Real-time machine learning modeling of pressure and density profiles on NSTX

J. Chadwick[1], M. D. Boyer[2]

1 Carnegie Mellon University

2 PPPL

Overview

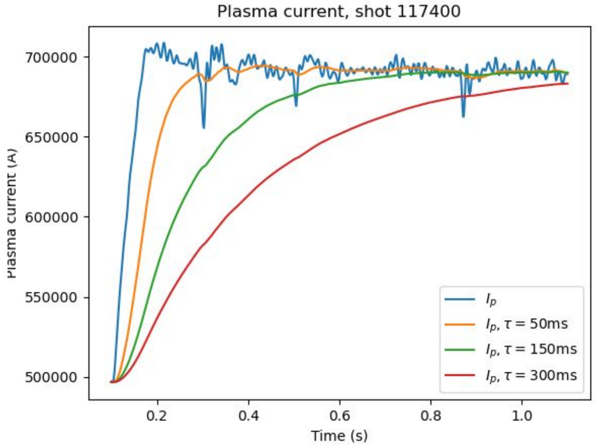
- Control systems for fusion reactors would greatly benefit from real-time density and pressure profile data during operation
  - These profiles cannot be measured directly
  - Existing code (TRANSP) can produce these profiles but is far too slow to do in real time
- A much faster alternative: machine learning! Neural networks can produce these profiles in real time with high accuracy
  - Approach has previously been applied to NSTX-U by M. Boyer, but NSTX has a much larger dataset with a wider range of experimental conditions
- Goals: determine optimal model parameters, analyze performance, develop measures of prediction uncertainty

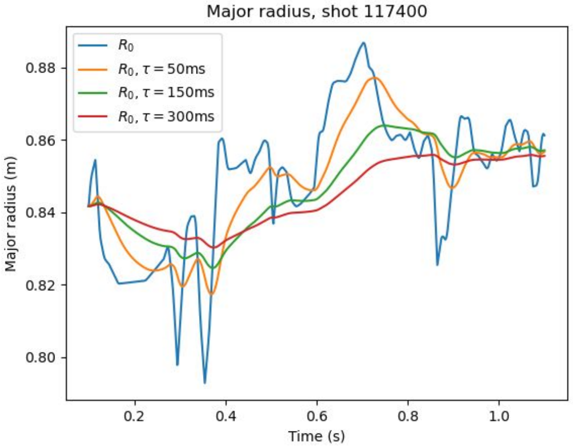
Time history of scalars

- Low-pass filter with a selection of time constants applied to input scalars
    - Instantaneous and filtered values are then all passed as inputs to net
- Gives the network a simple measure of the time history of each scalar
- Filter described by

$$x_{n+1} = \frac{u_n - x_n}{\tau}(t_{n+1} - t_n)$$

-
    - *x* are filtered values, *u* are instantaneous scalar values, τ is the filter time constant





Inputs, outputs, and model topology

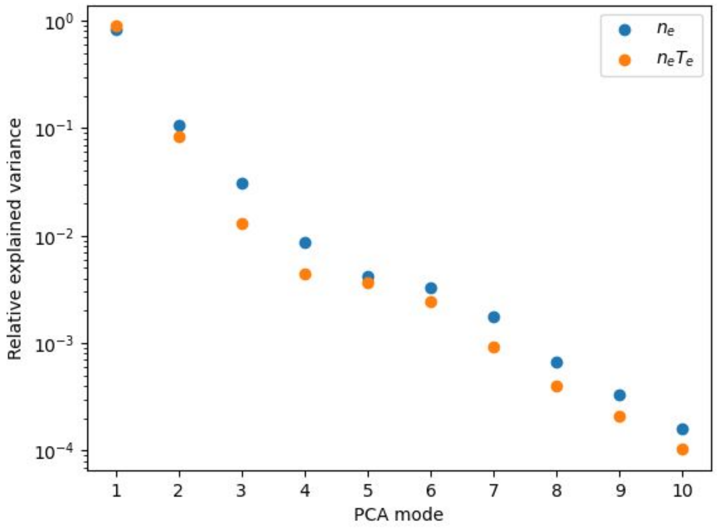
Inputs		Outputs	
Symbol	Name	Symbol	Name
R_0	Major radius	n_e	Electron density profile
a	Minor radius	$n_e T_e$	Electron pressure profile
$B_{\phi,v}R$	Vacuum toroidal field		
κ	Elongation		
I_p	Plasma current		
δ_u	Upper triangularity		
δ_l	Lower triangularity		
$\langle n_e \rangle$	Volume average electron density		
$\langle n_e T_e \rangle$	Volume average electron pressure		

