate is lower for amateurs

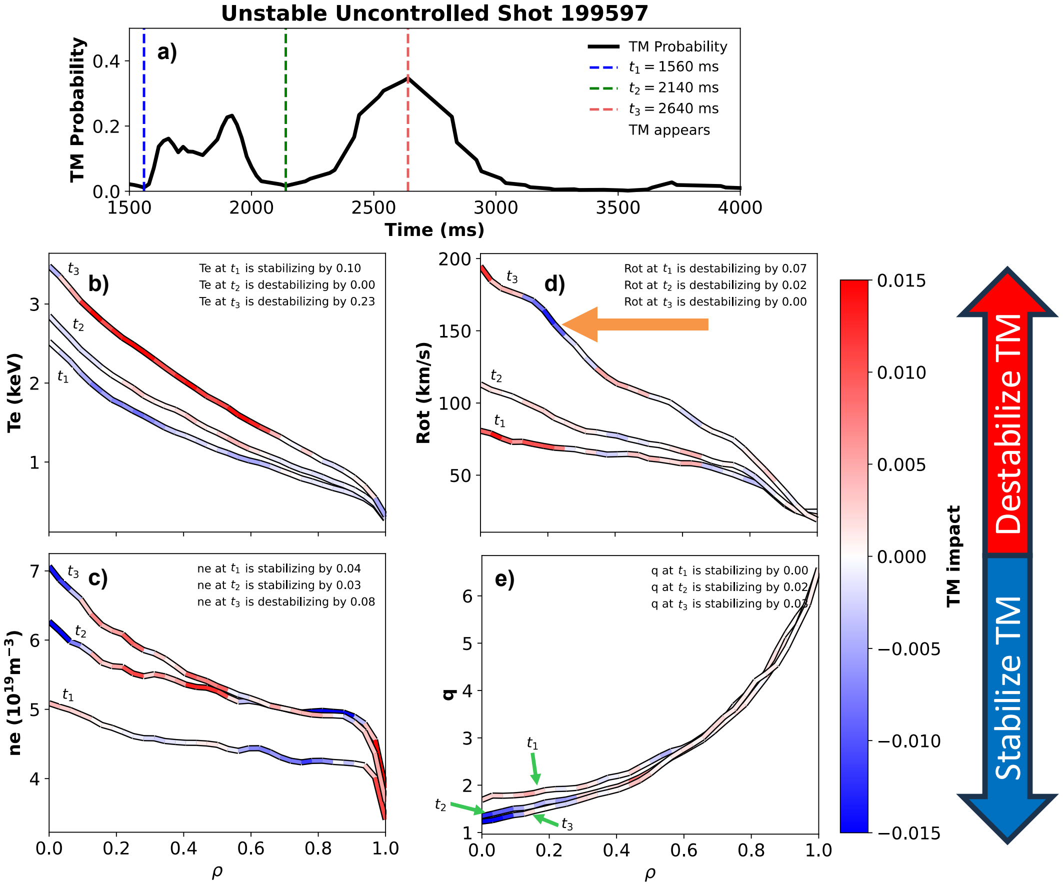
Shapley values tell us why unstable shots are unstable

- Background distribution:
  - All shots in DIII-D history
  - All advanced scenario shots
  - **All shots in single experiment**
- Higher T_e destabilizing



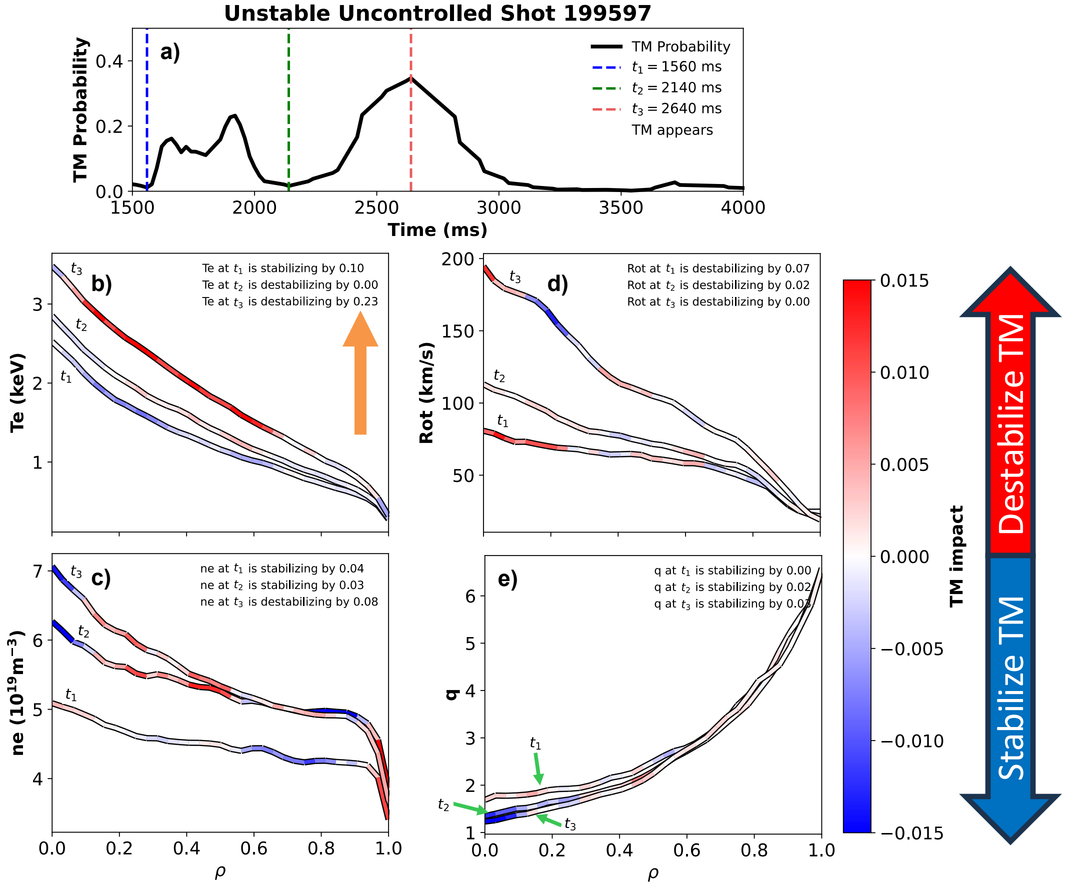
Shapley values tell us why unstable shots are unstable

- Background distribution:
  - All shots in DIII-D history
  - All advanced scenario shots
  - **All shots in single experiment**
- Higher T_e destabilizing
- Higher Rot (ω_{tor}) has stabilizing effect



