SAND2023-02518C

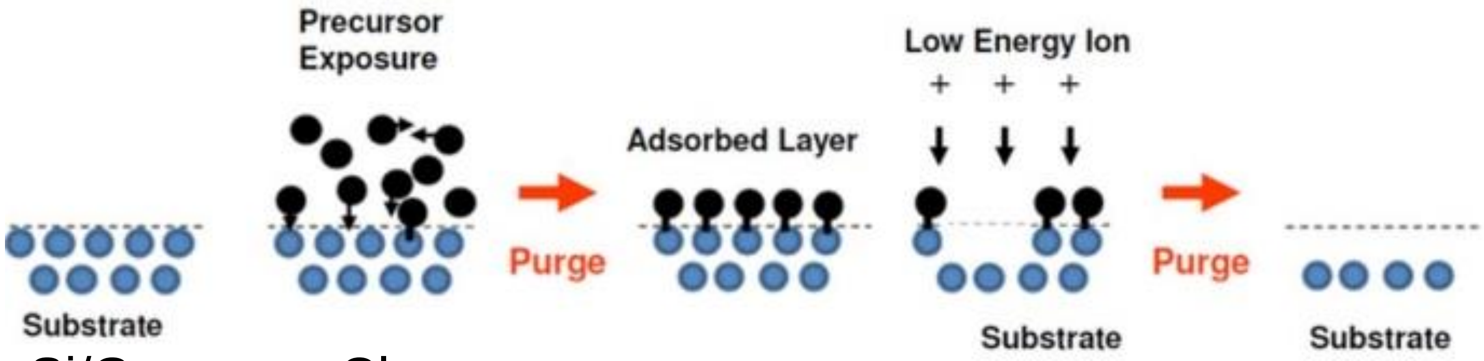
Large RF plasma data set for training plasma processing machine learning models

A. Fierro[1], T. Hardin[2], A. Lietz[3], M. Hopkins[2], A. Belianinov[2], B. Bentz[2]

[1]University of New Mexico, Albuquerque, New Mexico
[2]Sandia National Laboratories, Albuquerque, New Mexico
[3]North Carolina State University, Raleigh, North Carolina

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Atomic layer etching (ALE)



Si/Ge Cl_2 $SiCl_x/GeCl_x$ Argon Plasma

- The process beings with a semiconductor substrate that is exposed to a precursor
- Adsorbates form on the surface in a single layer
- An RF plasma produces low energy ions that remove the precursor and substrate
- True self-limiting etch of a single atomic layer across many process cycles requires better control of ions

Aleph Model

1D, 2D, 3D

Particle-in-Cell direct simulation Monte Carlo (PIC-DSMC)

- Charged species and neutrals are treated as computational particles.
- Computational particles represent varying number of physical particles according to their weight.

Electrostatic.

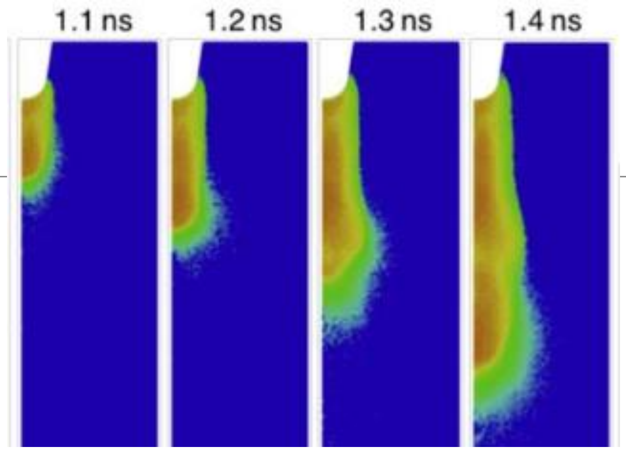
Unstructured meshes.

Merging algorithms limit the number of computational particles with minimal disturbance of velocity distributions.

Massively parallel.

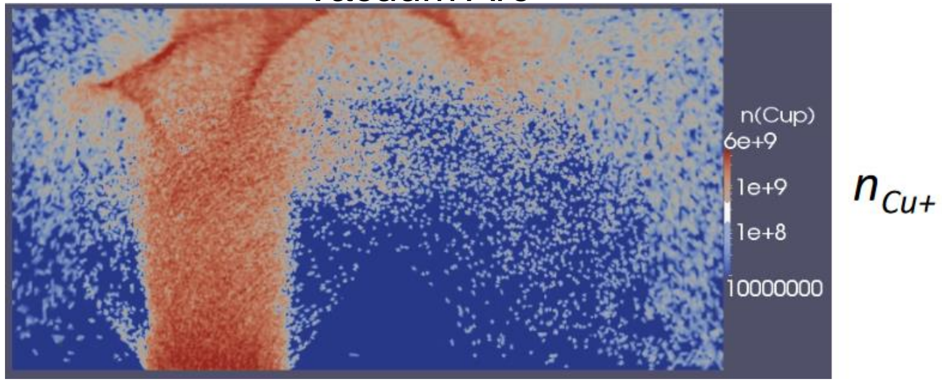
Used to address plasmas from vacuum arcs to atmospheric pressure streamers.

Streamer Propagation



A. S. Fierro, et al., Plasma Sources Sci. Technol. 27, 105008 (2018).

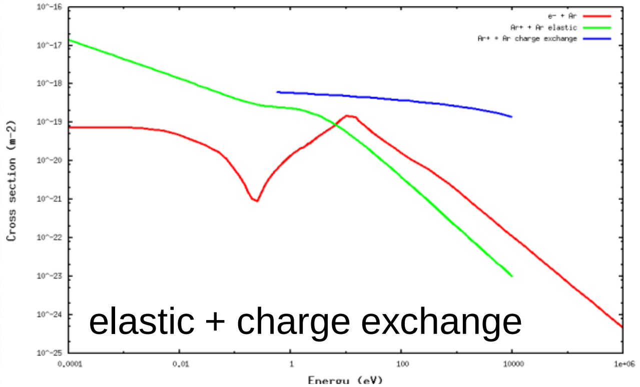
Vacuum Arc

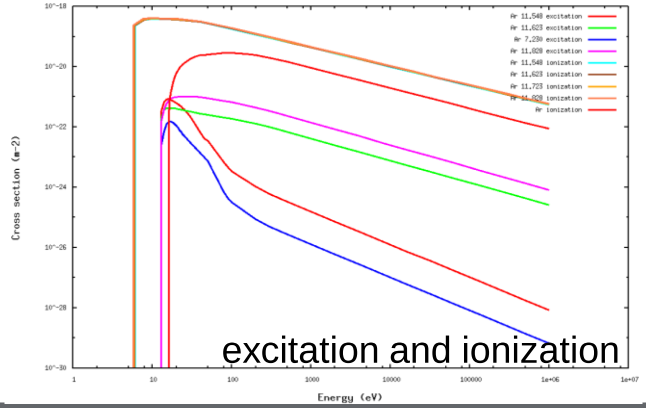


M. M. Hopkins, et al., International Conf. on Numerical Simulation of Plasmas, Beijing, China (2013).

Typical RF plasma discharge

The focus of this particular presentation is on the Argon, RF plasma discharge. A one-dimensional, RF discharge is simulated using particle-in-cell (PIC) and Direct Simulation Monte Carlo (DSMC)