- 7 scalar inputs are each low-pass filtered 3 times
  - Total of 30 scalar inputs to the model
- 4 fully connected layers of 100 nodes each
- 3-model ensemble
  - 3 models trained on overlapping subsets of training data
  - Final prediction = average prediction of the 3 models
  - Standard deviation can be used as a measure of uncertainty

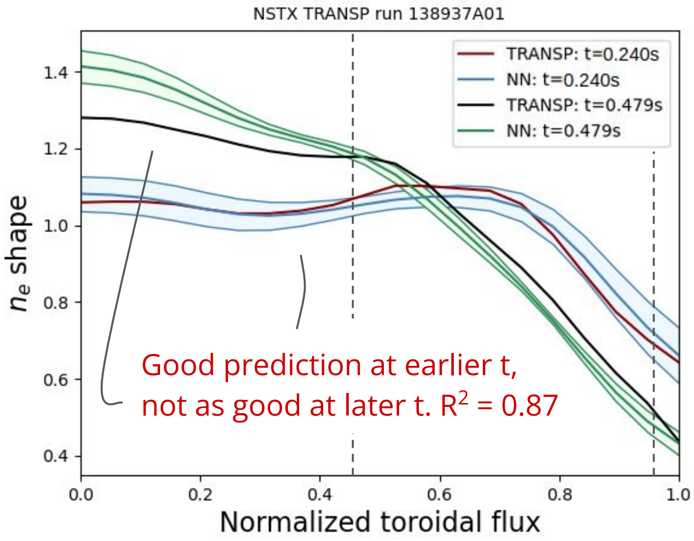
Dataset and preprocessing

- Dataset: NSTX TRANSP A01 runs from 2004-2011
- 1837 shots in total
- Total of ~995,000 time slices
- 49 data points per slice
  - 9 scalars, measured in real time
  - 2 profiles of 20 radial points each, calculated by TRANSP after the shot is over
- To reduce dimensionality, used principal component analysis (PCA) to project profile data onto a reduced number of modes (6) before training

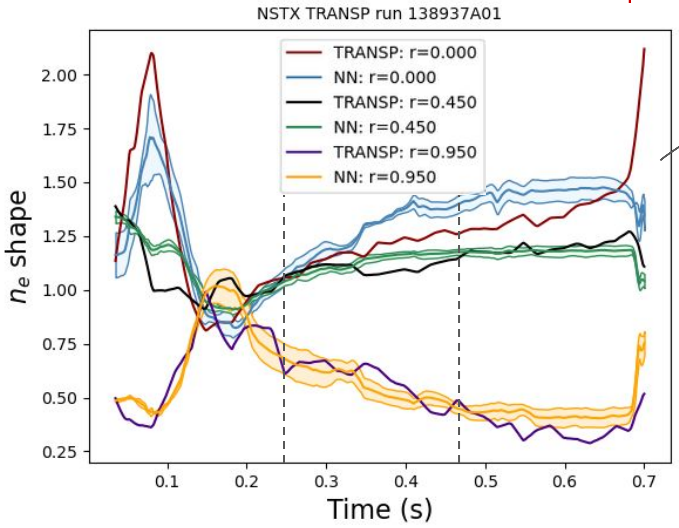
Keep first 6 modes
(explains ~99.7% of variance in data)