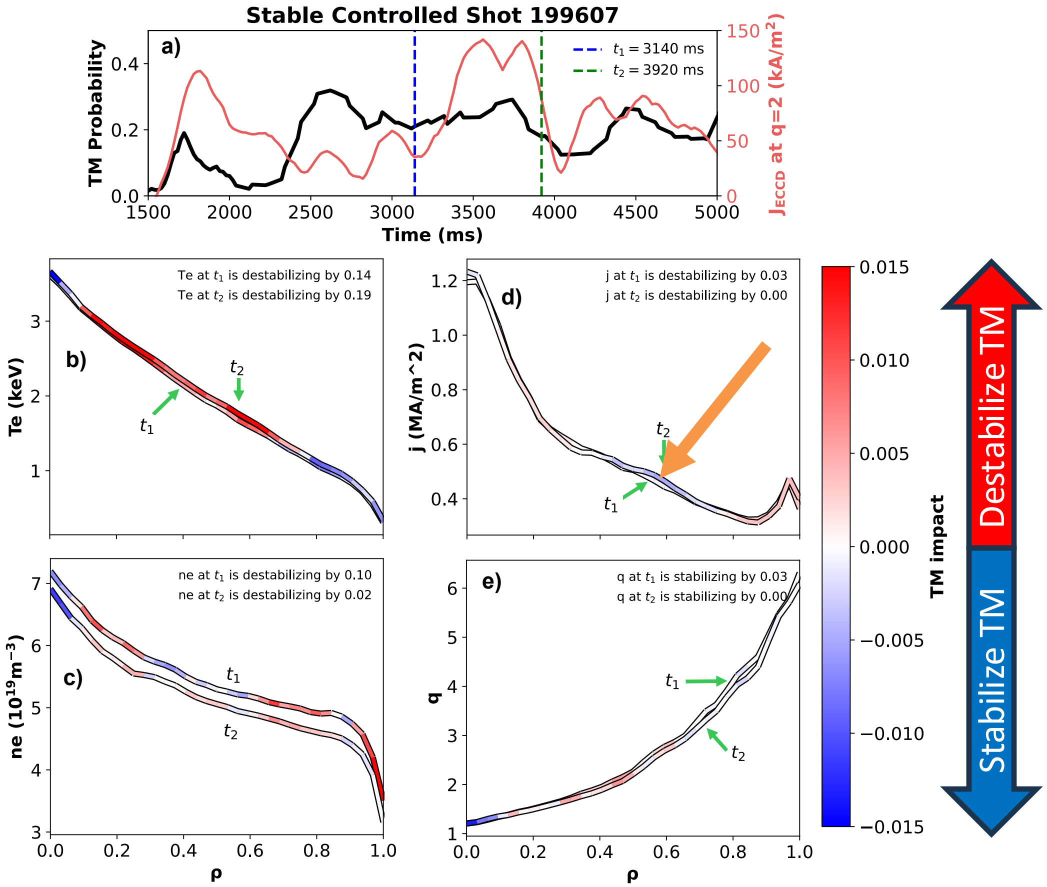
Shapley values tell us why unstable shots are unstable

- Background distribution:
  - All shots in DIII-D history
  - All advanced scenario shots
  - **All shots in single experiment**
- Higher T_e destabilizing
- Higher Rot (ω_{tor}) has stabilizing effect
- T_e profile as whole has largest Shapley values



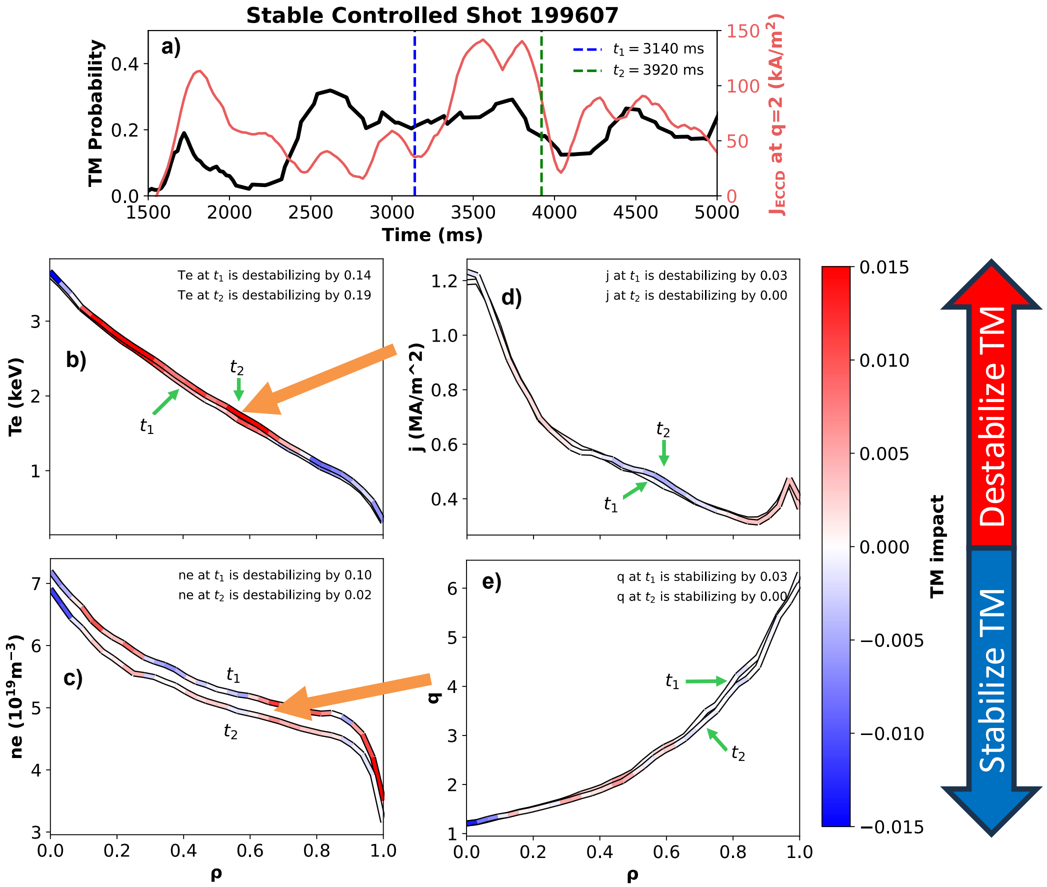
Shapley values show TM impact response to actuation

- Feedback control is changing localized ECH+ECCD deposition location
- Bump in current profile around $\rho \sim 0.6$ stabilizing



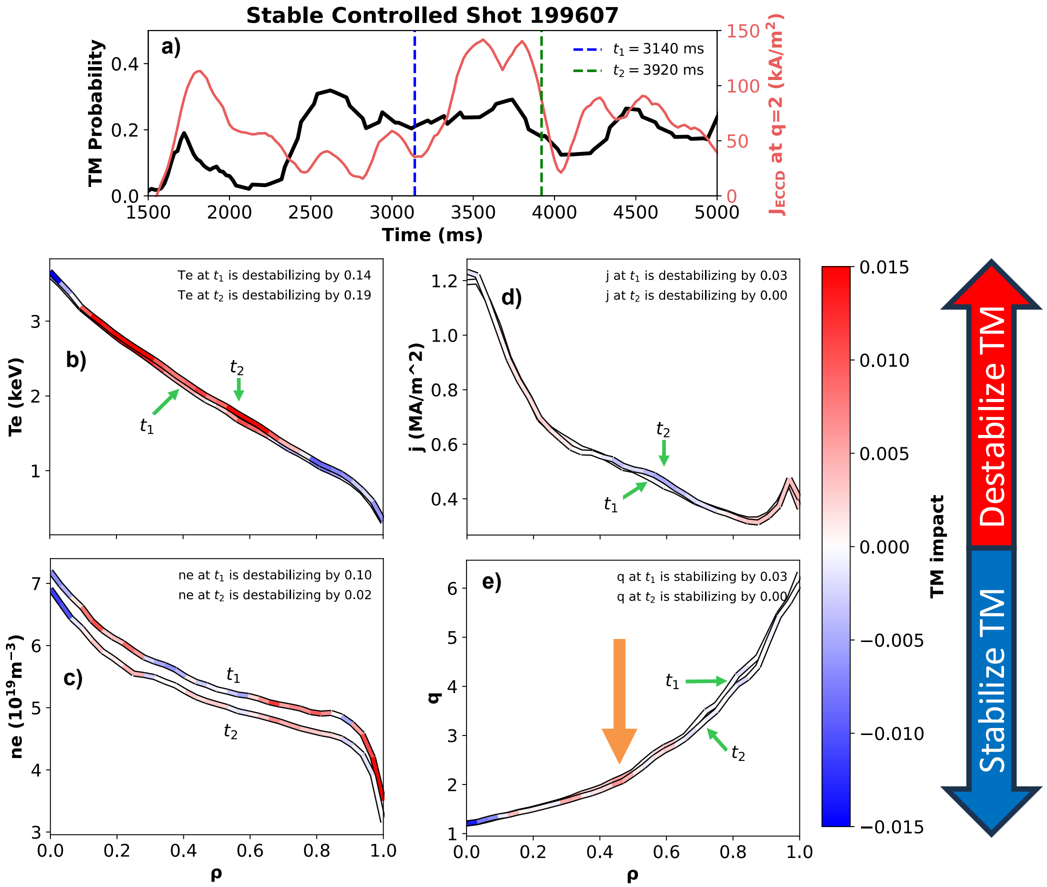
Shapley values show TM impact response to actuation