Model uses Argon cross sections with the following collisions considered

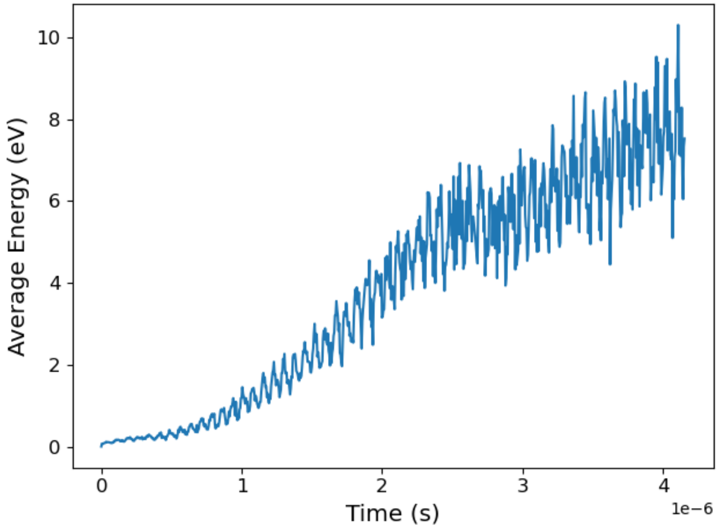
- Electron elastic
- Excitation
- Ionization
- Multi-step ionization
- Ion elastic
- Ion charge exchange

Numerical Parameters

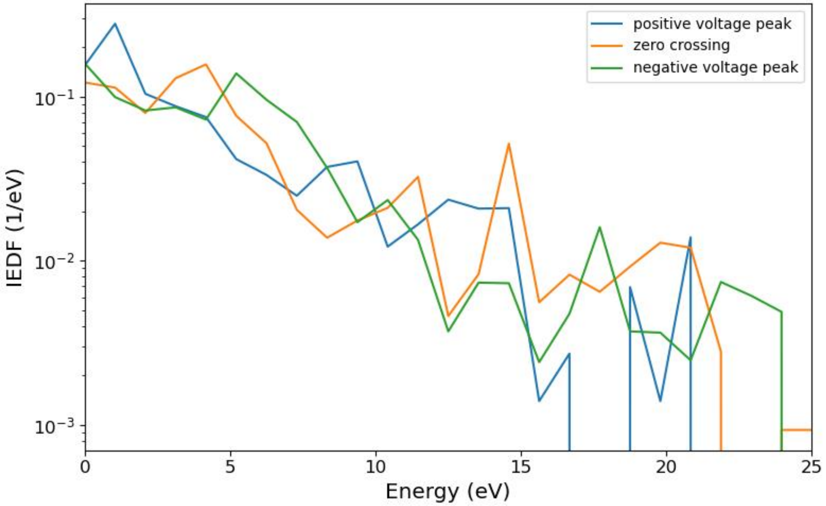
- $\Delta t = 1$ ps
- $\Delta x = 5\ \mu m$
- 200 charged particles per element
- f = 13.56 MHz
- P = 300 Pa

Typical RF plasma discharge

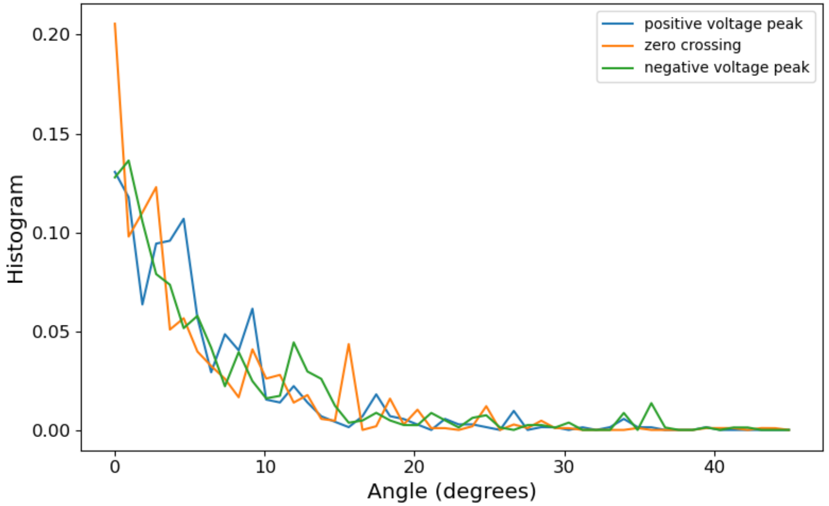
Ion characteristics can be determined by analyzing particles reaching the boundaries



