Machine-Learning-Assisted Analysis of Visible Spectroscopy in Pulsed-Power-Driven Plasmas

Rishabh Datta, Faez Ahmed, and Jack D. Hare, *Associate Member, IEEE*

***Abstract*— We use machine-learning (ML) models to predict ion density and electron temperature from visible emission spectra, in a high-energy density pulsed-power-driven aluminum plasma, generated by an exploding wire array. Radiation transport simulations, which use spectral emissivity and opacity values generated using the collisional-radiative code PrismSPECT, are used to determine the spectral intensity generated by the plasma along the spectrometer's line of sight (LOS). The spectra exhibit Al-II and Al-III lines, whose line ratios and line widths vary with the density and temperature of the plasma. These calculations provide a 2500-size synthetic dataset of 400-D intensity spectra, which is used to train and compare the performance of multiple ML models on a three-variable regression task. The AutoGluon model performs best, with an R^2-score of roughly 98% for density and temperature predictions. Simpler models [random forest (RF), k-nearest neighbor (KNN), and deep neural network (DNN)] also exhibit high R^2-scores (>90%) for density and temperature predictions. These results demonstrate the potential of ML in providing rapid or real-time analysis of emission spectroscopy data in pulsed-power-driven plasmas.**

***Index Terms*— Machine learning, magnetohydrodynamics (MHD), spectroscopy.**

I. Introduction

SPECTROSCOPY is a powerful technique for inferring plasma parameters from emitted electromagnetic radiation. For instance, line widths and line ratios can be used to determine electron density and temperature [1], [2], [3], velocity can be determined from the Doppler shift of spectral lines [1], [4], and magnetic field strength can be inferred from the Zeeman splitting of line radiation [5], [6]. The wide applicability of spectroscopy makes it an attractive tool for implementation in a variety of laboratory plasmas [1], [4], [7], [8], [9].

In emission spectroscopy, a typical intensity spectrum can contain several peaks (called emission lines) overlaid on a continuum [1]. The lines correspond to bound–bound electron transitions in the ions of the plasma, while the continuum emission results from free–free (Bremsstrahlung emission) and free–bound electron transitions (recombination radiation) [1], [4]. Line radiation generated by the plasma arises due to either collisional or radiative processes [1]. Collisional processes, such as electron impact excitation/de-excitation and three-body recombination, change the energy levels of bound electrons via collisions with other electrons [1], [4]. Similarly, radiative processes, such as photoexcitation/de-excitation, induce energy transitions due to the interaction of bound electrons with photons [1], [4]. Collisonal-radiative models balance the rates of excitation (and ionization) against that of de-excitation (and deionization), to determine the spectral emissivity and opacity of radiation emitted from the plasma [2].

The typical approach to determining the ion density n_i from emission spectra is to identify lines dominated by Stark (collisional) broadening, and then to compare the line widths with tabulated data [1], [9], [10], or with the predictions of collisional-radiative codes, such as PrismSPECT [11], [12]. Similarly, for the characterization of electron temperature T_e, we typically compare the intensity ratios of two or more lines (typically, interstage lines for which density changes have a small effect) with the predictions of collisional-radiative models [1], [3], [10].

When the plasma is not optically thin, radiation transport, which describes how the energy distribution of radiation changes as it propagates through an absorbing, emitting, and/or scattering medium, must be adequately modeled for accurate interpretation of emission spectra. The optical thickness of a material to radiation of frequency ω is characterized by $\tau \equiv \int \alpha(\omega, s) ds$, which is the line integral of the spectral opacity $\alpha(\omega)$ along the path s [4], [13]. When $\tau \ll 1$, the plasma is optically thin, and the output spectrum is simply the line-integrated emissivity $\epsilon(\omega)$ of the plasma along the path s. Similarly, for $\tau \gg 1$, the plasma is optically thick, and the output spectrum (for a plasma in local thermodynamic equilibrium) is the Planckian [4], [13]. In plasmas that are not optically thin, the radiation spectrum recorded by the spectrometer is significantly altered by radiation transport. Furthermore, if the plasma exhibits spatial inhomogeneity along the line-of-sight (LOS), the resulting spectrum may be dominated by strongly emitting or absorbing regions.

In high-energy-density pulsed-power-driven systems, the condition for optical thinness may not be satisfied [14], [15]. In this article, our focus is to diagnose pulsed-power-driven plasmas for laboratory astrophysics applications, which typically exhibit ion densities and electron temperatures between $n_i \approx 1 \times 10^{17} - 1 \times 10^{19}$ cm^{-3} and $T_e \approx 1 - 50$ eV,

Manuscript received 31 August 2023; revised 19 January 2024; accepted 3 February 2024. This work was supported in part by NSF and National Nuclear Security Administration (NNSA) under Grant PHY2108050, in part by EArly-concept Grants for Exploratory Research (EAGER) under Grant PHY2213898, and in part by the MathWorks Fellowship. The review of this article was arranged by Senior Editor F. Beg. *(Corresponding authors: Rishabh Datta; Jack D. Hare.)*

Rishabh Datta and Jack D. Hare are with the Plasma Science and Fusion Center, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: rdatta@mit.edu; jdhare@mit.edu).

Faez Ahmed is with the Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TPS.2024.3364975.

Digital Object Identifier 10.1109/TPS.2024.3364975

respectively [14], [16], [17], [18]. Pulsed-power devices generate plasma flows by driving large currents (1-30 MA) through thin ∼10–100 μm diameter wires. These plasmas are not optically thin to visible radiation; therefore, radiation transport modeling becomes important for spectral analysis [19], [20].