- Feedback control is changing localized ECH+ECCD deposition location
- Bump in current profile around $\rho \sim 0.6$ stabilizing
- ECH causes slight increase in T_e and density pump out effect



Shapley values show TM impact response to actuation

- Feedback control is changing localized ECH+ECCD deposition location
- Bump in current profile around $\rho \sim 0.6$ stabilizing
- ECH causes slight increase in T_e and density pump out effect
- Value around q=2 shows destabilizing effect



Outline

- Crash course to tokamak experiments
- Long term TM prediction and preemptive ECCD TM suppression and understanding ML predictions
- **Surrogate models for real-time control**

What is a surrogate model?

- We have a costly function $f(x)$
- Want to approximate as simpler function
$$\bar{f}(x) \approx f(x)$$
- For simple functions, Taylor series is an easy surrogate model
$$\bar{f}(x) \approx f(0) + f'(0)x + \cdots$$
- For more complex functions or models, need more complicated surrogates
  – Giant matrices with many basis functions (aka ML)

Outputs:
- Fitted profiles
- Transport results
- Results from physics codes

Inputs:
- Raw data
- Diagnostics

$$y = f(x)$$

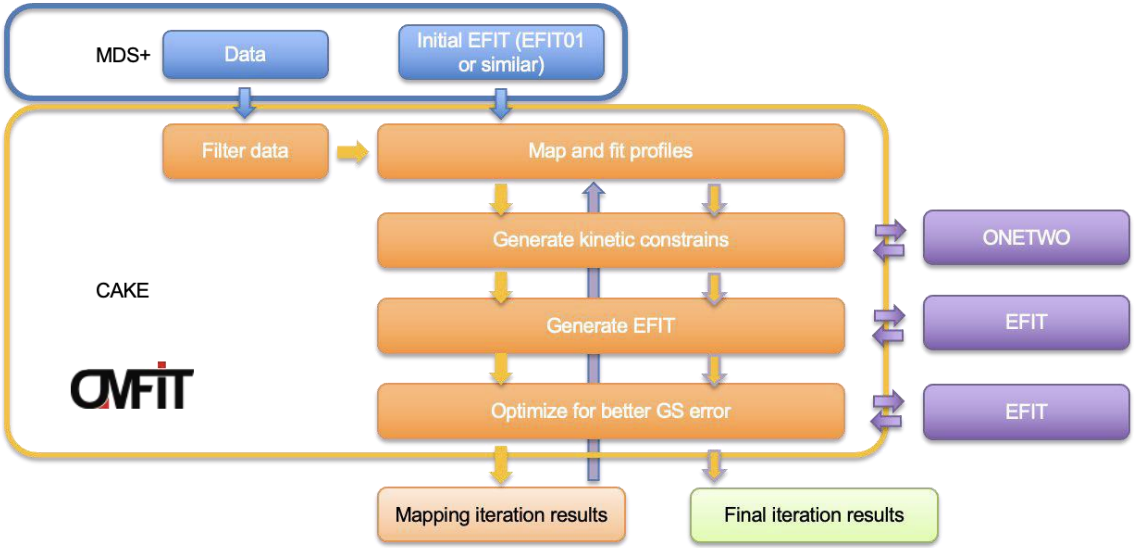
$$y \approx \bar{f}(x)$$

Surrogate model:
- Improve weights of basis functions iteratively by gradient descent
- Exact form of model "tuned" to improve accuracy

Plasma Control

Introduction to CAKE kinetic equilibria

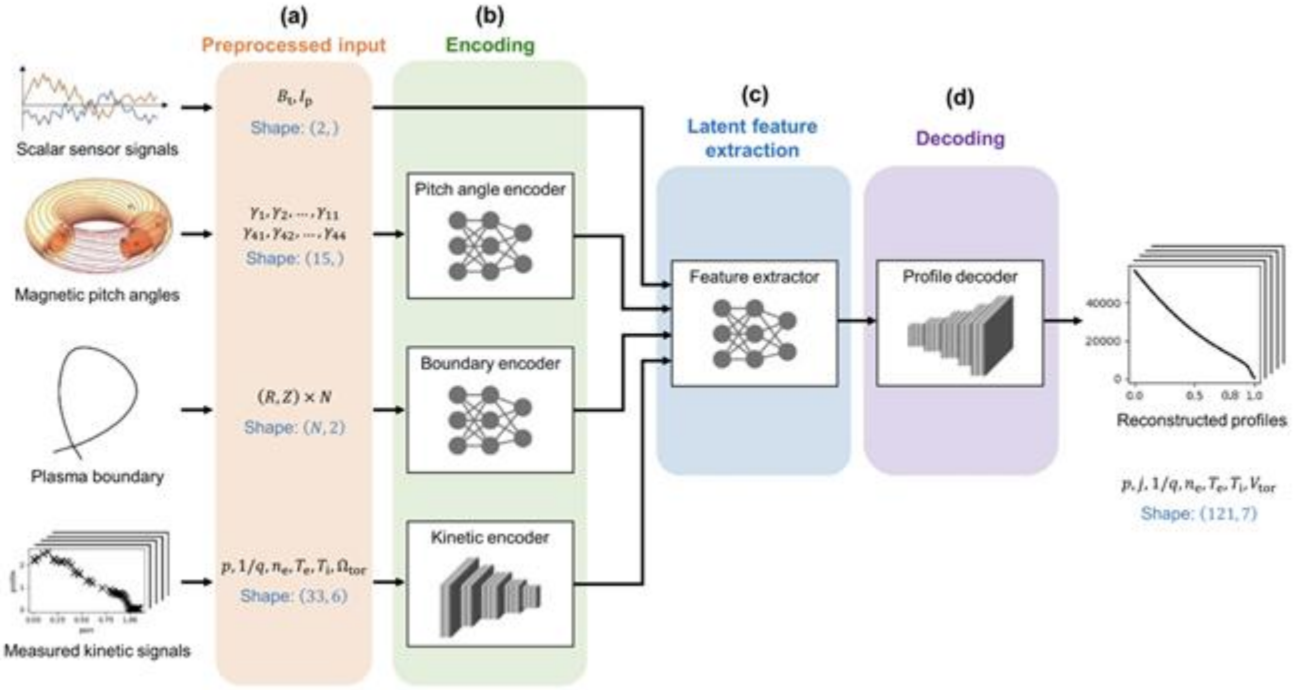
- Consistent Automatic Kinetic Equilibria (CAKE) provides automated, high-fidelity equilibria on a database scale [Xing et al. FED 2021]
- Great for database analysis (>67,000 equilibria) or generating consistent results to use with machine learning