Average Ion Energy in simulation volume



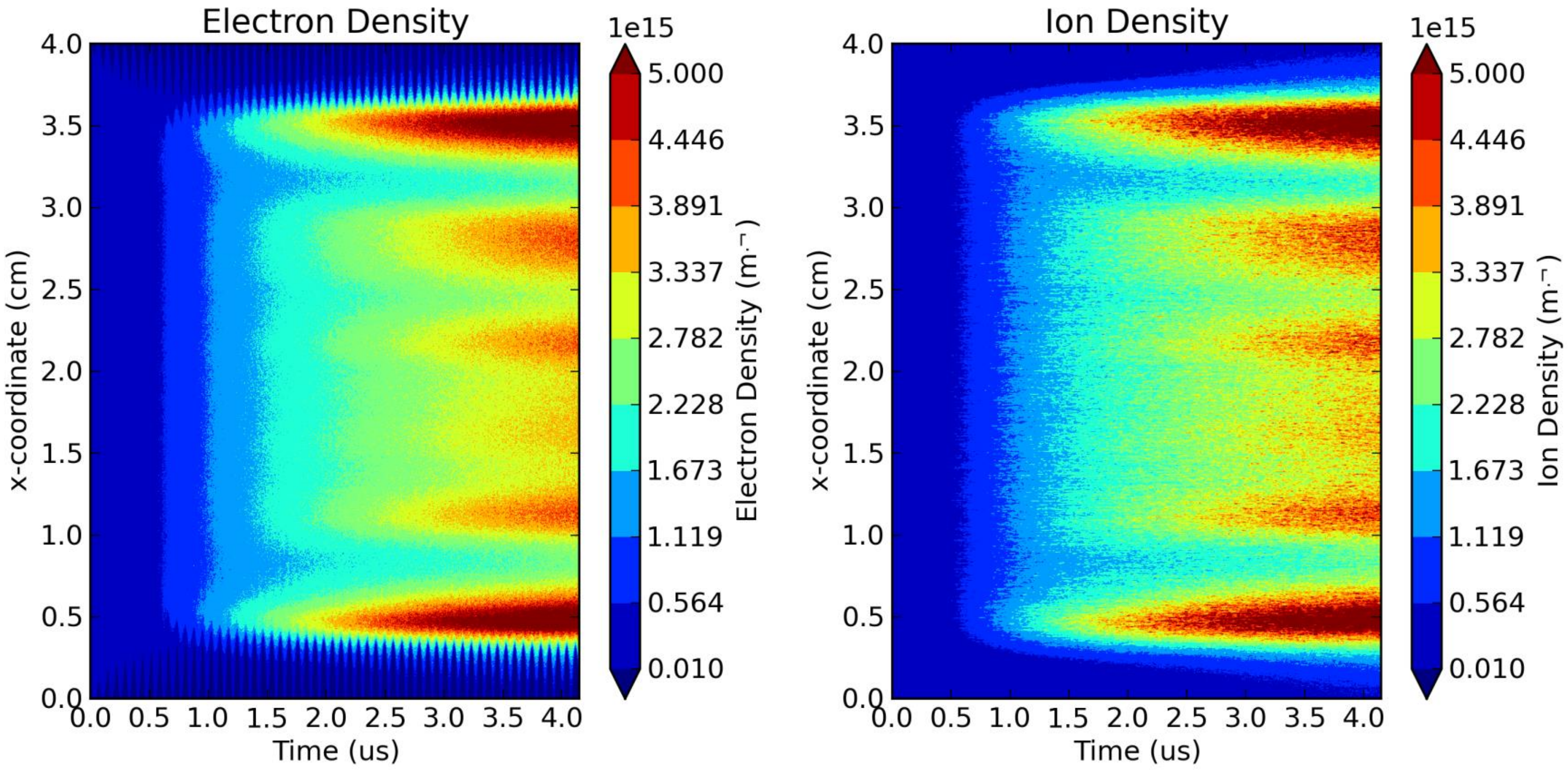
IEDF at one of the surfaces



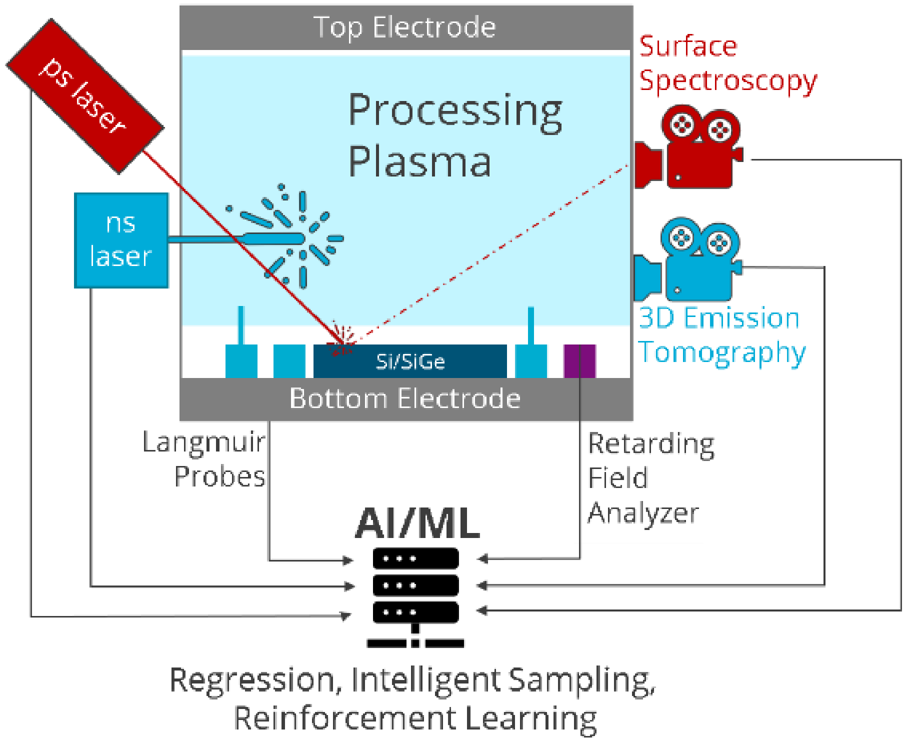
Ion angle of impact

Typical RF plasma discharge

Spatio-temporal development of electron and ion densities



Machine learning



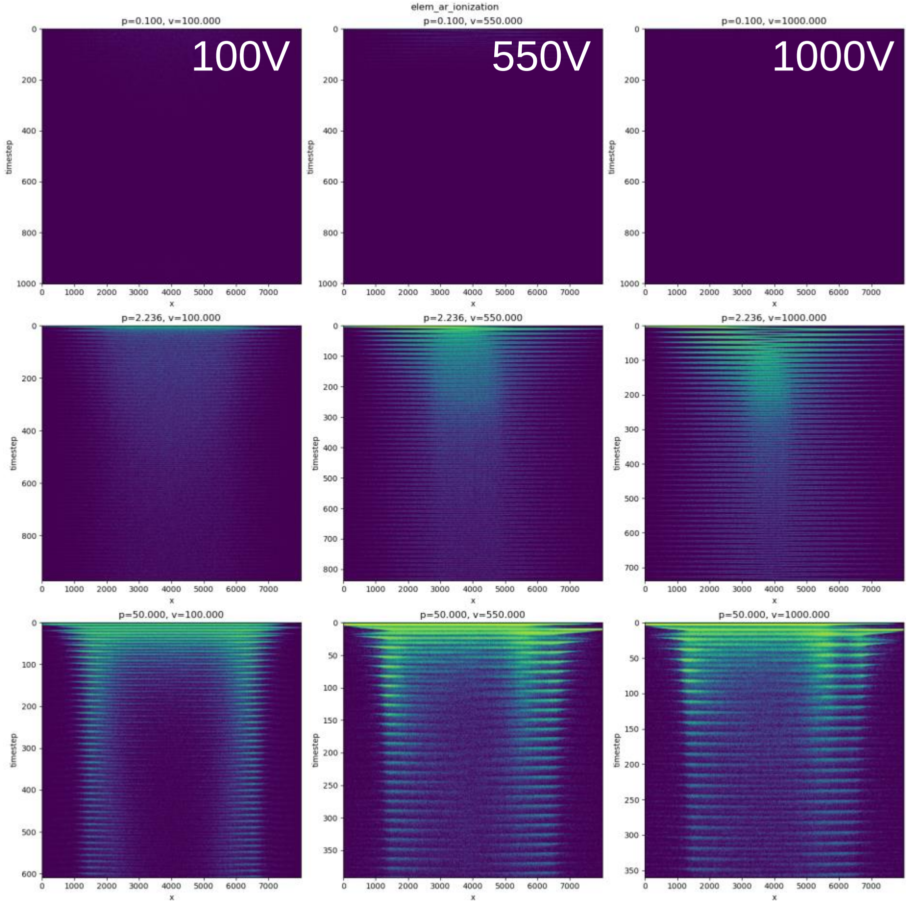
- Supervised learning of IEDF as a function of processing parameters
- “Hidden Field” inference as a function of measurable parameters
- Dimensionality reduction on state of plasma
- Surrogate model of system