A key drawback of the aforementioned emission spectroscopy analysis approach is that it requires significant collisional-radiative and radiation transport modeling for the analysis of a given spectrum. The use of machine-learning (ML) models can reduce the computation time required for spectral analysis, especially for large batches of spectral data, and provide rapid real-time results during experimentation. Spectroscopy has previously been combined with supervised ML techniques, primarily in lower density plasmas. Visible emission spectroscopy combined with regression methods and neural networks has been used to predict density and temperature in low-temperature low-density ($n_e \sim 1 \times 10^{10}\text{cm}^{-3}$ and $T_e \sim 1$ eV) laboratory plasmas [21]. Similarly, neural networks have been used to predict electron energy distribution in low-temperature nonthermal plasmas [22], and classifiers have been used for trace element and impurity detection in RF-generated plasmas [23]. Neural network regressors and classifiers have also been shown to accurately predict electron temperature and divertor detachment from UV/Extreme Ultraviolet (XUV) spectroscopy measurements in magnetic confinement fusion devices [24], [25], [26]. In the examples above, the labeled dataset for training is generated by simultaneous spectroscopy and independent density and temperature measurements with other diagnostics, like Thomson scattering or Langmuir probes [24], [25], [26]. This eliminates the need for complicated theoretical or computational modeling. However, simultaneous independent measurements using secondary diagnostics are not always possible; for example, in experiments with poor diagnostic access or diagnostic unavailability. Moreover, alternative diagnostics, such as Langmuir probes also perturb the plasma. Finally, in experiments with low repetition rates, such as in pulsed-power plasmas, it is challenging to generate a purely experimental dataset for the data-hungry supervised ML task. In such situations, we must, therefore, rely on synthetic data for training.

In this article, we use an ML approach to predict electron density and temperature profiles along the spectrometer line of sight (LOS) from visible emission spectroscopy data in a pulsed-power-driven aluminum plasma. In contrast to previous work [21], [22], [25], [26], which focuses on lower density and/or optically thin plasmas, here we aim to characterize high-density nonoptically thin spatially inhomogeneous plasmas, characteristic of pulsed-power-driven plasmas generated using exploding wire arrays [16], [17]. We approach the problem in two parallel ways. In the first approach, we frame the prediction problem as a single-objective optimization problem, where we minimize the deviation between a simulated spectrum and a target spectrum. We generate the simulated spectra using collisional-radiative calculations performed using PrismSPECT, which are then used to solve radiation transport in a spatially inhomogeneous plasma. In the second approach, we solve a multivariable regression problem, where we predict ion density $n_i(s)$ and electron temperature $T_e(s)$ as a function of position s along the spectrometer LOS from a given spectrum. We train multiple supervised ML models—linear regressor, k-nearest neighbor (KNN), decision trees (DTs), random forest (RF), deep and convolution neural networks, and AutoGluon—using synthetic data, generated from radiation transport simulations. The AutoGluon model performs best, with an R^2-score of roughly 98% for density and temperature predictions, and significantly reduces the computation time over optimization-based curve fitting methods. Our results demonstrate the potential of ML methods in providing rapid or real-time analysis of emission spectroscopy data in pulsed-power-driven plasmas.

II. Dataset Generation

A. Radiation Transport Modeling

We use PrismSPECT [12] to compute emissivity $\epsilon(\omega)$ and opacity $\alpha(\omega)$ values for an aluminum plasma, in the visible range of the electromagnetic spectrum ($400\,\text{nm} < \lambda < 700\,\text{nm}$). We use a steady-state nonlocal thermodynamic equilibrium model with Maxwellian free electrons. We run 10 000 PrismSPECT simulations, for electron temperatures T_e linearly distributed in the range [0.5, 25] eV, and ion density n_i logarithmically distributed between $[1 \times 10^{16}, 1 \times 10^{19}]\,\text{cm}^{-3}$.

Our in-house radiation transport solver computes the output intensity spectrum $I_\omega(s)$ given spatially varying emissivity $\epsilon_\omega(s)$ and opacity $\alpha_\omega(s)$ values, by solving the steady-state radiation transport equation along the 1-D path s [13]

$$\partial_s I_\omega(s) = \epsilon_\omega(s) - \alpha_\omega(s) I_\omega(s). \tag{1}$$

Emissivity and opacity values vary along the LOS due to spatial variations in density and temperature. PrismSPECT computes ϵ_ω and α_ω for spatially homogeneous plasmas. To construct the spatially varying emissivity and opacity, we first assume some m-dimensional density and temperature distributions $n_i(s)$, $T_e(s) \in \mathcal{R}^m$ along the LOS. Here, m is the number of points required to capture the spatial variation along the LOS. We then calculate the emissivity and opacity values at each position s along the LOS from the values of $n_i(s)$ and $T_e(s)$ at that location. The radiation transport solver then determines the output intensity distribution by solving (1). We use a spectral resolution of 3.75×10^{-3} eV for our radiation transport calculations, which results in a $[400 \times 1]$ dimensional intensity spectrum for each position.

In Fig. 1, we compare the output intensity generated by the radiation transport solver with that from a planar plasma simulation in PrismSPECT. Here, the ion density and electron temperature of the plasma are $n_i = 5 \times 10^{17}\,\text{cm}^{-3}$ and $T_e = 2.5$ eV, respectively, while the size of the plasma is 10 mm. The planar plasma simulation includes the effect of radiation transport, for the case of a homogenous (i.e., constant density and temperature) slab of a specified size. The output of the radiation transport solver agrees with that from PrismSPECT. Fig. 1 also shows the case for which the optical thickness is set to 0. For the $\tau = 0$ case, lines with high opacities are no longer damped by absorption from the plasma, and thus, the line ratios are significantly modified when compared with the case with radiation transport. This illustrates that in our

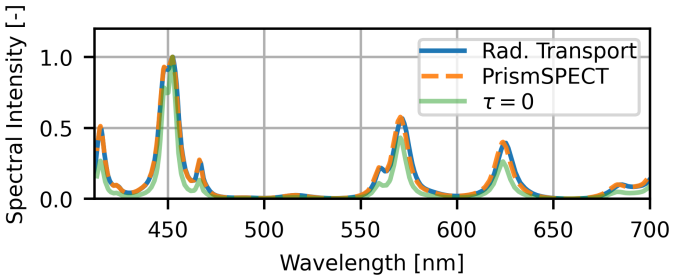

Fig. 1. Comparison of the output intensity calculated by the radiation transport solver with that from a planar plasma simulation in PrismSPECT ($n_i = 5 \times 10^{17}$ cm^{-3}, $T_e = 2.5$ eV, and $L = 10$ mm) and with a zero opacity $\tau = 0$ case.