[Courtesy of A. Xing]

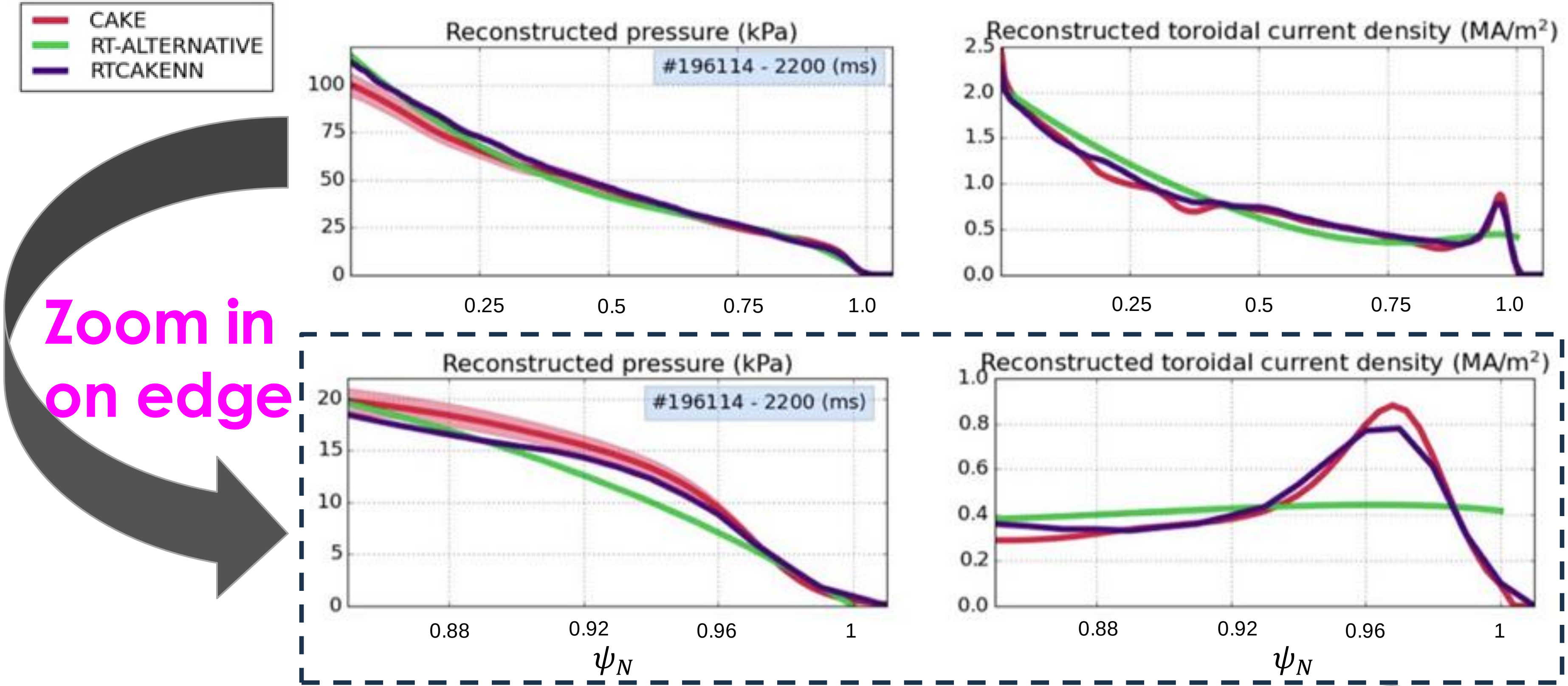
RTCAKENN: CAKE-quality profiles in real-time using ML

- ML model for real-time CAKE to derive high quality rt-profiles
- Trained on 19,000 CAKE equilibria database



Plasma Control

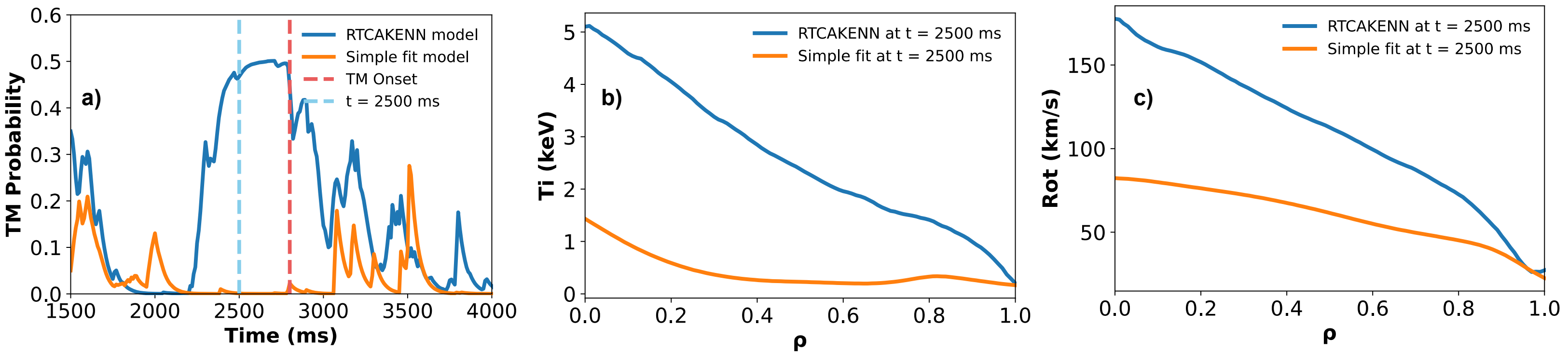
RTCAKENN gives improved real-time profiles



Plasma Control

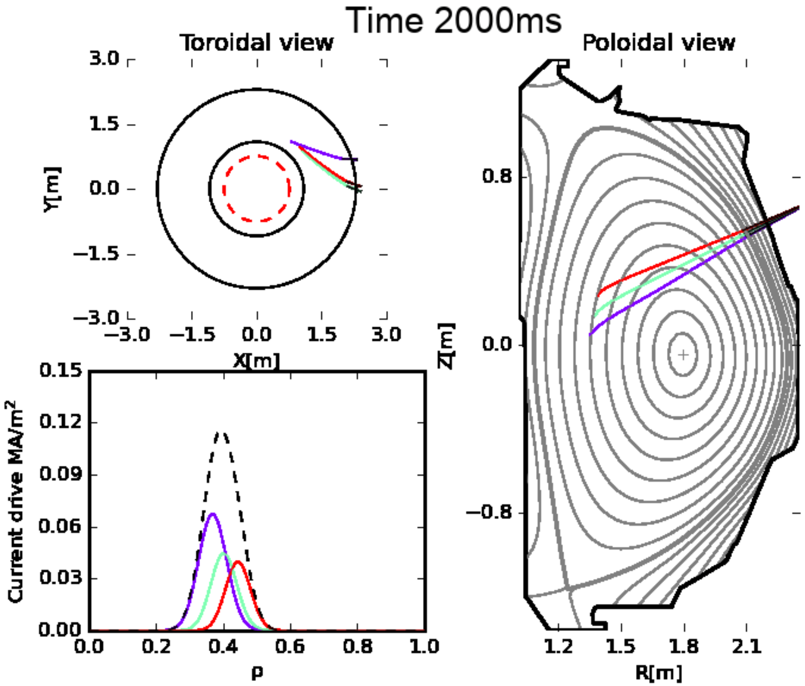
RTCAKENN models robust to diagnostic dropout

