MULTI-DIAGNOSTIC CLASSIFICATION OF ALFVÉN EIGENMODES USING MULTIMODAL MACHINE LEARNING

A. ROTHSTEIN, A. JALALVAND, A.V. GARCIA, M.E. AUSTIN, W.W. HEIDBRINK, and E. KOLEMEN

DIII-D NATIONAL FUSION FACILITY

PPPL PRINCETON PLASMA PHYSICS LABORATORY

Introduction

Alfvén Eigenmodes (AEs)

- AEs are driven by energetic particles that drive shear Alfven waves in magnetically confined plasmas
- AE modes are important as they will eject high energy particles out of the plasma, reducing confinement and possibly damaging chamber walls

How We Find AEs

- Since AEs cause magnetic field fluctuations, we can find them in the spectra of many diagnostics such as:
  - Electron Cyclotron Emission (ECE)
  - CO_2 Interferometry
  - $\dot{B}_P$ sensors (MPI)
  - Beam Emission Spectroscopy (BES)
- After a shot, we can analyze spectrograms of the signals to see where AEs were present (see Fig. 1)

Goals of Study

- Real-time detection of AEs cannot be achieved by regular controllers
- Answer the questions:
  - What diagnostics are best for real-time AE classification?
  - How robust will these detectors be to noise and detector dropout?

Previous Works and their Drawbacks

- ECE based Reservoir Computing Network classifier [1]
  - Uses raw ECE data and can run in real-time
  - Limited to single diagnostic
- ECE based classifier using spectrograms [2]
  - Achieves better performance and finds locations of AEs
  - Processing of spectrograms is not feasible in real-time and model is limited to only ECE

Labeled AE Mode Dataset

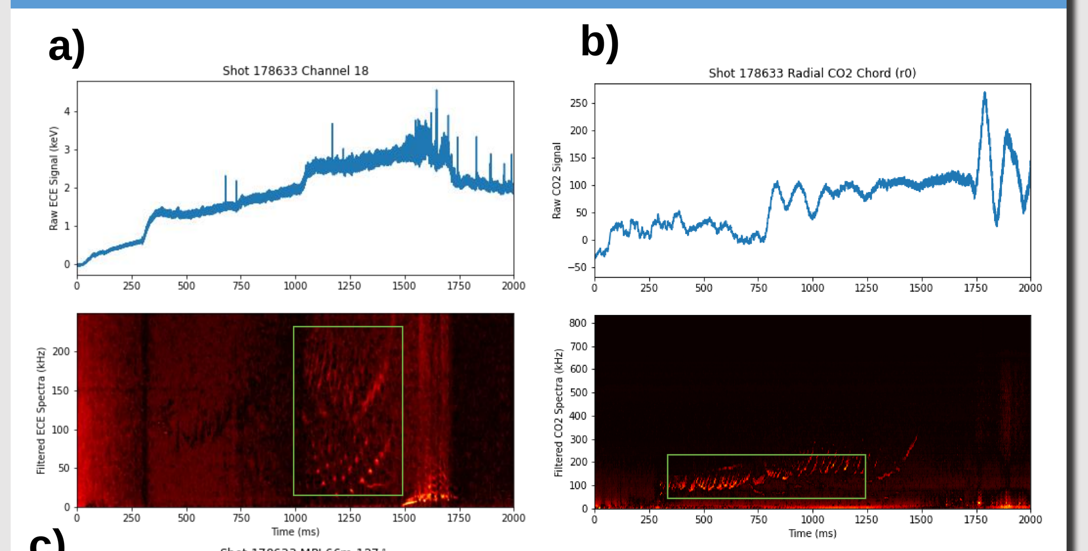
AE Labels

- Labeled dataset of ~1000 shots from Heidbrink
- Labels for:
  - Beta-induced Alfvén acoustic eigenmode (BAAE)
  - Beta-induced Alfvén eigenmode (BAE)
  - Ellipticity-induced Alfvén eigenmode(EAE)
  - Reversed shear Alfvén eigenmode (RSAE)
  - Toroidal Alfvén eigenmode (TAE)

Drawbacks of Dataset

- ~90% of the data is non-AE activity and is background noise
- Labels are inaccurate and have only rough time-label (single time value)
  - AEs are assumed to be present for $\pm 250ms$ for training data
  - Looser label of $\pm 500ms$ used to testing
- BES not available for most shots, so it is omitted

