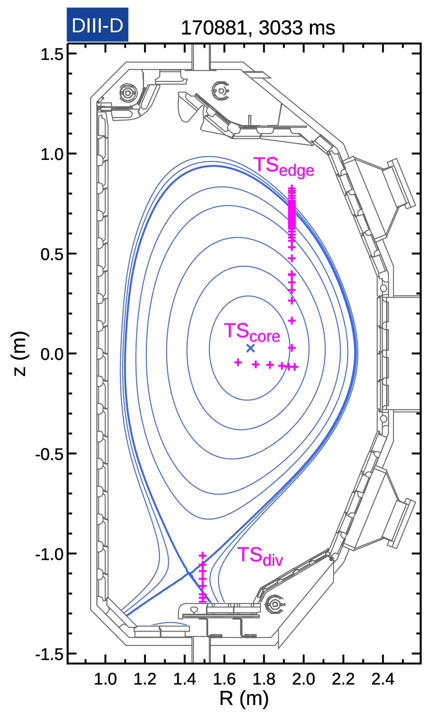
Experimental Based Pedestal Structure Prediction using Machine Learning

J. J. Zhao[1], E. Kolemen[2], X. Li[1], Y. Fu[2] and F. M. Laggner[2]

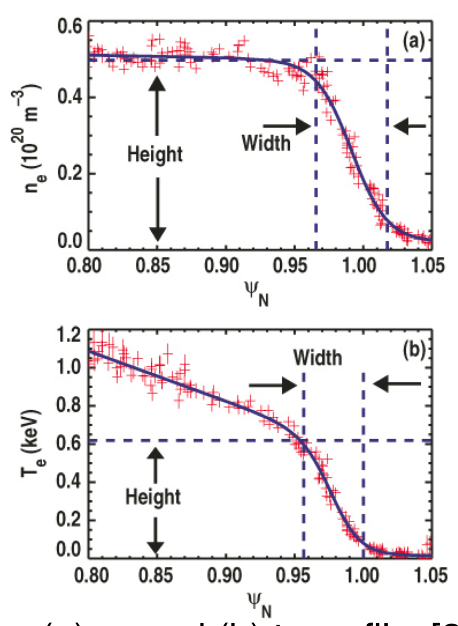
[1]Princeton University, [2]Princeton Plasma Physics Lab

1. Introduction and Motivation

- In tokamak reactors, the pedestal is the steep pressure drop at the plasma edge in high confinement mode (H-mode)
  - Over 30x increase in pressure across a 0.4-5 cm layer
- Importance of the pedestal:
  - Fusion power in the tokamak is strongly dependent on the pedestal top pressure
  - Edge localized modes (ELMs), expelling particles and heat from the confined plasma, originate from the pedestal layer
    - ELMs leads to machine wall deterioration
- Understanding and predicting pedestal behavior enables pedestal and fusion performance optimization
- Previous works use some pedestal features as inputs for other pedestal features
  - Alongside global plasma parameters such as β_N, EPED-NN[1] uses pedestal top electron density ($n_{e,ped}$) and effective charge ($Z_{eff,ped}$) for pressure and width
- **Goal of presented work:** Predict pedestal features based on operationally accessible 'machine' parameters



Cross-section of plasma



(a) n_e and (b) t_e profiles[2]

2. DIII-D Pedestal Database

- Two years of data from DIII-D tokamak in San Diego totaling **1092 shots with 43980 time slices** (TMS)
- Automated database creation and data fetching with custom OMFIT module [3]
- Fits to Thomson scattering data used to obtain pedestal parameters [2], and EFIT used to calculate machine parameters [4]
- **Included input parameters:**
  - Basic plasma: Plasma current, toroidal magnetic field, edge safety factor and normalized plasma pressure
  - Shaping: a, r, triangularity, elongation, wall clearance, separatrix distance
  - Heating: NBI heating power, beam fueling, ECH heating power
  - Applied gas puffs
- **Output parameters:**
  - $n_{e,ped}$, $n_{e,sep}$, $t_{e,ped}$, $t_{e,wid}$, $n_{e,wid}$