Developing large amounts of training data

- Would like to simulate a pressure range between 0.1 Pa and 300 Pa with driving voltages (peak to peak) of 200 V to 2000 V. These will be the only two variables that are swept in our initial training data set.
- Each simulation takes on the order of 4 to 10 (sim time of 5 us) days to run on 144 processors and 8000 elements. Particle counts are on the order of 5 million.
- To simulate 30 pressures and 30 voltages (900 Simulations), low-order estimate of 12 million CPU hours. Data size estimate ~ 100 TB
- Must use High-Performance Computing
  - Our research effort obtained priority on the Ghost Super Computer (740 Nodes, 26,640 Cores, 128 GB RAM/node) at Sandia National Laboratories to use 3000 processors for 5 months continuously.
- Simulations are run and managed through scripts to generate input decks and launch simulations on the cluster – Have to manage system downtime !

A small subset of data (3x3 data set)

Ionization rate

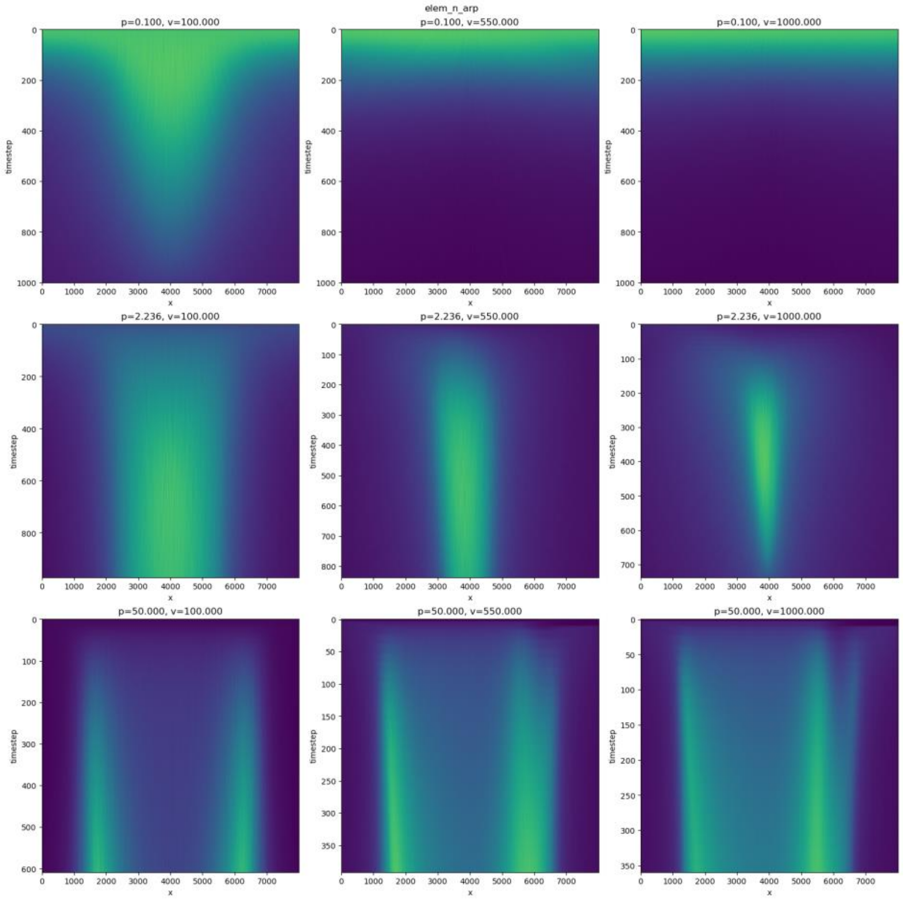


0.1 Pa

2.236 Pa

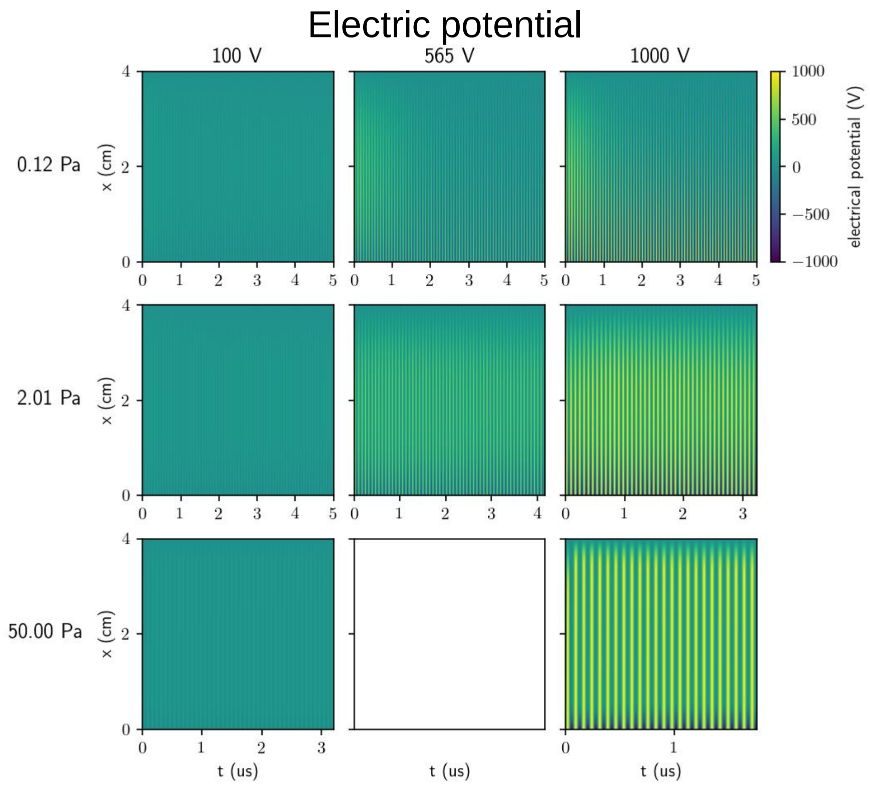
50 Pa

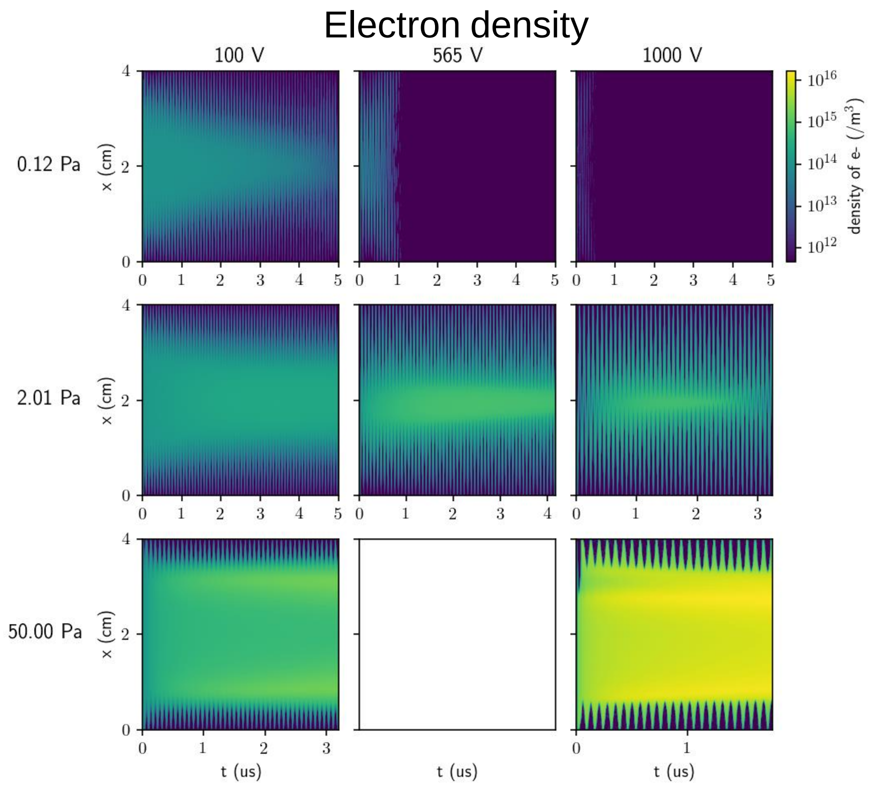
Ion density



The large data set (30x30)

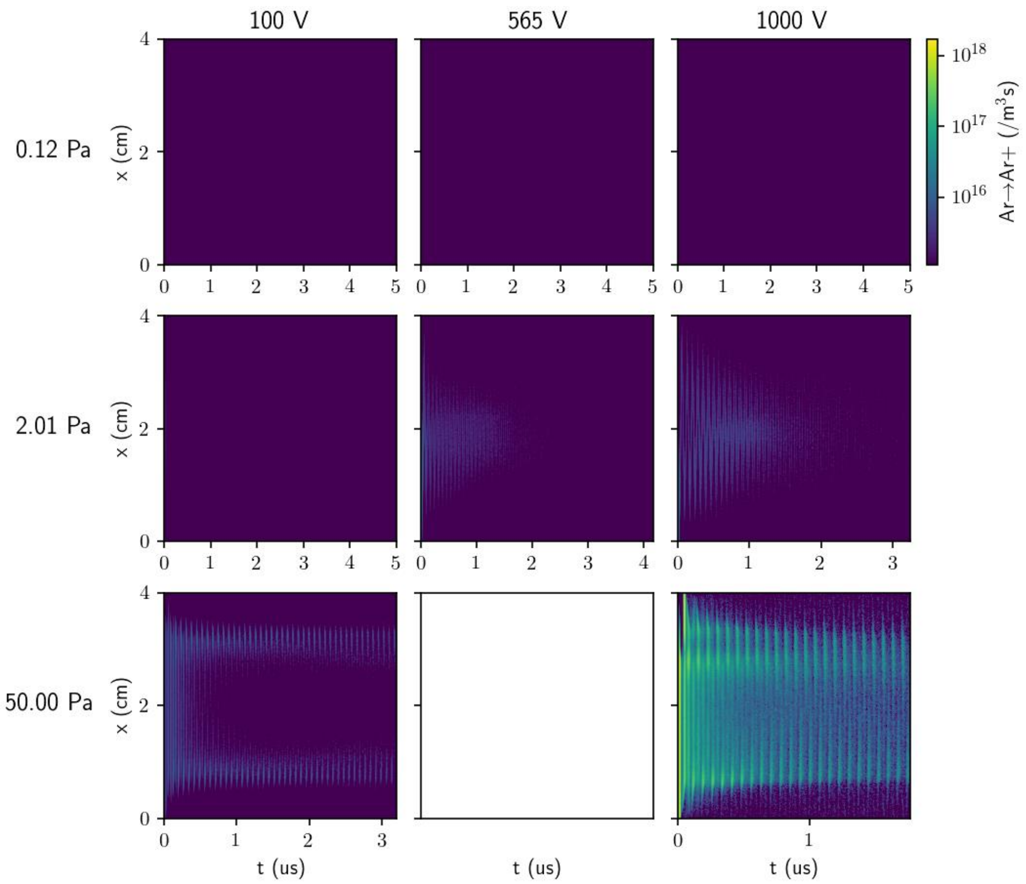
- The small data set was a test of the toolchain for input decks to be generated, submission scripts, and restart capabilities. Furthermore, it enabled development of data visualization and processing tools.