Diagnostics



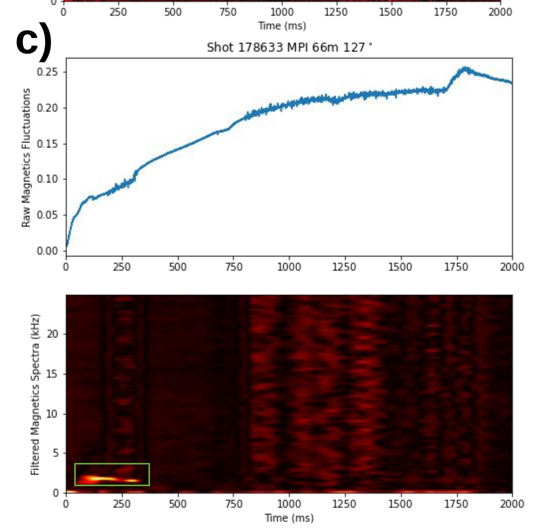


Fig. 1 Spectrograms showing (a) ECE, (b) CO_2, and (c) MPI raw signals (upper plot) and filtered spectra (lower plot). A combination of RSAE and TAE activity can be seen in the green boxes.

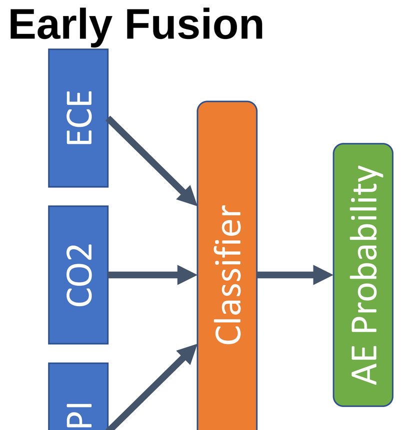
Diagnostic	ECE	CO2	MPI
Sample Rate	500kHz	1.67MHz	50kHz
# of Channels	40	4 Chords	8 Toroidal Angles
Global or Local?	Local	Global	Global

Signal Noise

- ECE has the problem of cutoff – at certain resonant frequencies, no ECE can be seen
- CO_2 can have linear drifts throughout a shot
  - These drifts can have a large range of slopes

Model Architecture

a) Early Fusion

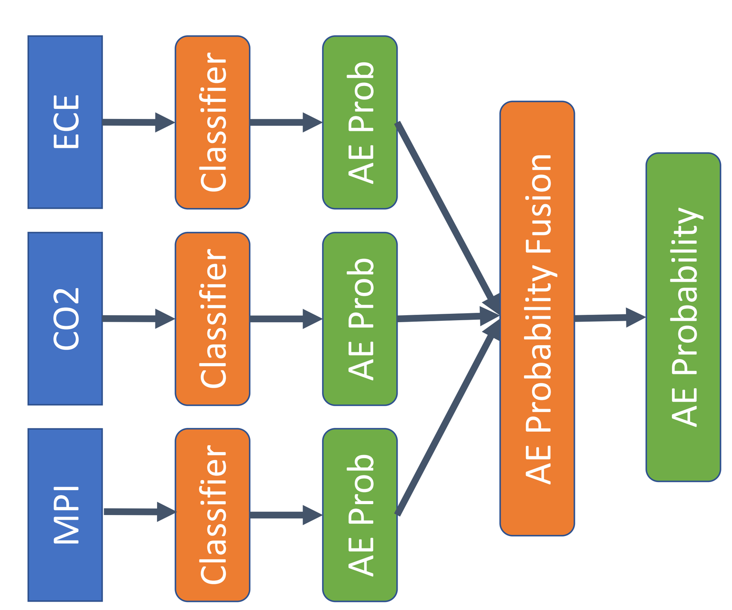


(a) The first model explored used early fusion to combine the diagnostics before learning to classify the AEs.

(b) The second tried to classify AEs based on each diagnostic, then the resulting probabilities were combined by another model to give a final probability of an AE.

Classifier layers were made up of 3-layer dense MLP network with nodes equal to approximately 2-4x input dimension. (ie for the 40 channel ECE classifier had dense layers of 128 nodes)