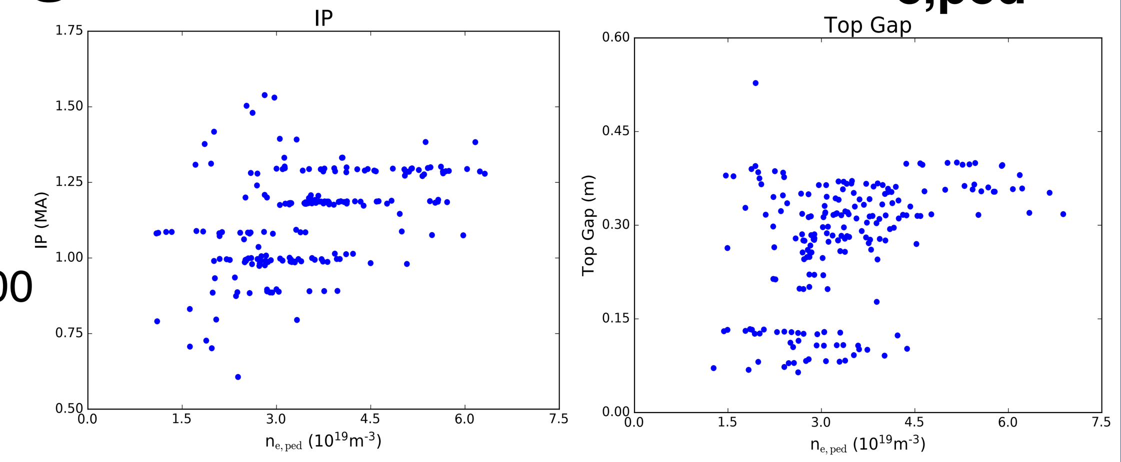
Acknowledgements

Part of the data analysis was performed using the OMFIT integrated modeling framework. This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Fusion Energy Sciences, under Awards DC-AC02-09CH11466, DE-SC0015878 and DE-SC0015480 using the DIII-D National Fusion Facility, a DOE Office of Science user facility, under Awards DE-FC02-04ER54698. DIII-D data shown in this paper can be obtained in digital format by following the links at https://fusion.gat.com/global/D3D_DMP.

Disclaimer: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

3. Preprocessing and Correlation with $n_{e,ped}$

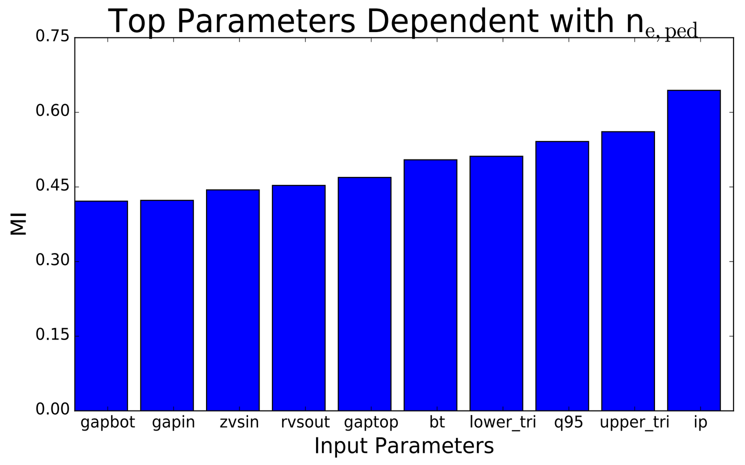
- Database Split for Neural Network Evaluation:
  - **Training dataset**: 892 shots with 35154 TMS
  - **Validation dataset**: 100 shots with 4413 TMS
  - **Testing dataset**: 100 shots with 4413 TMS



- Removed outliers and noise from parameters and TMS samples
- Parameters and outputs normalized
- **Parameter correlation with $n_{e,ped}$:** Mutual information (MI) measures dependency between parameters and $n_{e,ped}$

$$MI(x,y) = H(y) - H(y \mid x)$$

- H(x): Entropy



4. Neural Network Model

- Function at each node:

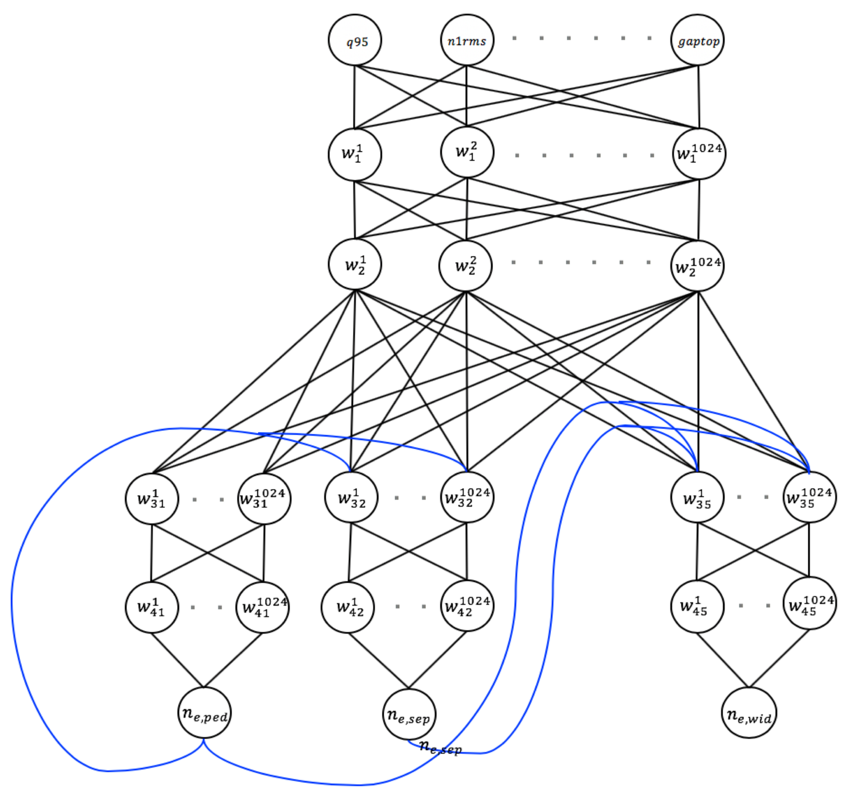
$$f(X) = ReLu(wX + b)$$

- ReLu: activation function
- w: matrix of learned weights
- X: input parameter matrix
- b: learned bias
- Node weights originally set with Xavier initialization to reduce training time [5]
- Training using Adam optimizer with exponential decay on loss function [6]:

$$L(y_i, \mathbf{y_i}) = |y_i - \mathbf{y_i}|^2$$

- y_i: vector of measured outputs of the i^{th} sample
- $\mathbf{y_i}$: vector of calculated outputs of the i^{th} sample
- Adam optimizer prevents staying in local minimums
- Exponential decay reduces training time and increases accuracy
- 15 epochs during training
  - Epoch: the number of times the algorithm sees the entire training dataset

Neural Network Architecture



- **Multitask neural network with output back-feeding:**
  - output parameters as input for other outputs
  - Inspired by strong correlations between pedestal features
- Multiple architectures and hyperparameters tested, and the final network with the best final result on the validation dataset chosen

References

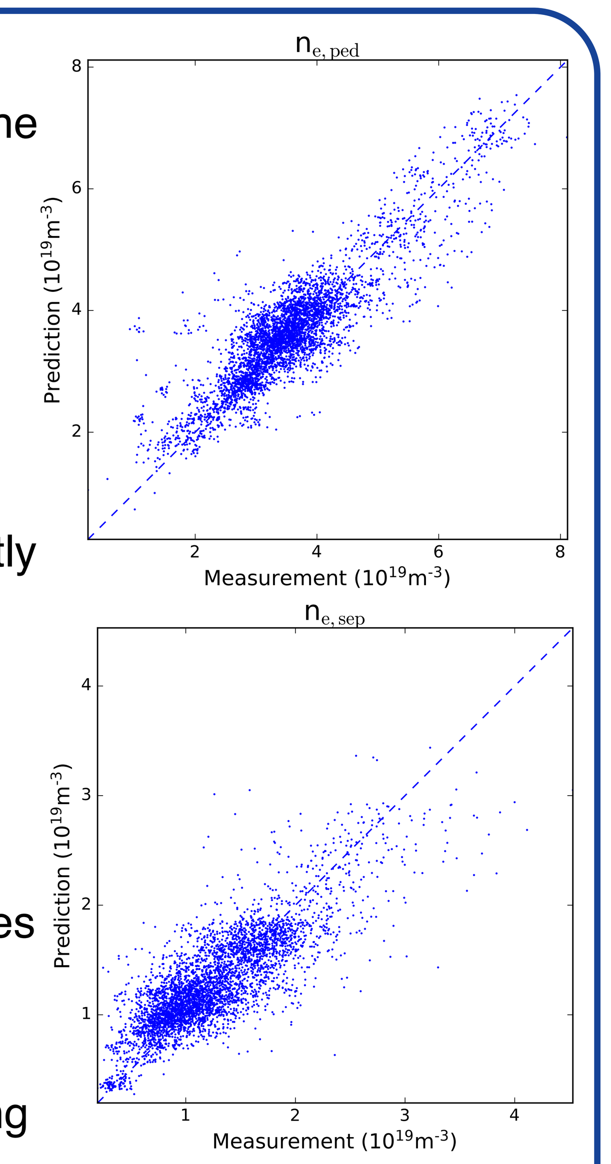
[1] O, Meneghini, et. al., Nuclear Fusion, Volume 57 (2017)
[2] R.J. Groebner et. al., Nuclear Fusion, Volume 49 (2009)
[3] B. Zywicki et. al., General Atomics SULI Poster Presentation (2017)
[4] L.L. Lao, et. al., Nuclear Fusion, Volume 30 (1990)
[5] Xavier Glorot et al. Proc. of the Thirteenth Int. Conference on AI and Stat., in PMLR 9:249-256 (2010)
[6] D. Kingma and J. Ba, Third Int. Conference for Learning Representations (2015)