- During experiment, CER diagnostic had failure causing poor fits for T_i and Rot (ω_{tor}) profiles
- RTCAKENN shows robustness to still predict accurate profiles when classical fitting routine fails
- TM prediction using RTCAKENN remain accurate while classical fitting fails



Torbeam code provides ray tracing and ECH deposition profiles

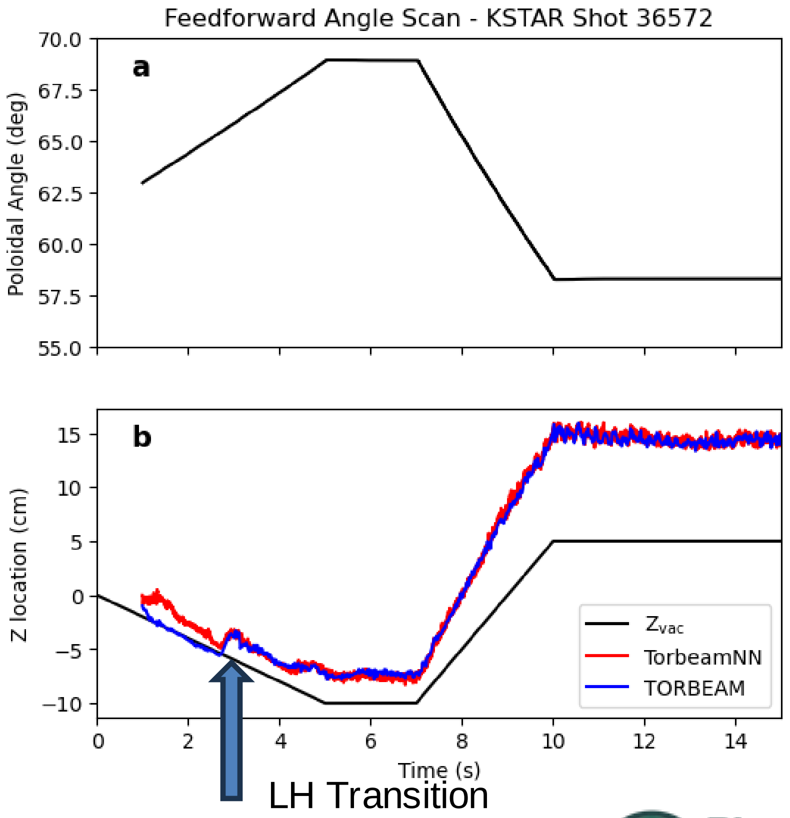
- Torbeam calculates trajectory of electron cyclotron heating (ECH) waves
  - Also calculates absorption resonant location and heating and current drive profiles
- Real-time version runs in 10-20ms with only information about maximum deposition location
  - No deposition profiles or CD efficiency



Plasma Control

TorbeamNN: ML surrogate model speeds up ray-trace calculation

- TorbeamNN on KSTAR provides accurate information about ECH absorption location
- Coming to DIII-D soon!
  - Include full ECH deposition profiles and current drive



Plasma Control

Summary

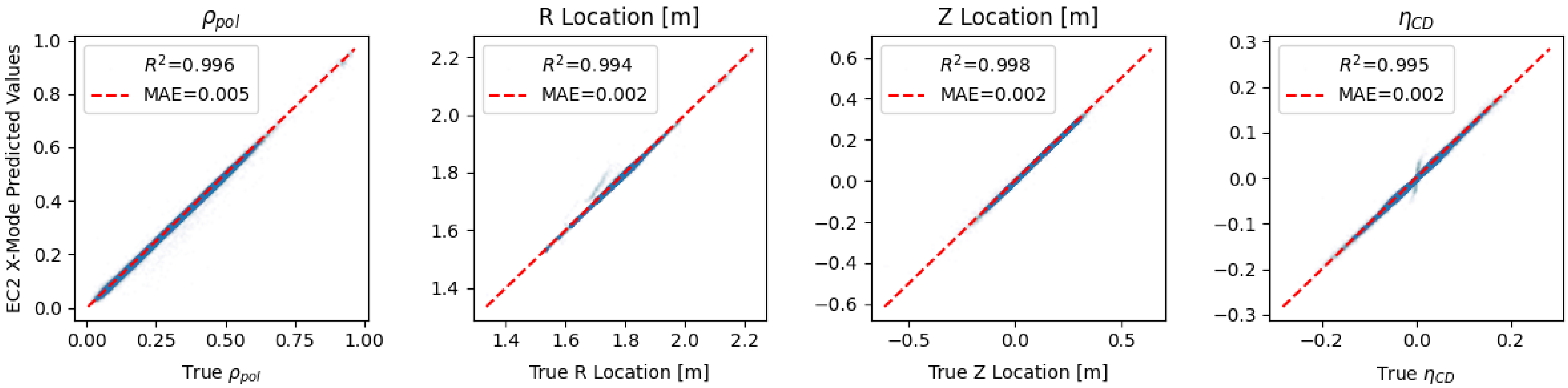
- TM prediction models provide long warning time for TM events and enables multi-tasking gyrotrons
  - Can now better balance TM avoidance with high performance scenarios
- Shapley values provide an understanding of how ML models make their predictions
  - Framework can be used for any ML models
- High quality real-time profiles provide more accurate real-time profiles
- Dynamics models provide full shot prediction of profiles

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Fusion Energy Sciences, using the DIII-D National Fusion Facility, a DOE Office of Science user facility, under Award DE-FC02- 04ER54698. In addition this material was supported by the U.S. Department of Energy, under Awards DE-SC0015480.