b) Late Fusion



Model Predictions

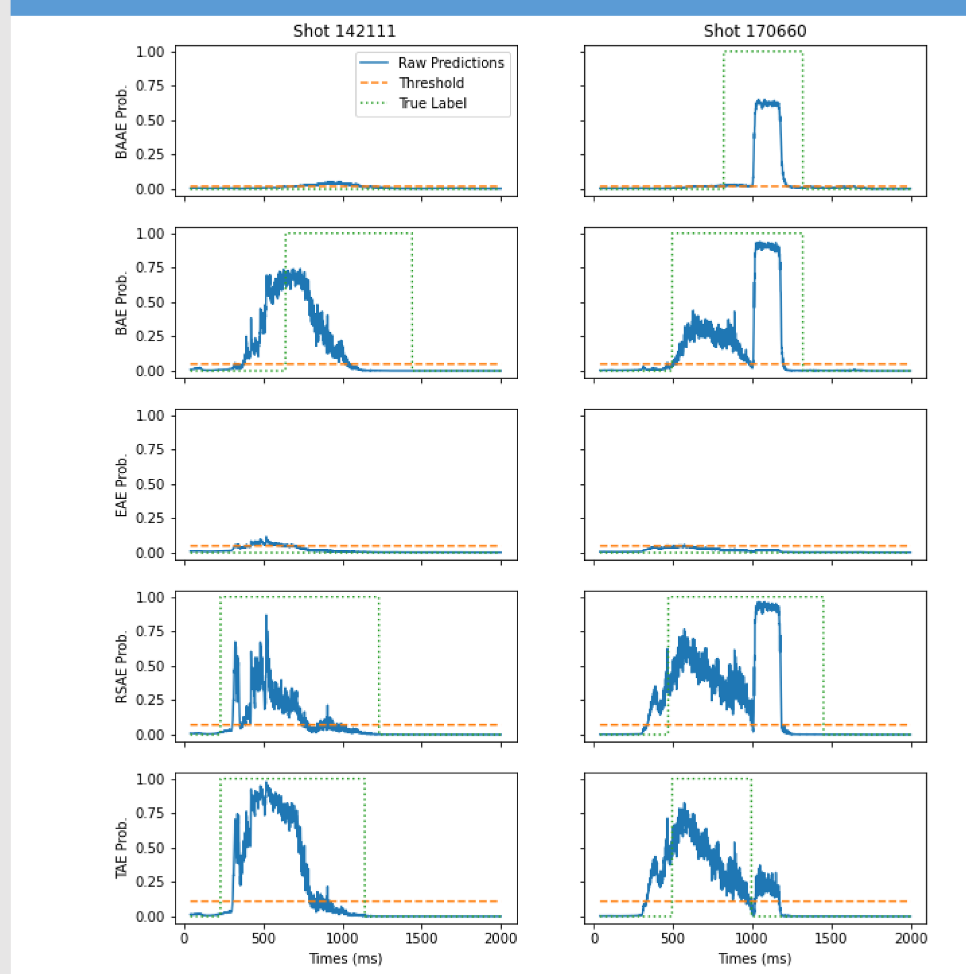


Fig. 2 Model predictions for testing shots 142111 and 170660 in the left and right columns, respectively. Vertically, we have our 5 AEs. In each plot, we have the true label (green dashed line), model prediction (solid blue line), and the threshold value (orange dashed line). When the prediction is above the threshold, our model is saying an AE is present.

Evaluating Models

Model	BAE	RSAE	TAE
Early Fusion	63.2%	63.5%	65.7%
ECE Only	41.7%	53.5%	51.1%
CO2 Only	23.1%	40.7%	51.4%
MPI Only	14.8%	40.7%	58.1%
Late Fusion	38.9%	57.1%	61.2%
ECE + CO2 TF	61.5%	62.5%	63.9%
ECE + MPI TF	57.7%	64.1%	61.8%
CO2 + MPI TF	28.2%	54.7%	65.6%

Table: Values in table are the TPRs with a threshold decided by an $FPR = 10\%$ in the validation data.

$$TPR = \frac{TP}{TP + FN}$$

$$FPR = \frac{FP}{FP + TN}$$

TPR – True Positive Rate
FPR – False Positive Rate

TP – True Positives
FP – False Positives
TN – True Negatives
FN – False Negatives

Since our dataset is skewed towards predominantly no AEs, we would find our model would perform "better" by always giving 0% AE probability. We use the TPR and FPR metrics to account for this uneven distribution of training data.

TF stand for Tensor Fusion – an input preprocessing that takes the outer product of the two signals to get the full cross-correlation matrix.