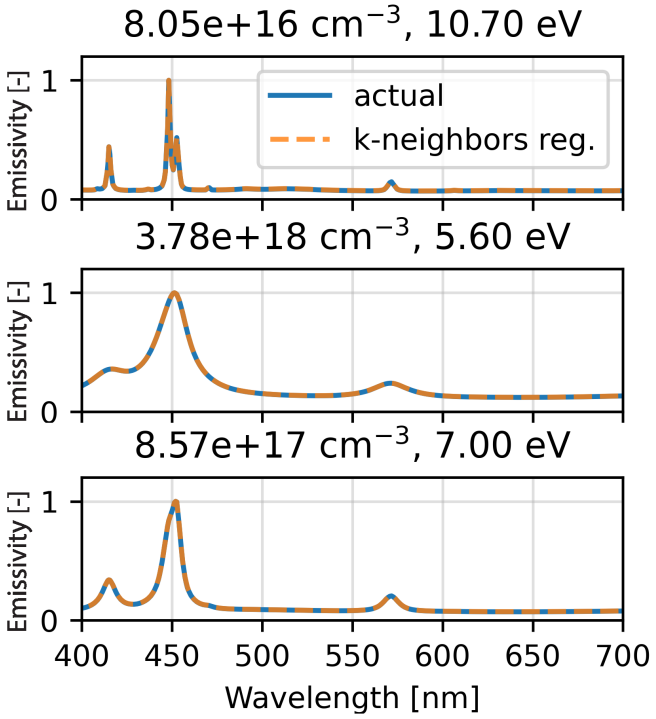

Fig. 2. Comparison of emissivity predictions made by a KNN model trained on PrismSPECT simulation results. We compare results for randomly chosen members of the test set. Emissivity is scaled to [0, 1].

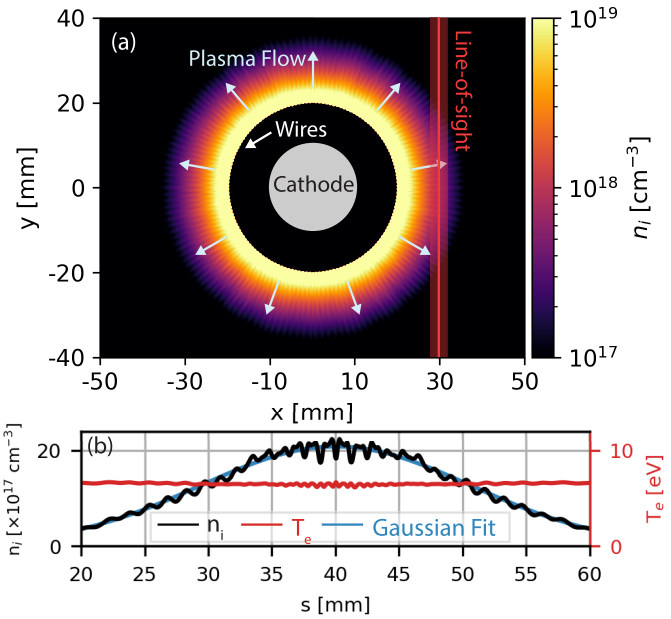

Fig. 3. (a) Simulated ion density at peak current, generated by a 40-mm diameter exploding wire array with 150 aluminum wires, driven by a 10-MA current pulse (300-ns rise time). This simulation was performed using GORGON, a two-temperature resistive MHD code. (b) Variation of density and temperature along a chordal LOS as shown in (a).

pulsed-power-driven plasma of interest, optical thickness is important, and must be included in the spectroscopy analysis.

The radiation transport solver requires emissivity and opacity calculated at each density and temperature value along the LOS as inputs to generate the intensity spectrum. We interpolate the emissivities and opacities for the intermediate values not simulated with PrismSPECT, using a weighted KNN regressor with $k = 4$, trained on the output of the 10 000 PrismSPECT simulations. The output is the weighted average of the four closest neighbors, based on Euclidian distance in $[T_e^*, n_i^*]$ space. Note that T_e^* and n_i^* are normalized values of the temperature and density based on (3), and the weights are inversely related to the distance to the neighbors. To evaluate the performance of the regressor, we compare the predicted spectra with the previously unseen emissivity spectra in the test set. As observed from Fig. 2, where we compare the predicted emissivity with the actual emissivity for randomly chosen members of the test set, the predictions agree well with the actual spectra. The coefficient of determination (also called the R^2-score) is a commonly used metric to characterize the performance of regression, and is defined as follows:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}. \tag{2}$$

Here, y_i is the predicted value, $\hat{y}_i$ is the actual value, and $\bar{y} \equiv 1/n \sum_{i=0}^{n} y_i$ is the mean of the actual values. For the problem above, the KNN regressor exhibits an R^2-score of 99.62%, showing that the model accurately reproduces the emissivity and opacity spectra for density and temperature values not included on the simulation grid in PrismSPECT.

B. Density and Temperature in Exploding Wire Arrays

Exploding wire arrays, which consist of a cylindrical cage of wires around a central cathode, are commonly used sources of pulsed-power-driven plasma for laboratory astrophysics experiments [14], [16], [17], [19]. The magnetic field is oriented in the azimuthal direction inside the wires, and results in a $\mathbf{j} \times \mathbf{B}$ force that accelerates the ablating plasma radially outwards from the wires. Due to radially diverging flows, the density decays rapidly with distance from the wires. This can be observed from Fig. 3(a), which shows the simulated ion density distribution generated by a 40-mm diameter exploding wire array with 150 aluminum wires, driven by a 10-MA current pulse with a 300-ns rise time. This simulation was performed using GORGON—a two-temperature Eulerian resistive magnetohydrodynamic (MHD) code [20], [27].