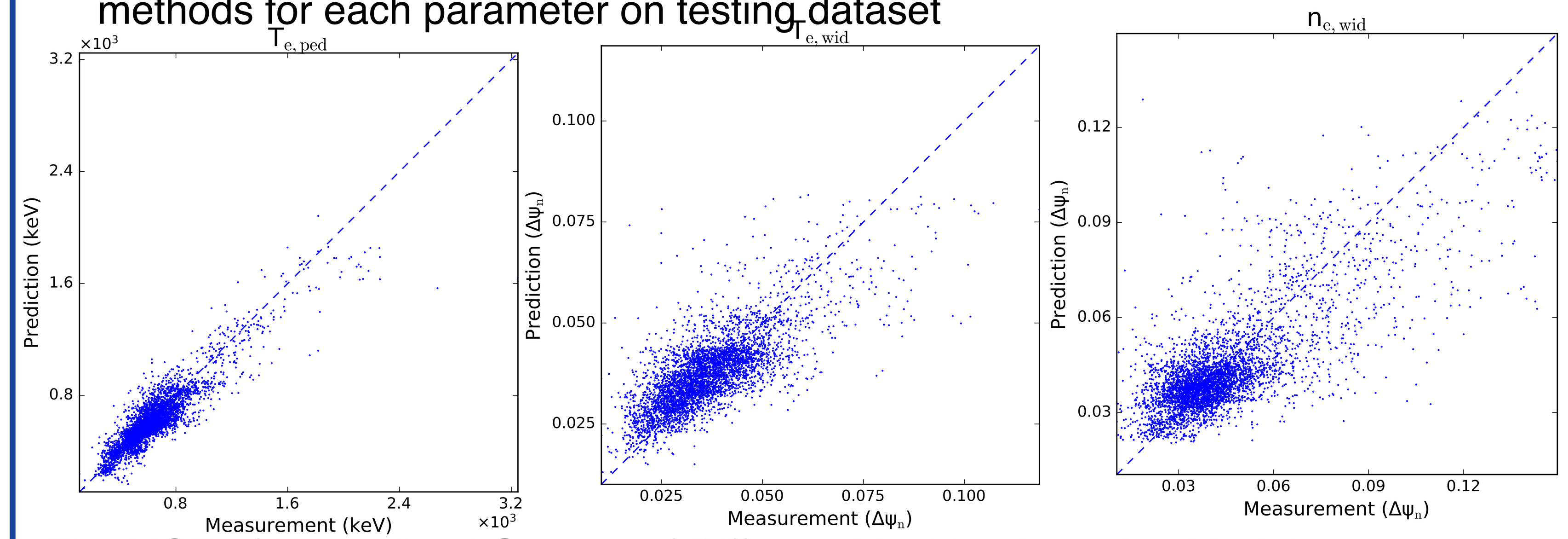
4. Prediction Results

- Graphs show predictions of the neural network on the testing dataset
- Good prediction of $n_{e,ped}$, $n_{e,sep}$ and $t_{e,ped}$
  - This suggests that pedestal features correlated with basic machine parameters
  - However, machine learning cannot distinguish correlation from causation
- $t_{e,wid}$, and $n_{e,wid}$ show correlation but have significantly worse predictions on outliers
  - Unbalanced datasets with little variation
    - Only < 10% of $n_{e,wid}$ is greater than 0.08
  - Large uncertainty in the measurement of both values
- Mean Squared Error (MSE) of j^{th} output on n samples

$$MSE(y^j, \mathbf{y}^j) = \frac{|y^j - \mathbf{y}^j|^2}{n}$$

- Neural network performs better than shallow learning methods for each parameter on testing dataset





The MSE of Normalized Outputs of Different Machine Learning Models

Model	$n_{e,ped}$	$n_{e,sep}$	$t_{e,ped}$	$t_{e,wid}$	ne_wid
Linear Regression	0.0064	0.0054	0.0039	0.0029	0.0119
Random Forest	0.0041	0.0043	0.0027	0.0026	0.0108
AdaBoost	0.0073	0.0073	0.0040	0.0048	0.0174
Neural Network	**0.0034**	**0.0042**	**0.0016**	**0.0023**	**0.0095**

5. Summary

- New neural network model based on multitask architecture shows significant accuracy on pedestal features despite not using pedestal information
  - Better performance compared to shallow machine learning models
- Demonstrated relationship between external parameters and $n_{e,ped}$

6. Outlook

web version

- Predictions of the pedestal on future tokamak experiments
- Extend the database to more machines
- Which parameters correlate with the pedestal (especially $n_{e,ped}$) and how are they correlated?
  - Further analysis using neural networks

corresponding author: jinjinz@princeton.edu

60th Annual Meeting of the APS Division of Plasma Physics, Portland, OR, November 8, 2018, JP11.00054