Profile example



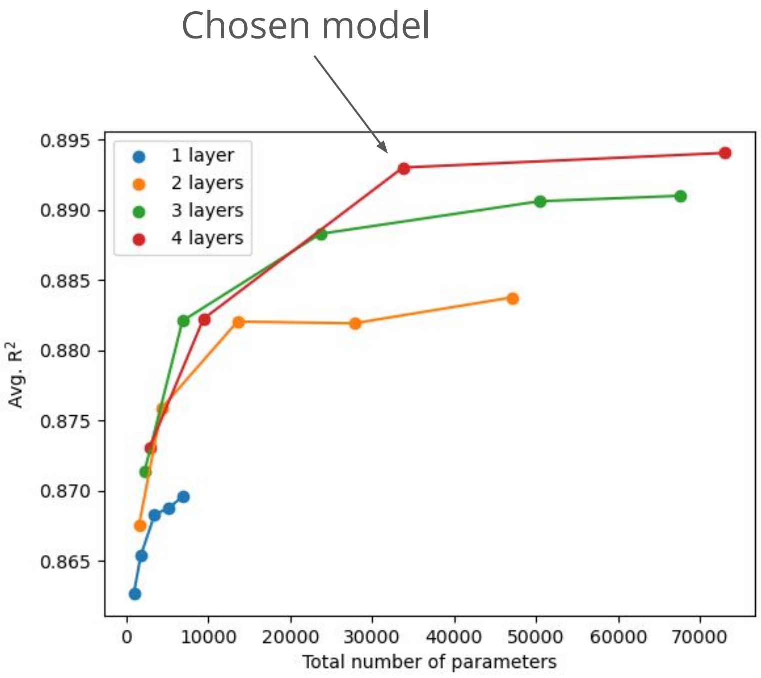
Bad predictions
at disruption



- Left: profile at two different times in shot. Right: Three fixed radial points tracked throughout the whole shot
- Shaded area is one ensemble standard deviation

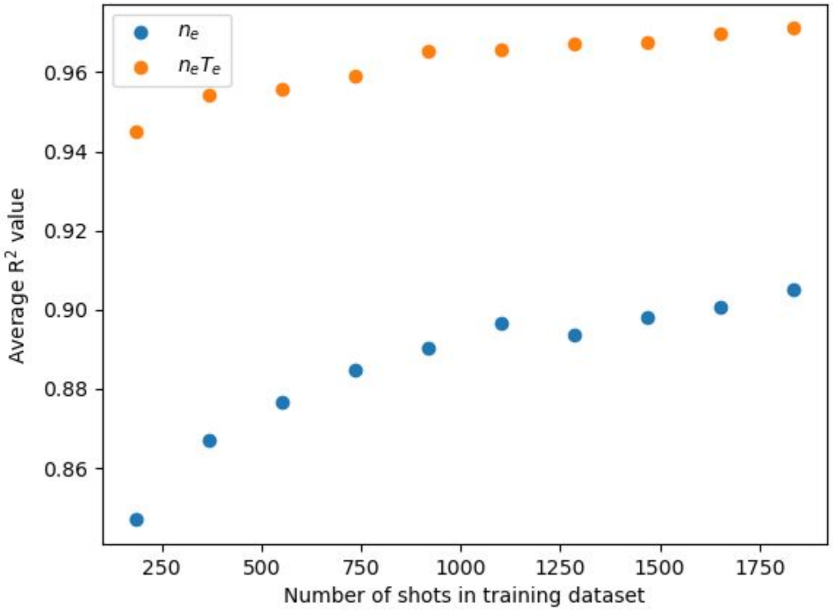
Network architecture choice

- To determine ideal architecture, train many with different parameters and test on the same dataset
    - Each model had *n* layers of *m* nodes each
- Diminishing returns after ~30,000 parameters
    - Parameter = total num. of nodes and weights
    - Too many parameters → long runtime
- Selected architecture consisting of 4 layers of 100 nodes each (35,854 parameters) to balance accuracy and complexity