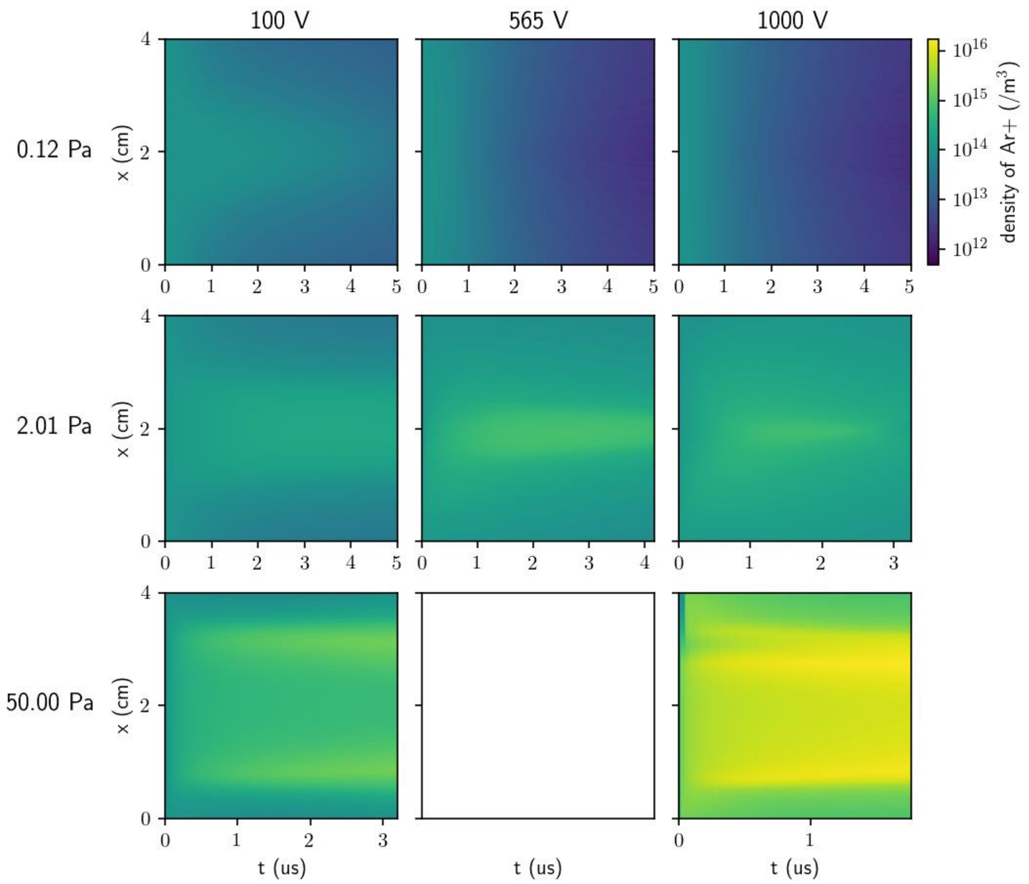


The large data set (30x30)