As discussed before, we require m-dimensional arrays to fully capture the spatial variation in density and temperature

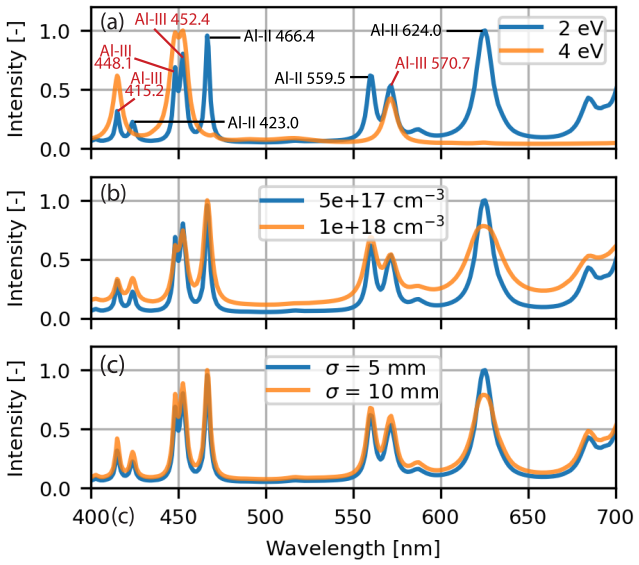

Fig. 4. Normalized spectral intensity simulated by the radiation transport solver. In (a)–(c), the blue curves correspond to output spectra generated for a Gaussian density variation ($n_0 = 5 \times 10^{17}\,\mathrm{cm}^{-3}$, $\sigma = 5\,\mathrm{mm}$), and constant temperature ($T_e = 2$ eV) along the spectroscopy LOS. (a) Change in the intensity spectrum with increasing temperature. (b) Change in the intensity spectrum with increasing density n_0. (c) Change in the intensity spectrum with increasing σ.

along the LOS. However, we can use simplifying assumptions to make the problem more computationally tractable. If the emission is recorded along a chordal LOS, as shown in Fig. 3(a), the density variation can be approximated as Gaussian [see Fig. 3(b)], i.e., $n_i(s) = n_0 \exp[(s - s_0)/2\sigma^2]$. Here, n_0 is the peak density, σ is the standard deviation of the Gaussian function, and we set the mean s_0 to half the total path length. The density distribution shown in Fig. 3 is consistent that measured experimentally with laser imaging interferometry in previous pulsed-power experiments [16]. Our MHD simulations [Fig. 3(b)] and previous experimental measurements also show little spatial variation in the temperature due to short thermal diffusion time in pulsed-power plasmas [17]. Therefore, we approximate the temperature to be constant along the LOS, i.e., $T_e(s) = T_0$. This allows us to reduce our $2 \times m$ dimensional problem to just three variables—n_0, T_0, and σ. Both the density and temperature in Fig. 3(b) also exhibit small amplitude modulations, which arise due to oblique shocks resulting from the azimuthal expansion of plasma from the discrete wires [28]. Our radiation transport calculations, however, show that the effect of these modulations on the recorded intensity spectrum is small.

Fig. 4 shows a synthetic intensity spectrum, generated by the radiation transport solver, with values $n_0 = 5 \times 10^{17}\,\mathrm{cm}^{-3}$, $T_e = 2$ eV, and $\sigma = 5\,\mathrm{mm}$. Here, we normalize the spectrum between [0, 1] by dividing by the maximum intensity. The spectrum exhibits Al-II and Al-III lines, which correspond to transitions in singly- (Mg-like) and doubly-ionized (Na-like) aluminum respectively. When the temperature is increased [Fig. 4(a)], the relative intensity of the Al-III lines compared to the Al-II lines increases. This is expected because the ionization is higher at a higher temperature, and thus, the relative population of the higher Z Al-III ions increases relative to the singly ionized Al-II ions. The Al-II and Al-III lines only appear simultaneously between 1.5 and 3.5 eV. In Fig. 4(a), at 4 eV, the Al-II lines are completely suppressed. When we increase the density [Fig. 4(b)], the lines not only become broader (due to Stark broadening) but also the line ratios change as well. This is because increasing the density also increases the optical thickness, and the optically thick lines are damped more strongly. This can be observed from Fig. 4(b) when we compare the relative intensity of the Al-II 466.4 nm line (which is relatively optically thin) with that of the higher opacity Al-II 624.0 nm line. The more optically thick Al-II 624.0-nm line is strongly damped at higher densities. Finally, changing the value of σ [Fig. 4(c)], also changes the line ratios because the optical thickness increases with the size of the plasma; however, the sensitivity of the spectrum to changes in σ is relatively smaller than that in density and temperature.

To generate our training dataset for the ML task, we randomly sample values of $1 \times 10^{16} \le n_0 \le 1 \times 10^{19}\,\mathrm{cm}^{-3}$, $5 \le \sigma \le 30\,\mathrm{mm}$, and $0.5 \le T_0 \le 25\,\mathrm{eV}$ from a uniform distribution. Our radiation transport solver then uses these sampled values to calculate intensity spectra for each n_0, σ, and T_e. We generate a total of 2500 [400 × 1] intensity spectra to use as training data for our three-variable regression problem. Finally, we scale and normalize the values of n_0 , σ and T_0 so that they lie within the interval [0, 1]. We use linear scaling for the temperature and σ, and logarithmic scaling for the density

$$\begin{aligned} n_0^* &= \frac{\log_{10}\left(n_0[\mathrm{cm}^{-3}]\right) - \log_{10}\left(1 \times 10^{16}\,\mathrm{cm}^{-3}\right)}{\log_{10}\left(1 \times 10^{19}\,\mathrm{cm}^{-3}\right) - \log_{10}\left(1 \times 10^{16}\,\mathrm{cm}^{-3}\right)} \\ T_0^* &= (T_0[\mathrm{eV}] - 0.5\,\mathrm{eV})/(25\,\mathrm{eV} - 0.5\,\mathrm{eV}) \\ \sigma^* &= (\sigma[\mathrm{mm}] - 5\,\mathrm{mm})/(30 - 5\ \mathrm{mm}). \end{aligned} \tag{3}$$

We also scale the intensity output of the radiation transport solver to the range [0, 1]. This means we only use the shape of the intensity spectrum for our prediction, which obviates the need for absolute intensity calibration.