Backup Slides

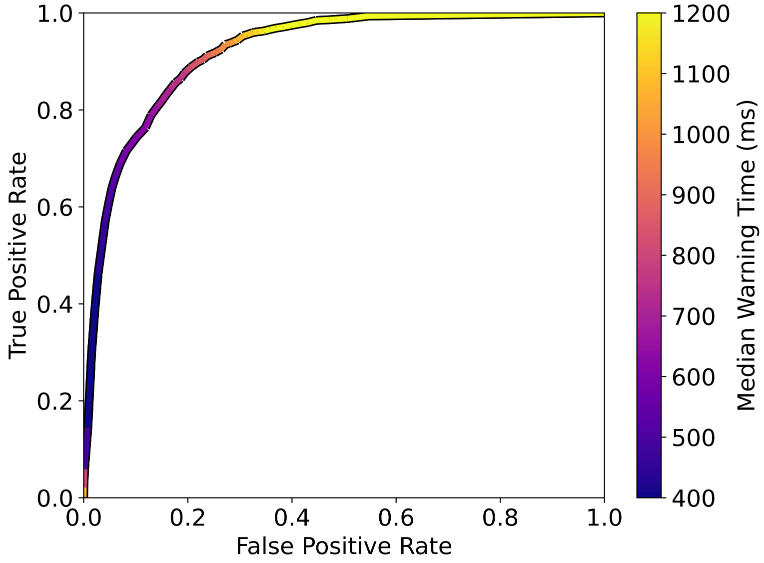
Misc Background

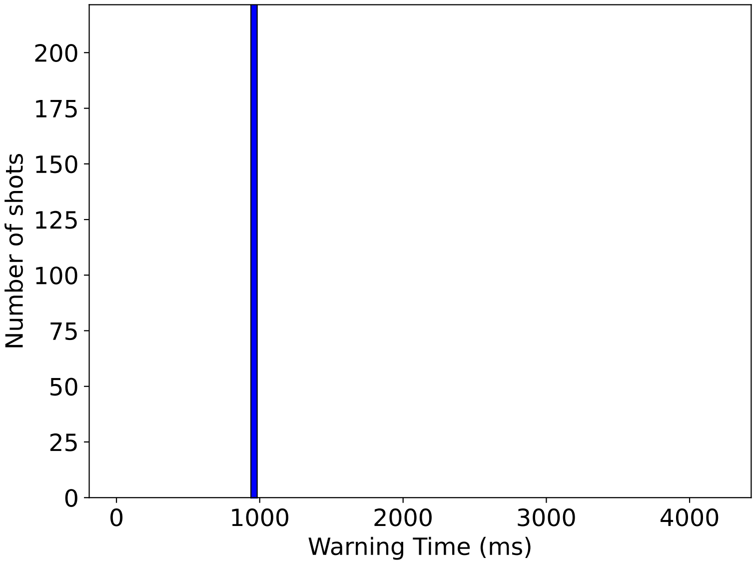
KSTAR TorbeamNN Accuracy



AUROC + Warning Times Histogram for TM predictor

- Vertical blue line is median

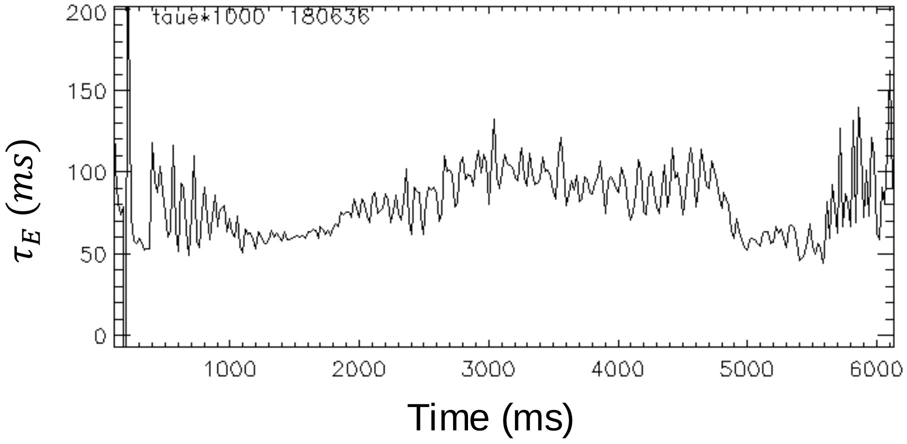




Experimental Timescales

Profile Evolution
- τ_E: 50-100ms
- τ_R: $\approx$ 1s

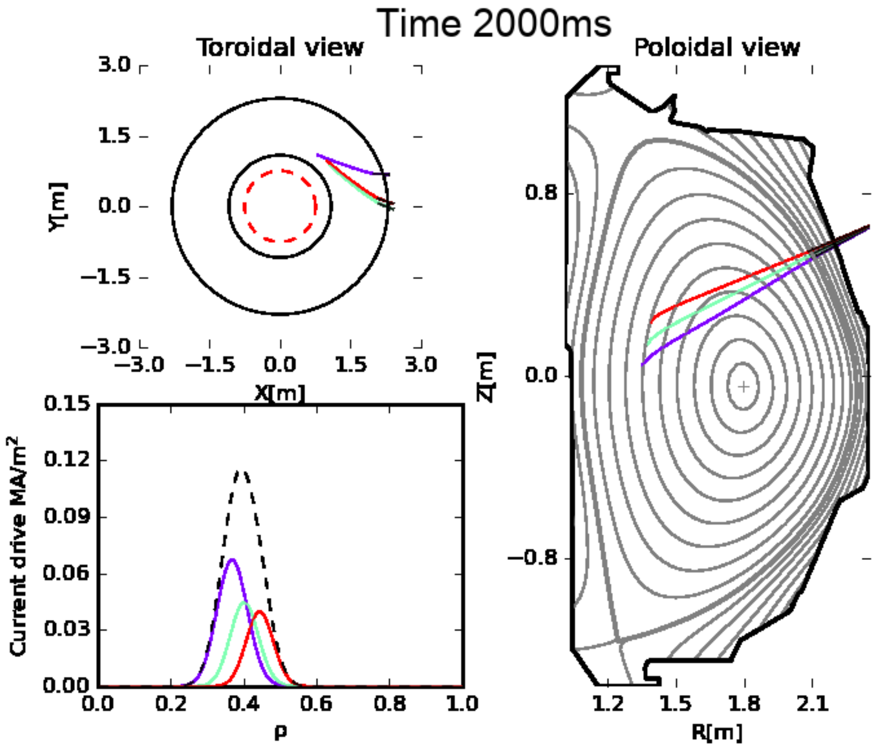
Instabilties
- Tearing Modes: 1-10ms
- VDEs: μs scale
- Disruptions: $\approx$ 1ms



Real-time control system
- Shape control: <1ms
- NBI heating: 50ms
- ECH heating: 50ms
- ML models: 1-10ms
- Magnetic diagnostics: <1ms
- Profile diagnostics: $\approx$ 20ms

Real-time TORBEAM developed and demonstrated in DIII-D

- Gyrotrons start spread out for broad, off-axis CD
- As shot progresses and becomes unstable, gyrotrons clump around q=2 location ($t = 3500ms$)
- After stabilizing, gyrotrons spread out again