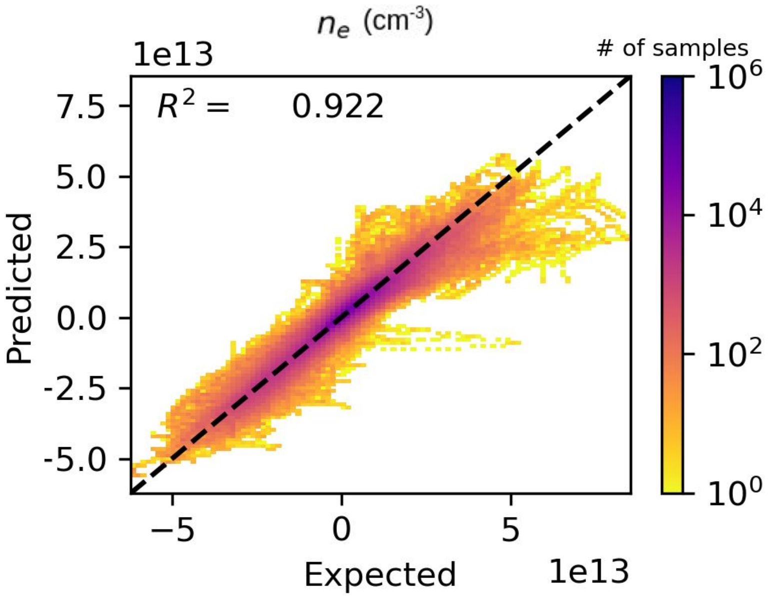


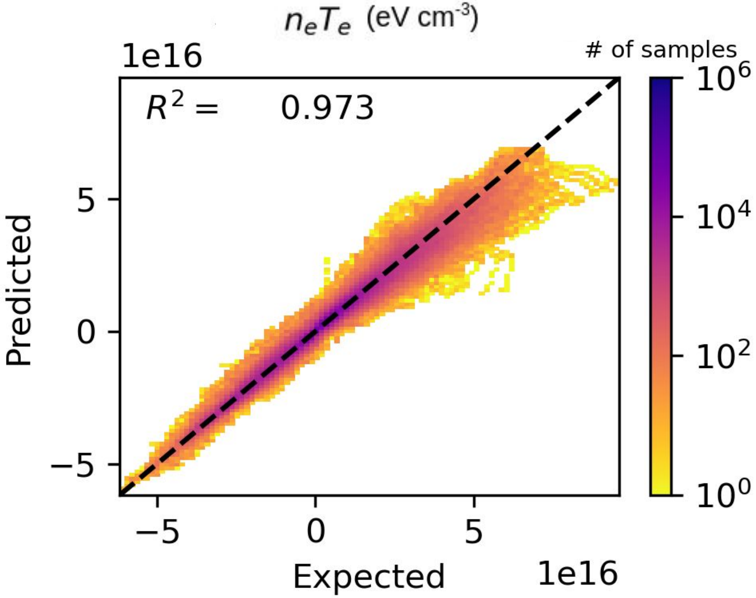
Training set size

- Trained on increasing subsets of training data and tested on full test set
- Results: Continually improves with more data, as expected
  - More data is always good, but can get reasonable results without all of it
- Pressure R^2 values are regularly higher than density