III. Methodology

A. Single-Objective Optimization

Our goal is to predict density and temperature profiles given a measured intensity spectrum I_{target}. One way to frame this problem is as a single-objective optimization problem

$$\begin{aligned} \min_{\mathbf{x}} :\ & f(\mathbf{x}) = \mathrm{MSD}\left(I_{\mathrm{target}} - I(\mathbf{x})\right) \\ & \text{where:}\quad \mathbf{x} = [n_0, \sigma, T_0] \\ \text{s.t.} :\ & 1 \times 10^{16} \le n_0\left[\mathrm{cm}^{-3}\right] \le 1 \times 10^{19} \\ & 5 \le \sigma\ [\mathrm{mm}] \le 30 \\ & 0.5 \le T_0\ [\mathrm{eV}] \le 25. \end{aligned} \tag{4}$$

We minimize the mean squared deviation (MSD) between the target and predicted intensities. The objective function can be represented as follows:

$$\mathrm{MSD} = \frac{1}{N}\sum_{i}^{N}\left[I_{\mathrm{pred}}(\omega_i) - I_{\mathrm{target}}(\omega_i)\right]^2. \tag{5}$$

Here, $I_{\text{pred}}(\omega_i)$ and $I_{\text{target}}(\omega_i)$ are the simulated and target intensities at frequency ω, and $N = 400$ is the size of the intensity spectrum. We use the radiation transport solver described in Section II-A to generate the predicted intensities.

We perform the optimization using a $(\mu + \lambda)$ genetic algorithm (GA) implemented using the pymoo package in Python. The GA optimization algorithm iteratively searches for solutions that minimize the objective function over multiple generations [29]. In each generation, the best-performing solutions are selected and included in the population for the next iteration. Solutions are combined in each iteration in a process called crossover to create offspring solutions. The solutions are also subject to random changes in the values of the variables $\mathbf{x}$ to increase the diversity of the solutions, in a process called mutation. For the optimization here, we use a randomly generated initial population size of 150, with simulated binary crossover and polynomial mutation (probability $= 0.5$ and distribution index $\eta_c = 1$). We terminate the optimization when the MSD of the solutions becomes lower than a specified threshold.

To determine the uncertainty in the optimization results, we repeat the optimization 50 times with different starting seeds to construct a family of optimal solutions. We exclude solutions from runs in which the GA gets stuck in a local minimum, where the objective function does not converge to a value below the required threshold. Repeating the optimization with different starting seeds (hence, different initial populations) allows us to construct distributions of n_0, T_0, and σ, which we then used to estimate the uncertainty in the solutions. In Section IV-A, we report the mean and the range determined from these distributions.

B. Multivariable Regression

An alternative approach is to formulate the problem as a multivariable regression problem. Given an input of an unseen intensity spectrum I_{target}, we predict the corresponding (normalized) values of n_0^*, T_0^*, and σ^* using ML-based regression models. Here, we compare the performance of multiple regressors for our three-variable regression task—linear regression (LR), KNN regressor, DTs, RF, deep neural network (DNN), and a 1-D convolution neural network (1-D CNN). The choice of regression algorithm often represents a trade off between model precision and interpretability. Simpler models, such as LR and KNN models, are relatively easier to understand and interpret, whereas DNNs and AutoML, which often provide high performance, require more training time, and are challenging to interpret [30].

The LR, KNN, DTs, and RF models are implemented using the scikit-learn [31] package in Python. The KNN algorithm predicts values based on the distance from the k nearest data points in the training set [30]. Here, we use $k = 8$ and Euclidian distance for our KNN regressor. DTs follow a flowchart-like structure and make predictions by asking questions at each level [30]. We use a DT regressor with a depth of 5 and a minimum sample split of 5. RFs are ensembles of many DTs [30]. Our RF regressor uses 140 estimators; the minimum samples required for a split is 4, and the minimum samples per leaf are 5.

We use the TensorFlow [32] package to construct a three-layer DNN. Neural networks consist of multiple "hidden" layers, consisting of several neurons, sandwiched between an input and output layers [33]. In our DNN architecture, each layer is a fully connected dense layer with 100 neurons (i.e., each neuron is connected to every other neuron in the previous and next layers), with ReLU activation functions, and padded with a batch normalization layer and a dropout layer ($p = 0.3$) to help prevent overfitting. The final layer consists of a 3-D dense layer. Similarly, the 1-D CNN consists of two 1-D convolution layers (kernel size $= 3$) with filter sizes of 8 and 16, respectively, and leaky ReLU activation layers. The convolution layer performs a convolution operation using the specified kernel on the input from the previous layer [33]. Each convolution layer is padded with a batch normalization layer and a $p = 0.1$ dropout layer. The convolution layers are followed by a 200-neuron dense fully connected layer and a 3-D output layer. Both neural networks use the Adam optimizer with a 0.5×10^{-4} learning rate and a mean squared error (MSE) loss function. The key difference between the DNN and CNN architectures is that the DNN treats the input vector as a 400-D vector of parameters, whereas the CNN treats it as a 1-D image, and therefore, has information about the relative spectral location of each element in the input vector.

Finally, we also implement a tabular AutoGluon model using Python's AutoGluon package [34]. AutoGluon provides an automated approach to ML, by automatically comparing and combining the performance of many different models. The performance of AutoGluon has previously been shown to exceed that of more traditional ML models [34].

Each regressor is trained on the 2500×400 synthetic spectra with a 2000:500 split between the training and test sets, and $k = 3$ stratified k-fold cross-validation for the training set. Members of the test set are not shown to the ML model during the regression task and are used to evaluate the regression performance after training. The hyperparameters described in this section were determined using hyperparameter optimization implemented using the Python package Optuna [35].

We use two metrics to characterize the performance of the different regression models. These are the coefficient of determination (also called the R^2-score), and the MSD from the simulated curve. The R^2-score is defined in (2), and is calculated from the predicted and actual values of n_0^*, T_0^*, and σ^*. Similarly, the MSD (5) measures the deviation of the predicted intensity spectrum from the actual spectrum. Here, the predicted spectrum is determined by feeding the predicted values of n_0^*, T_0^*, and σ^* from the ML model into the radiation transport solver. The R^2-score and MSD are calculated for the $N = 500$ test set, which contains spectra not previously seen by our ML models. For good performance, we aim to maximize the R^2-scores and minimize the MSD. The R^2-score may provide spurious performance metrics in case of nonunique solutions; whereas, the MSD characterizes how close the predicted spectra are to the actual spectra, allowing us to overcome this issue.