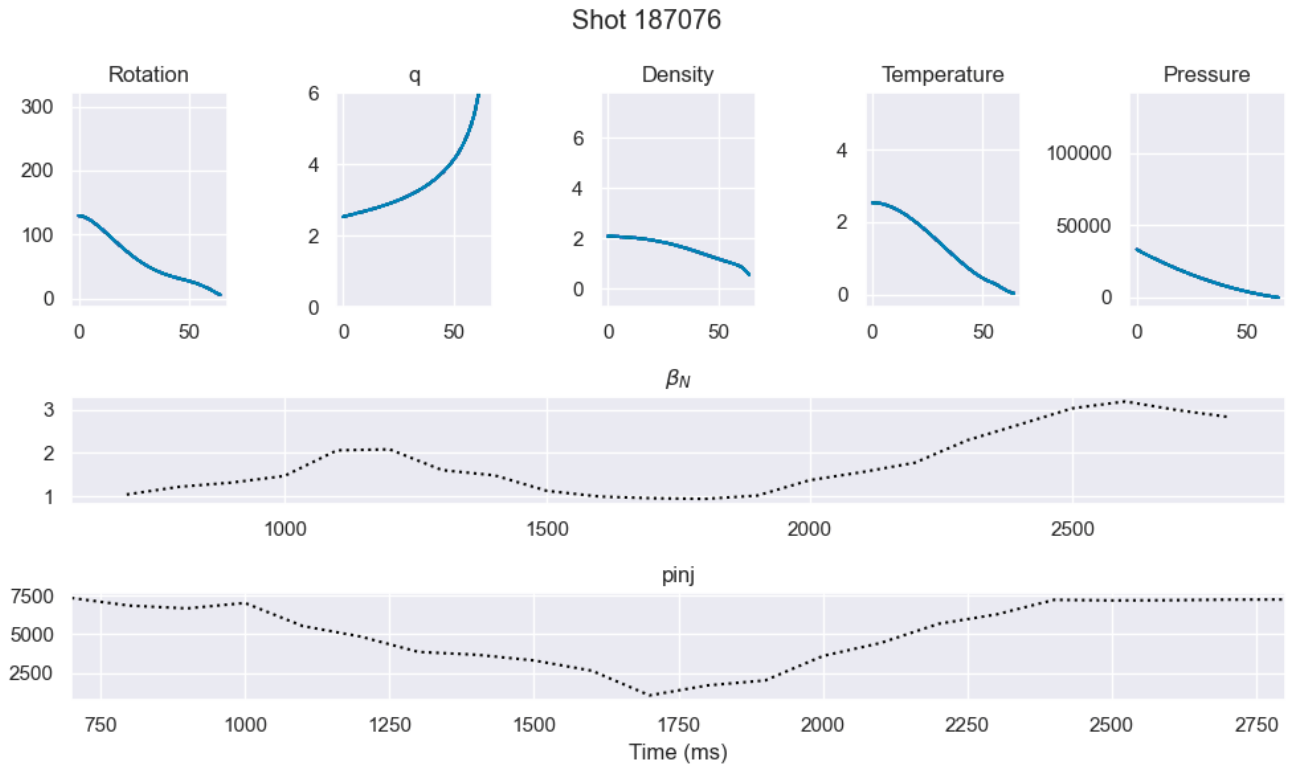
Plasma Control

Profile Prediction

Predict full shots with actuator trajectories

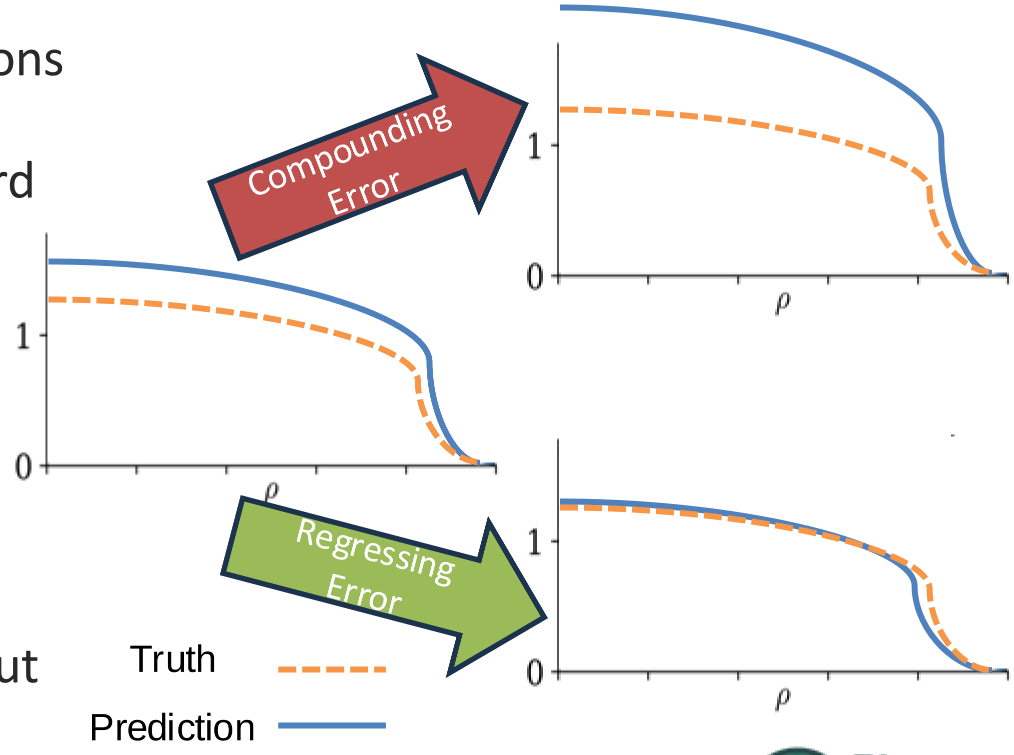
Ground Truth

Predictions



How to keep long-term predictions stable

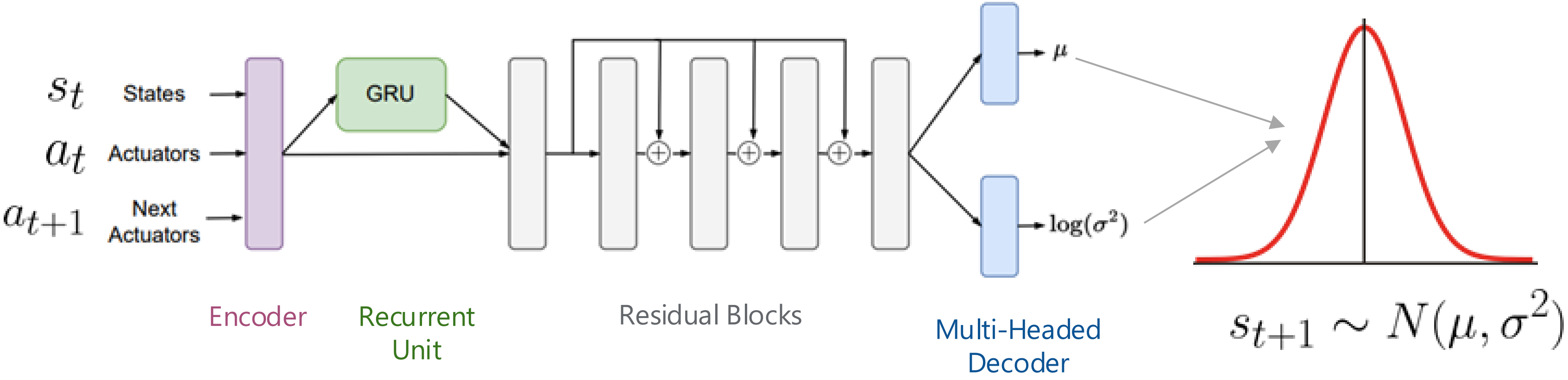
- Predicting off previous predictions causes errors to compound
- Need to have ”regression toward the mean”
- Solution 1: Uncertainty predictions
  - Predict (μ, σ)
- Solution 2: model ensembling
  - Multiple models = further averaging
- Solution 3: autoregressive rollout



Plasma Control

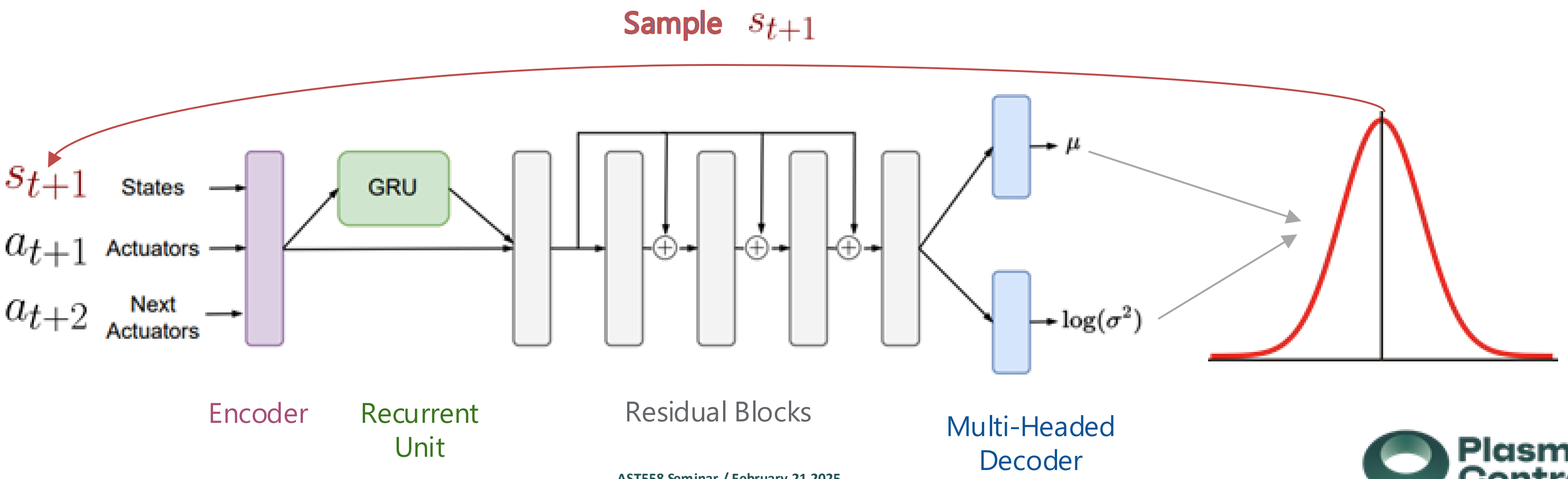
Model Architecture

- Predictions are made 25ms into future
- Model predicts a Gaussian distribution of the next state