Results: regression plots

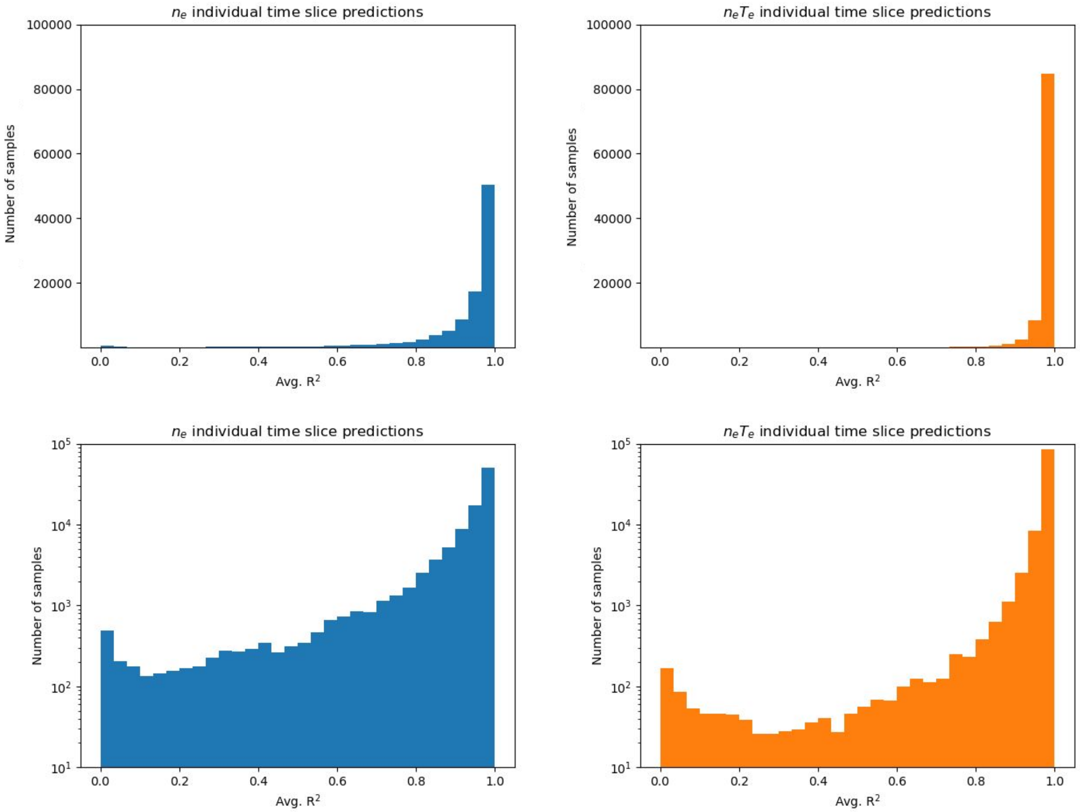




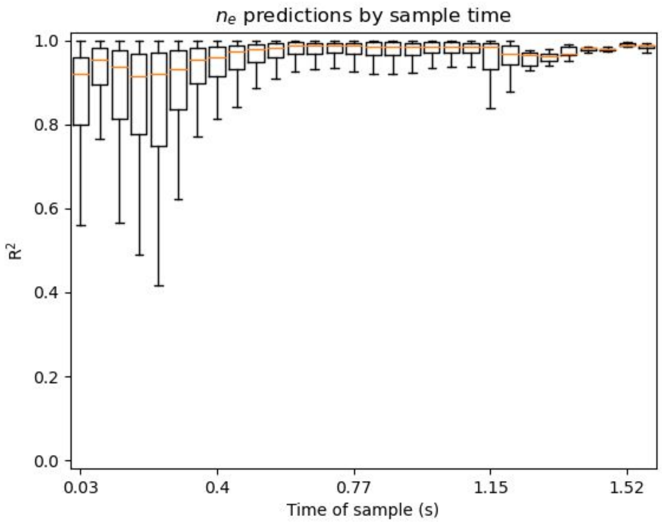
- Comparison between actual and predicted values for each radial measurement of each profile in testing set
  - Good predictions overall, with a small number of outliers