TABLE I
COMPARISON OF THE PERFORMANCE OF DIFFERENT ML MODELS ON THE REGRESSION TASK

Model	Training Time (s)	R2-Score(%)	2 Var. R2-Score(%)	Median MSD [$\times 10^{-4}$]
Linear Regression (LR)	0.3	-41.54 ± 80.76	64.00 ± 10.2	1.1
Decision Tree (DT)	0.35	53.22 ± 4.5	89.9 ± 2.57	5.3
k-Nearest Neighbor	0.02	67.89 ± 2.8	96.25 ± 1.8	0.4
Random Forest (RF)	6.41	67.06 ± 1.93	96.58 ± 0.5	0.4
Deep Neural Net. (DNN)	240.12	71.54 ± 1.2	94.52 ± 1.1	4.2
1D Conv. Neural Net. (1D-CNN)	260.17	65.34 ± 3.2	91.69 ± 1.4	5.7
AutoGluon	**4760.11**	$\mathbf{74.20 \pm 2.1}$	$\mathbf{98.98 \pm 1}$	**0.3**

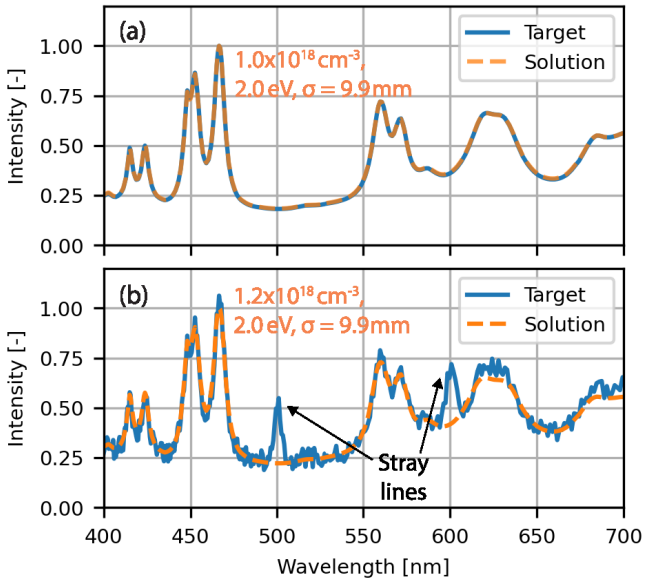

Fig. 5. (a) Comparison of target intensity spectra with spectra determined via optimization for the test case $n_0 = 1 \times 10^{18}\,\mathrm{cm}^{-3}$, $T_0 = 2\,\mathrm{eV}$, and $\sigma = 10\,\mathrm{mm}$. The target spectrum is the blue solid curve, while the predicted spectrum and the predicted values are in orange. (b) Robustness of the fitting to noise and stray lines.

IV. RESULTS AND DISCUSSION

A. Optimization Results

Using GA-based optimization, we predict values of n_0, T_0 and σ for several test cases. Fig. 5(a) shows the comparison of the target intensity spectra with spectra determined via optimization for a randomly selected test case ($n_0 = 1 \times 10^{18}\,\mathrm{cm}^{-3}$, $T_0 = 2\,\mathrm{eV}$, and $\sigma = 10\,\mathrm{mm}$). The predicted values of n_0, T_0 and σ reproduce the target spectrum well. For the test case shown in Fig. 5(a), the MSD from the target spectrum is roughly 6×10^{-5}, and the predicted values are $n_0 = 1.0 \times 10^{18}\,\mathrm{cm}^{-3} \pm 7\%$, $T_0 = 2.0\,\mathrm{eV} \pm 1\%$, and $\sigma = 9.9\,\mathrm{mm} \pm 20\%$. The range in the value of σ for the family of optimal solutions is comparatively higher than that in the density and temperature. This is consistent with our observation in Fig. 4(c), which shows that changes in σ generate linear changes in the optical depth, whereas those in n_0 and T_0 result in larger nonlinear changes in opacity and relative intensities.

In many real situations, the experimental data that we want to fit may be noisier than the synthetic spectrum shown in Fig. 5(a). Furthermore, the spectrum may also be contaminated by stray lines, generated by impurities in the plasma or by radiation emitted from other photoionized surfaces. In order to test the robustness of the prediction, we add noise and stray lines to the synthetic target spectrum [see Fig. 5(b)]. The added noise is randomly sampled from a continuous uniform distribution $U(-0.1, 0.1)$. The predicted solution reproduces the target spectrum well, despite the added noise and stray radiation. In this case, the MSD of the optimum solutions is about 3×10^{-3}, which as expected, is larger than that for the smooth case. The predicted solutions are $n_0 = 1.2 \times 10^{18}\,\mathrm{cm}^{-3} \pm 20\%$, $T_0 = 2.0\,\mathrm{eV} \pm 2\%$, and $\sigma = 9.9\,\mathrm{mm} \pm 30\%$. There is a larger uncertainty in the solutions for the noisy target spectrum when compared with the smooth target spectrum [Fig. 5(a)]; however, the predicted solutions still include the values used to generate the test case.

Although optimization using GA provides good results, the computational time is high (>10 min per prediction). This is primarily because the GA has to simulate a large number of potential candidates iteratively using our radiation transport model over several generations before convergence is achieved. Although we use GA as an example here, other optimization-based curve-fitting algorithms can also exhibit high computation times. This method can be highly effective with small datasets; however, it can be less attractive in cases where we must analyze large datasets of spectra, or when we require quick or real-time analysis of spectral information. The use of ML-based regression models, discussed in the next section, can be more useful in such situations.