Predicting Full Shots

- Next state sampled from Gaussian and fed back into model
- Actuators can be taken from historical data ("replaying a shot") or provided by some optimization algorithm

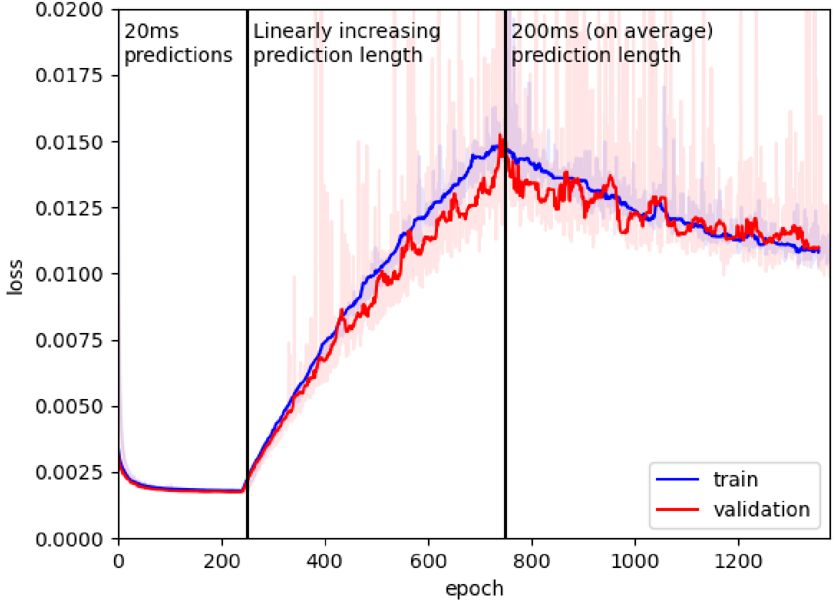


Plasma Control

Curriculum Learning stabilizes long-term predictions

- Curriculum learning (aka autoregressive rollout) is a staged approach to reach full shot profile predictions
- Start by having model predict $\mu = 1$ time steps into future
  - Use time t to predict $t + 1$
- Ramp prediction horizon from $\mu = 1$ to $\mu = 10$
- Continue training at $\mu = 10$

Training + Validation Loss



Plasma Control

MPC Control

Plasma Control

Model Predictive Control

$$x_{t+1} = Ax_t + Bu_t$$

Predicted state

Current state

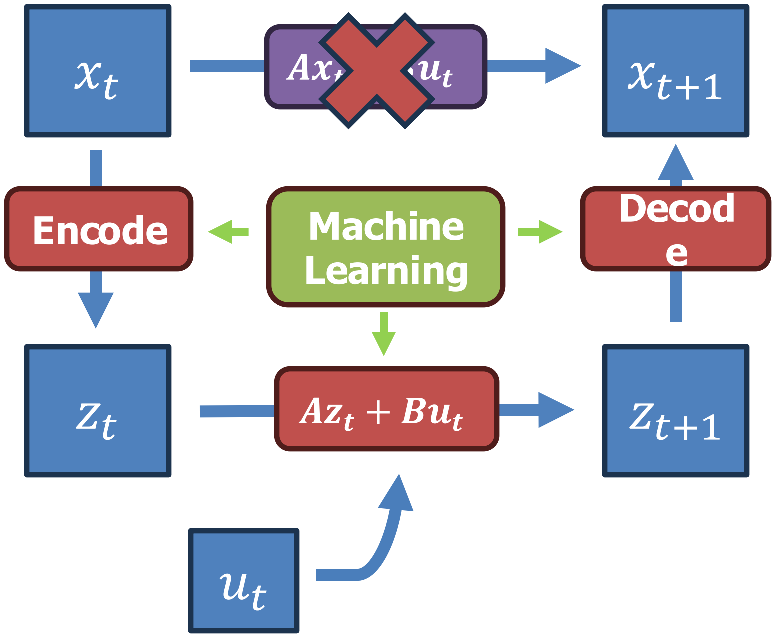
Control Actuators

$$Cost = \sum_t \overbrace{(x_{target} - x_t)^T Q (x_{target} - x_t)}^{\text{Tracking Error}} + \overbrace{u_t{}^T R u_t}^{\text{Control Effort}}$$

- MPC efficiently finds the optimal (cheapest) actuator trajectory to reach a desired state
- Requires linearized dynamics model of the plasma, but we know plasmas are strongly nonlinear!
- How can we control in real-time?

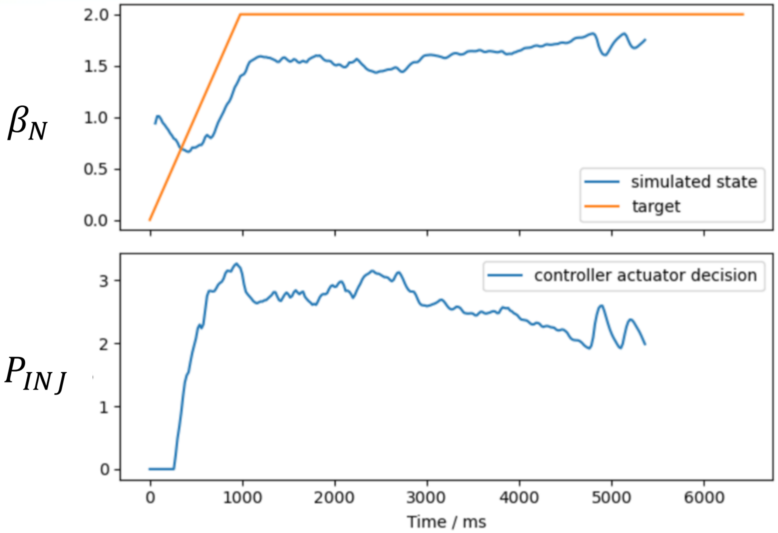
ML Linear Projection

- Nonlinear plasma behavior can be approximately mapped to a larger linear space
- The encoder, decoder, A, and B matrices are learned from DIII-D data
- MPC can be applied to this linear model to find optimal actuator trajectories



[M. Watter, 2015]

Testing out MPC Controller



Predicted result (using PP model)

β_N target

MPC Controller NBI trajectory

- Proof of concept: control β_N with NBI heating
- General behavior reasonable
  - Gets stuck with steady-state error
- Working on full profile controller given a broader set of actuators
  - NBI power and torque, ECH heating, Ip, Bt, shaping and gas injection