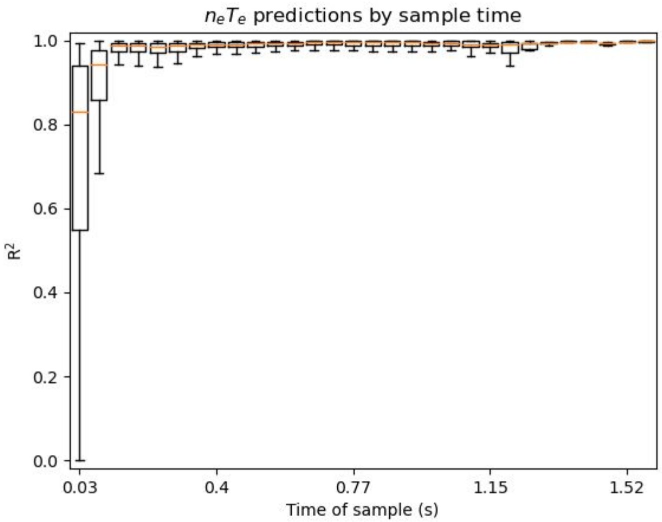
Results: R^2 values by time slice

- Plots: R^2 for each slice in testing set, on linear and logarithmic scales
- Vast majority of time slices are well predicted
- Can detect some poor predictions with measures of uncertainty
- Need to investigate source of very low ($R^2 < 0.2$) predictions

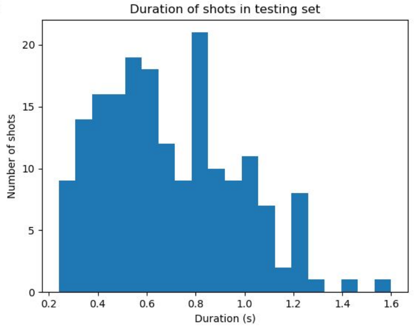


Results: predictions by sample time



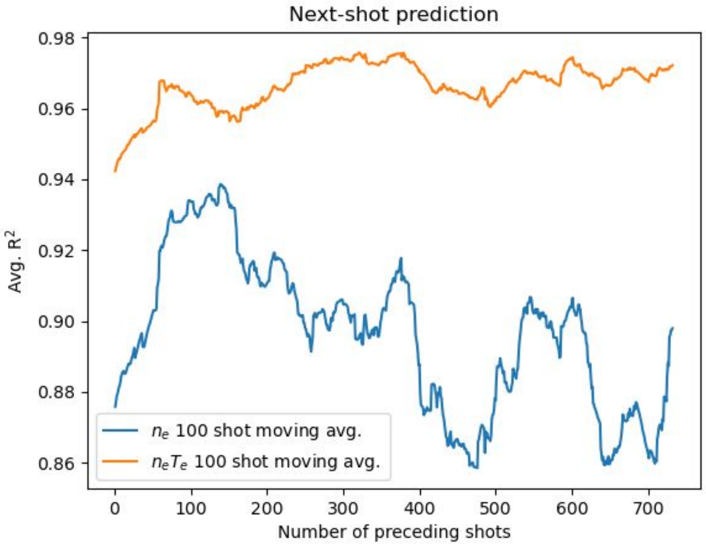


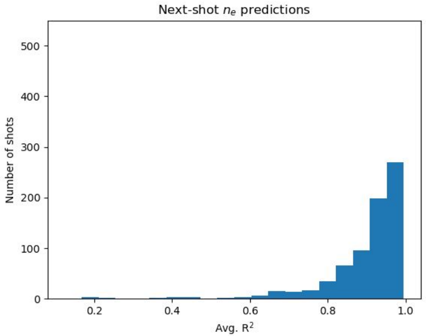
- Early times seem to be the hardest to predict
    - note: not all shots ended at the same time
- Need to investigate early predictions of pressure