B. Multivariable Regression Using Machine Learning

The training of ML models can be computationally time-intensive; however, once the training is complete, large datasets can be evaluated rapidly. The training time of models typically depends on the complexity of the model. Table I lists the comparison of the performance of the ML regressors described in Section III-B. Here, the R^2-score is computed using the predicted and actual values of all three variables, while we compute the two-variable R^2-score using only the value of density n_0^* and temperature T_0^*. The uncertainty in the R^2-score was determined from the range of scores calculated for each split in the cross-validation process. AutoGluon was the best-performing model, with an R^2-score of $74.20\% \pm 2.1\%$. Since AutoGluon trains and compares the performance of multiple different models simultaneously, the training time was high compared with the other models. The next best-performing models (DNN, RF, and KNN) exhibited R^2-scores between 67%–71% and shorter training times. The two-variable R^2 scores, calculated for density n_0^* and temperature T_0^* only, are roughly 30% higher than the three-variable

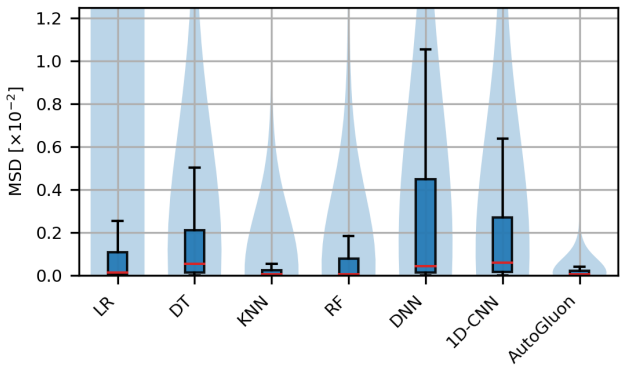

Fig. 6. Violin plots of the MSD of the predicted intensity spectra from the actual intensity spectra in the test set. The red lines represent the medians of the distribution, the blue rectangle represents the interquartile range, and the end-caps represent 1.5× the interquartile range. The blue-shaded regions show the shape of the full distribution.

R^2-scores, which shows that the models predict n_0^* and T_0^* with better accuracy than the third variable σ^*. There is a larger uncertainty in the prediction of σ^*, because as observed from Fig. 4(c) and in Section IV-A, changes in σ^* generate linear changes in the optical depth, whereas those in n_0^* result in larger nonlinear changes. We also find that in cases where σ^* is under predicted, the density is slightly over-predicted and vice versa, indicating that the model relies on small changes in density rather than on σ^* to incorporate the effects of changing optical depth. The majority of models, with the exception of LR and DTs, exhibit two-variable R^2-scores >90%, showing the effectiveness of these models in predicting n_0^* and T_0^*, despite the relatively poorer prediction of σ^*. AutoGluon, once again, exhibits the highest two-variable R^2-score (98.98% ± 1%).

In these models, we only use the shape of the intensity spectrum to make predictions. This can be advantageous as it obviates the need for absolute intensity calibration of the intensity spectra. However, adding information about the absolute intensity can potentially improve the predictions, as the absolute intensity also depends on the temperature, density, and size of the plasma. This would, however, require accurate modeling of intensity attenuation and losses in the optics used for the experiment. When we include the absolute intensity in the training data (that is, we do not normalize the intensity between [0, 1]), the models show improved prediction of σ, and the R^2-scores increase by roughly 3%–7%.

Fig. 6 shows the comparison of the MSD of the predicted intensity spectra from the actual intensity spectra in the test set for the different ML models. Here, we use the predictions of n_0, T_0, and σ from the ML regressors as inputs into the radiation transport solver to determine the predicted spectra. The red lines represent the medians of the distributions, the blue rectangles represent the interquartile range, and the shaded regions show the shape of the full distribution. As expected, the predictions of the AutoGluon model, which exhibit the highest R^2-score, also exhibit the lowest MSD, i.e., the predictions match the original intensity spectra well. The median MSD is relatively small ($<1\times10^{-3}$) for all the models; however, for models with lower R^2-scores, the distribution of the MSD exhibits a larger spread and extends to higher values. Interestingly, although the DNN exhibits R^2-scores similar to the RF and KNN models, its MSD distribution is significantly wider. For the DNN, the distribution is also wider than for the 1-D CNN model, which exhibited lower R^2-scores. This may indicate that although the predictions of n_0, T_0, and σ made by the DNN are close to the actual values, the values are consistently off, resulting in a relatively larger deviation between the predicted intensity calculated from these values and the actual intensity used to predict the values.

To gain further insight into the performance of the models, we plot the distribution of predicted values against the actual values. For the best-performing AutoGluon model (Fig. 7), we find that predicted values of n_0^* and T_0^* from the test set exhibit the least deviation from the diagonal, whereas predicted values of σ^* exhibit a relatively larger deviation from the diagonal, as expected. However, AutoGluon is approximately able to capture the relative trend in the values of σ^*, as seen in the positive slope of the distribution in Fig. 7(b). This is in contrast to other models, which exhibit larger deviations in the predictions of σ^*. In Fig. 7(a) and (c), deviations from the diagonal also become more pronounced at higher densities and temperatures (indicated by the red rectangles in these figures). Our radiation transport simulations show that at these higher temperatures ($T_0 \gtrsim 20\,\mathrm{eV}$) and densities ($n_0 \gtrsim 1\times10^{18}\,\mathrm{cm}^{-3}$), continuum emission begins to dominate over line emission, and the spectrum can become completely devoid of lines. This may make it harder to predict n_0 and T_0 from the shape of the spectrum.

To compare the performance of ML methods against GA optimization, we randomly select a 25-member subset from the test set used for ML regression. Using this subset, we make predictions with both AutoGluon and GA optimization. The evaluation time for this subset using the trained AutoGluon model was ∼1 s, while the optimization took significantly longer (>1.5 h). For AutoGluon, the median MSD was 5×10^{-5}, and the two- and three-variable R^2-scores were 98% and 72%, respectively, similar to that of the larger test set (see Table I). The GA predictions demonstrated a median MSD of about 4×10^{-4}, which is similar to that of the AutoGluon model. AutoGluon thus demonstrates a performance close to optimization-based curve fitting in minimizing the deviation between actual and predicted spectra but requires significantly less computational time. The two-variable R^2-score for density and temperature prediction was about 98% for the optimization, similar to that of AutoGluon, once again showing the excellent prediction capability for these variables. However, in predicting σ, AutoGluon outperformed the optimization, which demonstrated a three-variable R^2-score of about 65% for this subset.