Ionization rate

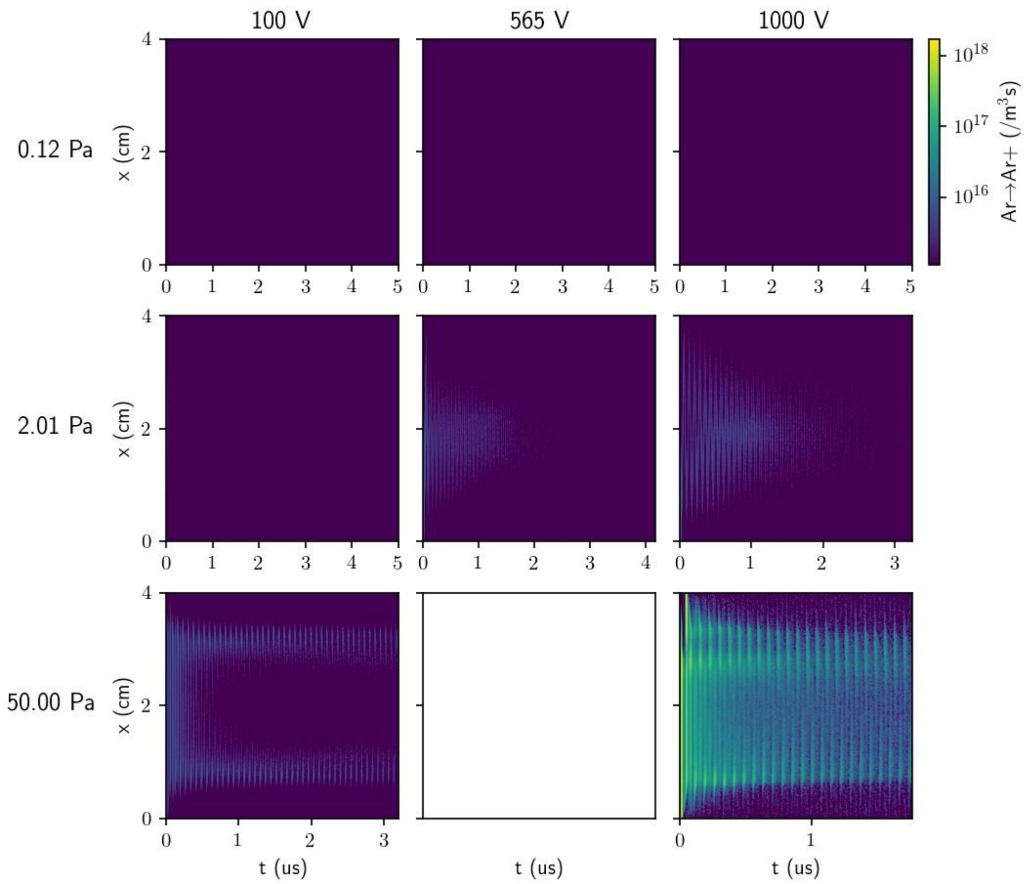


Ion density

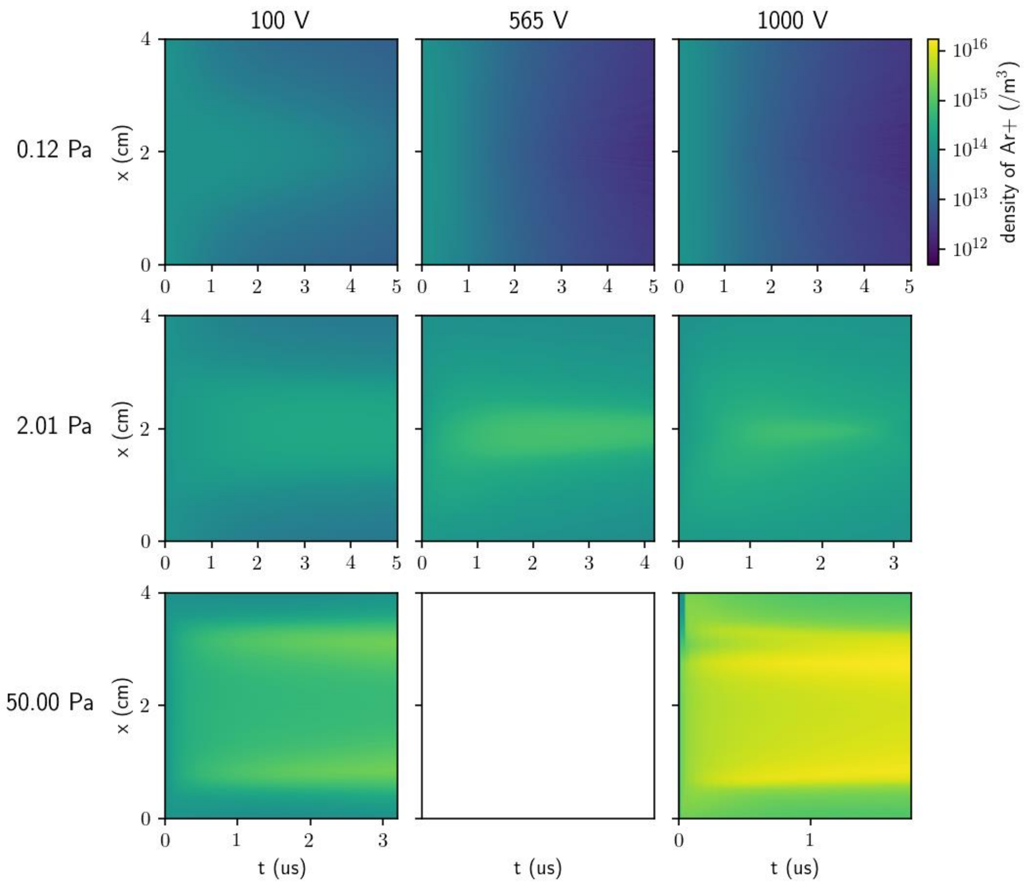


The large data set (30x30)