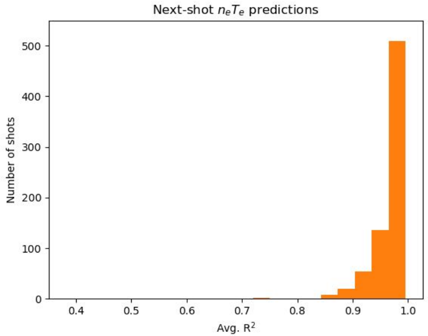


Next-shot predictions outside of training space

- More realistic test: predict each shot based only on previous data
    - Weight most recent 250 shots 3x to account for long-term changes in physical design of reactor or in experimental goals of shots
- Results: Good average predictions throughout dataset
    - Accuracy improves quickly at the start
    - Density predictions less reliable than pressure

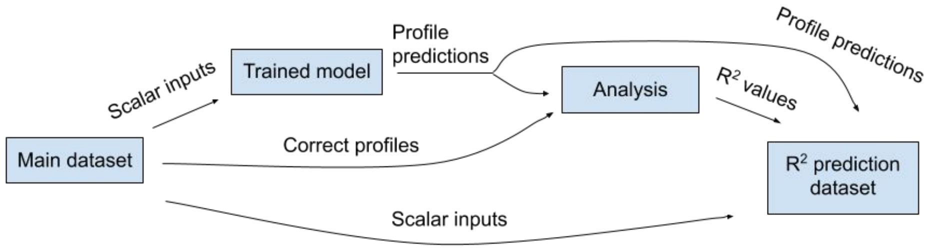






Predicting inaccuracies

- Goal: predict R^2 value based on inputs and profile predictions
  - Purpose: detect poor predictions before they would be used in a control system
- Used a second neural network, trained on the R^2 values of the main net
  - Inputs: original input scalars (9 scalars), predicted profile PCA components
  - Outputs: model prediction R^2 values (for density and pressure)
  - Training data: 90% of test set results (~85k time slices)
  - Test data: remaining 10% (~9.5k time slices)
- Looked at how often the model will correctly predict that an R^2 value is lower than a threshold (e.g. $R^2 < 0.80$)



Predicting inaccuracies: results

Threshold: $R^2 \leq 0.80$

R^2 category	Total expected below threshold	Total predicted below threshold	Num. overlapping	Num. false positives	**False positive rate**	Num. false negatives	**False negative rate**
n_e	1690	2478	1495	983	**0.11**	195	**0.12**
n_eT_e	309	484	279	205	**0.02**	30	**0.10**

- Results: good, but can be improved
  - False positive and false negative rates are relatively good but not perfect
- Combining this technique with ensemble uncertainty (and possibly other measures of uncertainty in the future) has potential to greatly increase reliability of main neural network

Conclusions

- A neural network is capable of reliably reproducing TRANSP profile predictions for most shots in the dataset with high accuracy
    - Promising for control system applications
- Approach was effective on both NSTX and NSTX-U
    - Promising for use with other reactors as well
- Model predicts electron pressure well, but we still need to improve density predictions
- Most poor predictions are at early times in each shot
- Model is capable of predicting future shots that are not in the training space
- We have reasonable measures of model confidence

Future work

- Develop improved estimation of volume-averaged electron density and pressure
  - Possibly a secondary neural network
- Attempt to further improve density prediction
  - Determine effect of different filtering τ values on beginning-of-shot prediction
  - Potentially need new measured quantities - could help guide future reactor design
- Improve measures of model confidence
  - Improve prediction of R^2 values
  - Try Monte Carlo dropout for uncertainty
  - Find ways to rigorously define the training set parameter space, so we can know when we are outside of it
- Test technique on other machines (DIII-D etc.)

Acknowledgements

This work was made possible by funding from the Department of Energy for the Summer Undergraduate Laboratory Internship (SULI) program. This work is supported by the US DOE Contract No. DE-AC02-09CH11466.