Sensitivity Analysis

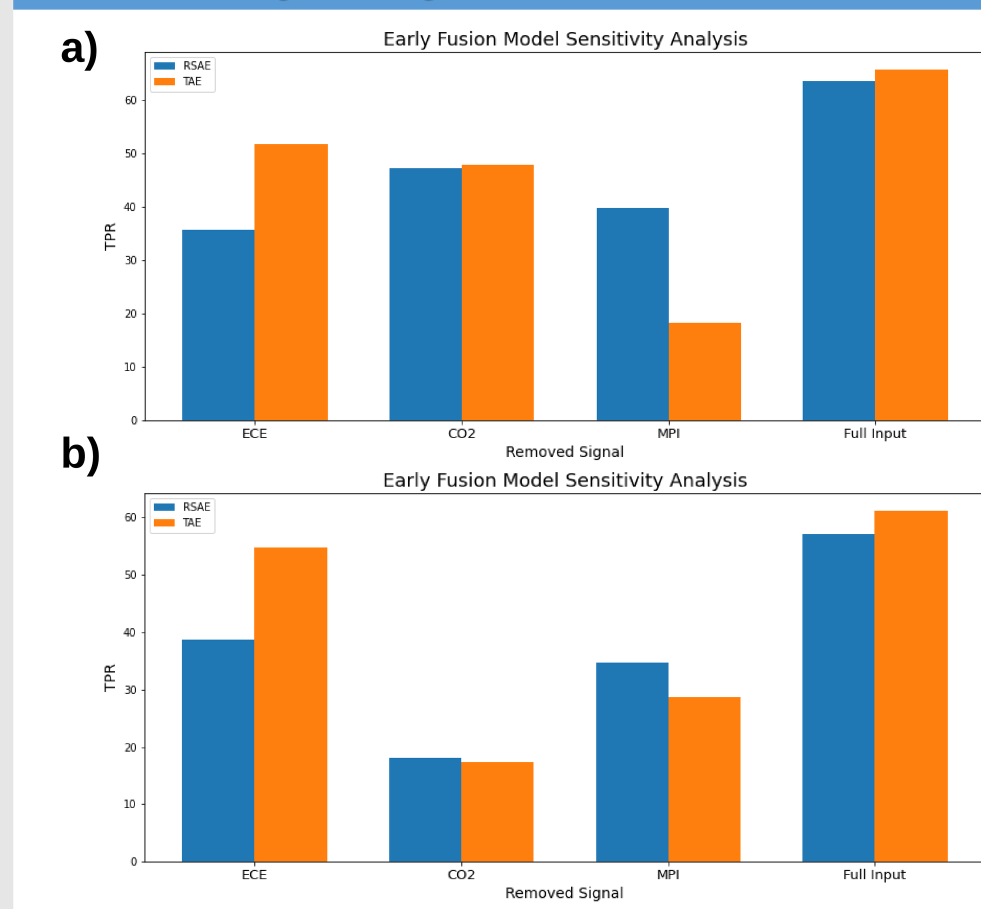


Fig. 3 We compare the model's TPR as diagnostics dropout – all channels set to 0. The x-axis tells us which diagnostic has had all inputs set to 0. The threshold is adjusted so that each has a $FPR = 10\%$ on the validation data. (a) The early fusion model. We have the TPRs for the full input on the right and we see how it changes when we set ECE, CO_2, or MPI to 0. (b) Late fusion model with the same input modifications. .

Discussion

Conclusions

- Magnetics (MPI) are extremely important
- ECE provides strong benefit while CO_2 is much less helpful in early fusion models
- Late fusion provides classification equal to early fusion, although more sensitive to signals as a whole

Future Possibilities

- Turn detector model into predictor. With a predictor, we can train a reinforcement-learning based controller to adapt and avoid AEs during shots
- Explore spectrogram-based models that could use FPGAs to be run in real-time

References

[1] Jalalvand, Azarakhsh, et al. "Alfvén Eigenmode Classification Based on ECE Diagnostics at DIII-D Using Deep Recurrent Neural Networks." *Nuclear Fusion*, vol. 62, no. 2, Feb. 2022, p. 026007. *DOI.org (Crossref)*, https://doi.org/10.1088/1741-4326/ac3be7.

[2] Kaptanoglu, Alan A., et al. "Exploring Data-Driven Models for Spatiotemporally Local Classification of Alfvén Eigenmodes." *Nuclear Fusion*, vol. 62, no. 10, Oct. 2022, p. 106014. *DOI.org (Crossref)*, https://doi.org/10.1088/1741-4326/ac8a03.

Acknowledgments

Part of data analysis is 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, DOE ARPA-E, using the DIII-D National Fusion Facility, a DOE Office of Science user facility, under Awards DC-AC02-09CH11466, DE-SC0015480, DE-SC0015878, DE-FC02-04ER54698, DEAR0001166 and Field Work Proposal No. 1903.

Princeton Plasma Control
control.princeton.edu