Ionization rate

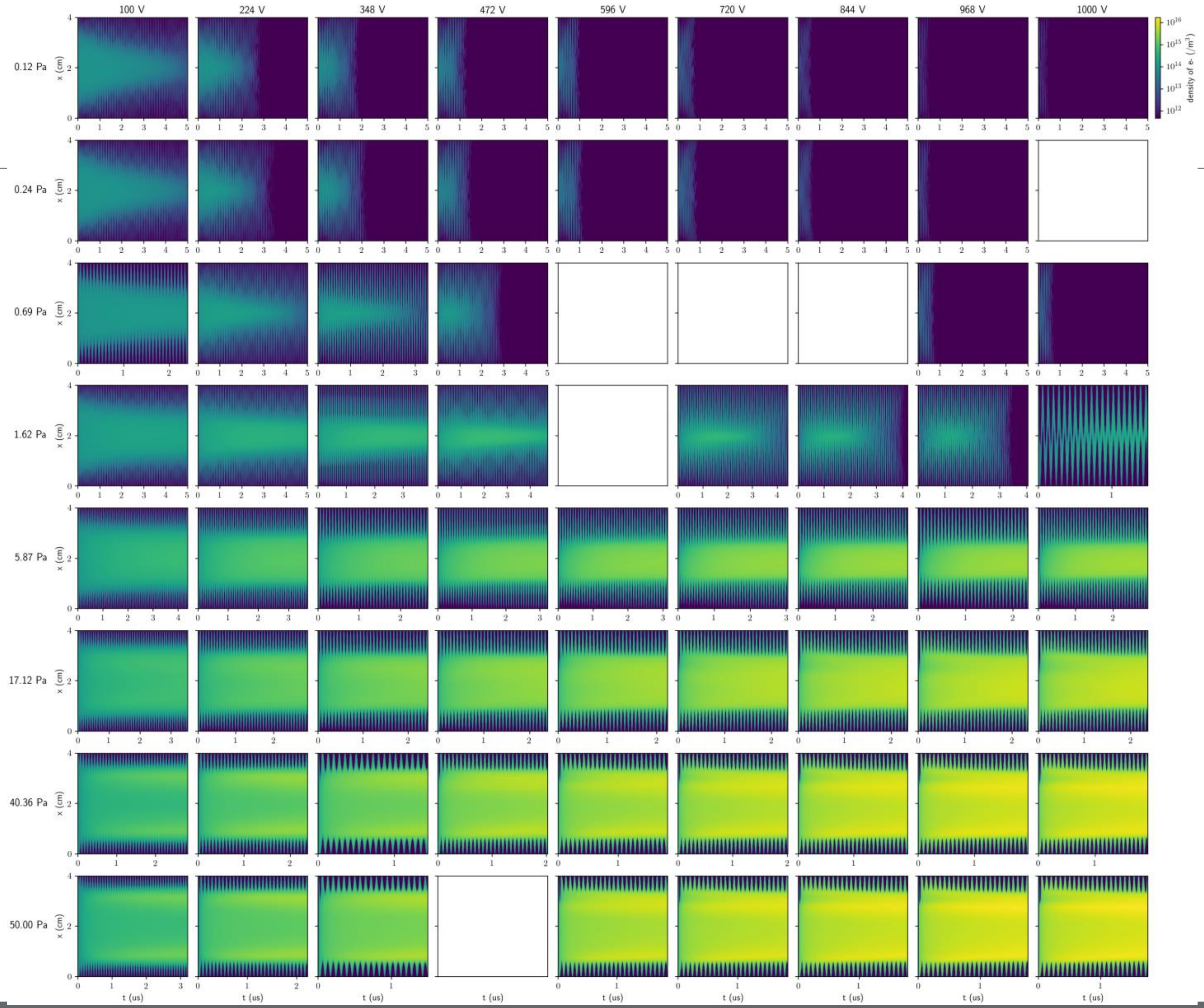


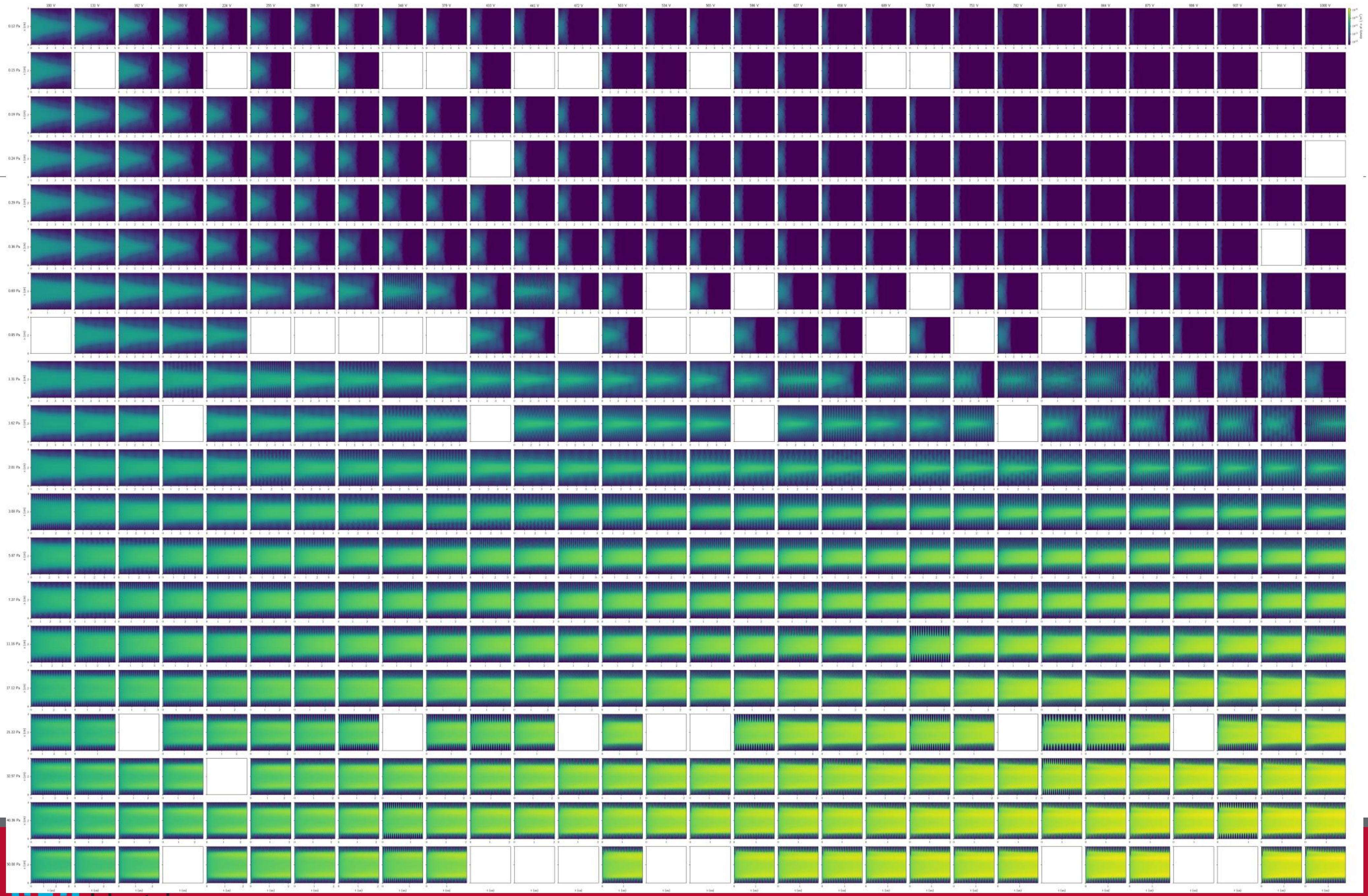
Ion density



Increasing Voltage

Increasing Pressure

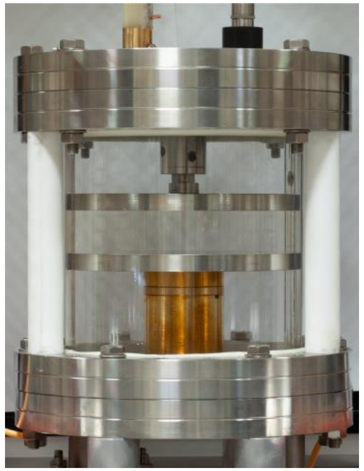


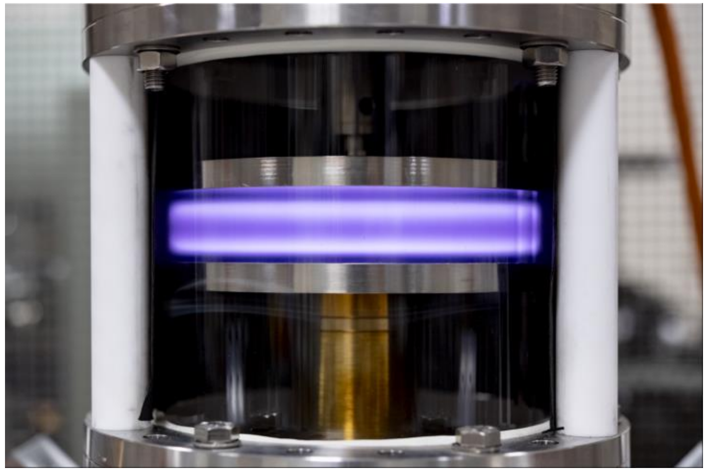


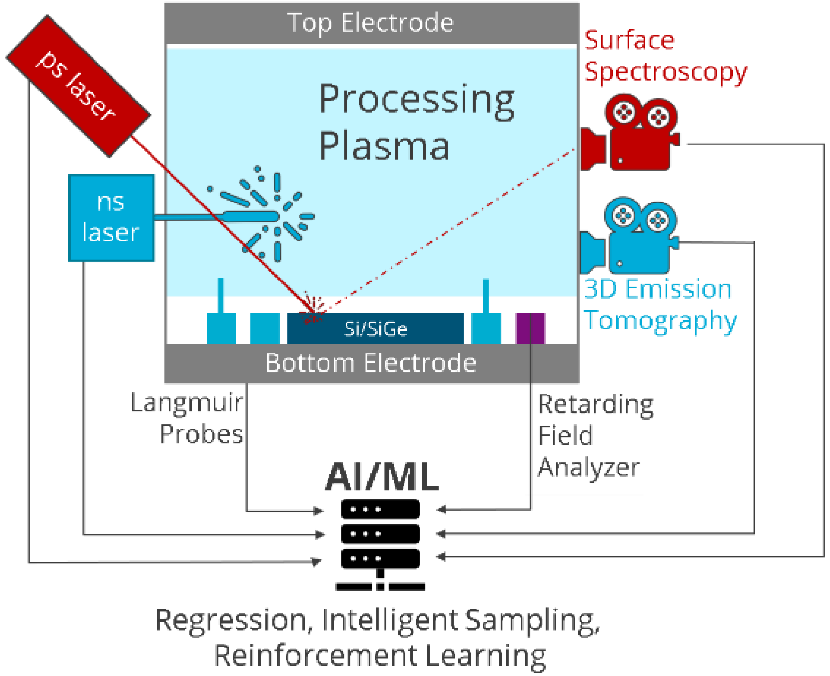
Laboratories

An experimental approach

An analogous experimental setup is being constructed based upon the Budapest Cell* to not only provide validation data for the models, but also serve as an input to the machine learning model as well.







* B. Horváth, A. Derzsi, J. Schulze, I. Korolov, P. Hartmann, and Z. Donkó. *Plasma Sources Science and Technology* 29, no. 5 (2020): 055002.

Conclusions

- Have developed a one-dimensional PIC/DSMC model of an RF plasma operating in Argon gas to analyze plasma characteristics relevant to atomic layer etch.
- This model is being used to generate a large amount of data for training a machine learning model.
- This data will be published in an online database for others to use
- Simultaneous experimental development of an Argon plasma reactor is being stood-up to provide simulation validation and further data input for the machine learning model.

Acknowledgements

- This work used resources of the Sandia Laboratory Plasma Research Facility (DOE FES) and was supported by the Laboratory Directed Research and Development (LDRD) program at Sandia National Laboratories
- Sandia is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.