A key challenge with more complicated ML models, such as neural networks and AutoGluon, is the lack of interpretability. To gain insight into features that inform the prediction, we investigate the relative importance of different parts of the intensity spectrum for the prediction of density and temperature. In order to do so, we calculate the sensitivity of the density and temperature predictions to perturbations in the spectral intensity value at different wavelengths. Fig. 8 shows the sensitivity map computed for the AutoGluon model for a

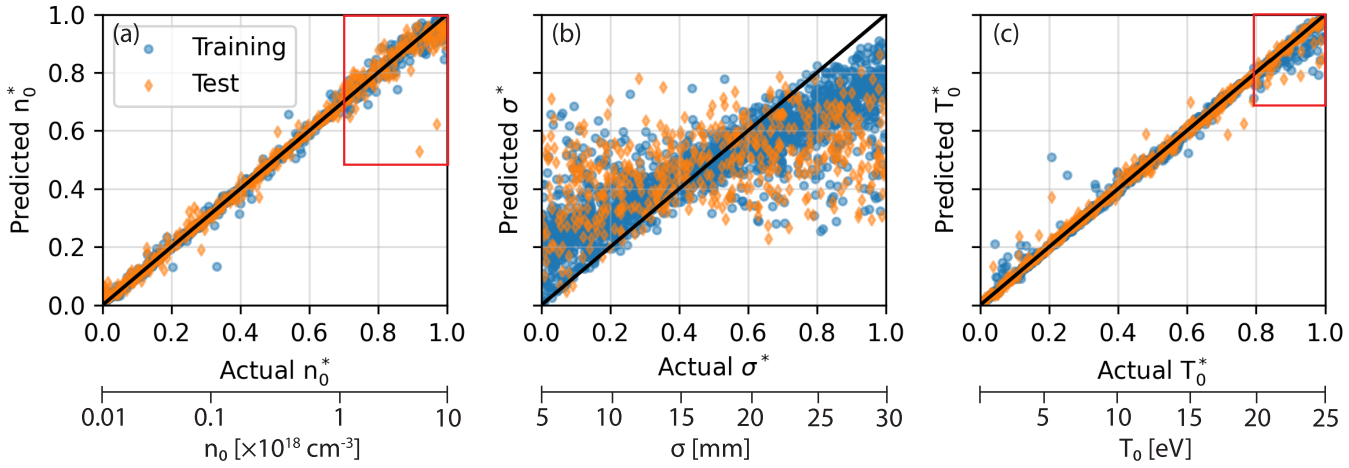

Fig. 7. Distribution of predicted values of (a) n_0^*, (b) σ^*, and (c) T_0^* against the actual values for the AutoGluon Model. Here, the values of peak density, temperature, and the spread of the density function σ have been normalized. Blue points compare the values for the training set, while orange points do so for the test set.

randomly selected member of the test set. For this intensity spectrum, the actual values of density, temperature, and σ are 1.4×10^{18} cm^{-3}, 7.3 eV, and 25 mm, and the predicted values are 1.5×10^{18} cm^{-3}, 6.9 eV, and 22 mm. We approximate the gradients in density and temperature ($\partial n_0^*/\partial\epsilon$ and $\partial T_0^*/\partial\epsilon$) by perturbing the input intensity spectrum at a given wavelength by a small value ϵ, and dividing the change in the predicted value ($|\delta n_0^*|$ or $|\delta T_0^*|$) by the perturbation ϵ. Here, we pick the perturbation ϵ from a Gaussian distribution of amplitude 0.1, and the mean gradients are calculated over multiple iterations to determine the final value. Such sensitivity maps are commonly used in image classification problems to identify parts of an image that may contribute to the final classification [36]. Here, we use it to identify features of the spectra that contribute to the final predictions in our regression problem. In Fig. 8, as expected, the lines (Al-III 415.2 nm, Al-III 448.1 nm, Al-III 452.5 nm, and Al-III 570.7 nm) contribute significantly to the prediction of density and temperature, whereas parts of the spectrum that correspond to the continuum are less important. This can be observed from the large gradients $\partial n_0^*/\partial\epsilon$ and $\partial T_0^*/\partial\epsilon$ at the positions of these lines. The Al-III 570.7-nm line appears to be particularly important for the temperature prediction, while the Al-III 448.1-nm and Al-III 452.5-nm lines (which have merged at this higher density) are relatively more important to the density prediction. In addition to the locations of the Al-III lines, smaller peaks in the gradients, particularly in temperature, also appear at parts of the spectrum devoid of lines. The locations of these peaks correspond to Al-II emission lines that appear predominantly at lower temperatures (0.5–1.5 eV), indicating that the absence of these lines contributes, although less significantly, to the temperature prediction.

Another limitation of these models is the reliance on synthetic data, which, in turn, is affected by the uncertainty and assumptions in the theoretical modeling used to generate the synthetic dataset. Benchmarking the results using independent diagnostics can be one way to probe the applicability of these ML models to real experimental data. However, as mentioned earlier, independent measurements are not always possible,

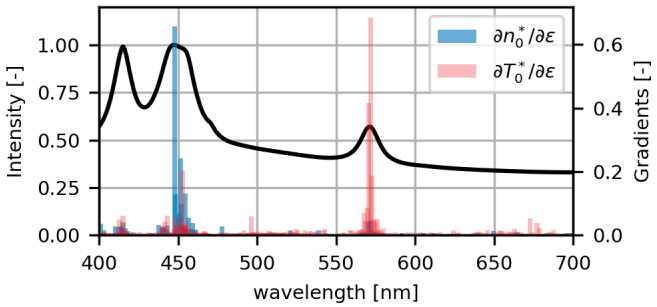

Fig. 8. Sensitivity of the predictions of density and temperature to variations in the spectral intensity value at different wavelengths for a randomly chosen intensity spectrum from the test set.

and the uncertainty in collisional-radiative modeling is inherent in the analysis of most spectroscopic data. As discussed in Section IV-A, the experimental data can also be noisy, and include background radiation and contamination by impurities. When we use the trained ML models to make predictions from noisy spectra contaminated with stray radiation, the R^2-scores typically fall by 10%–20%. For a given experiment, the instrument response, the bit depth of the spectrometer, and attenuation by the optics may also need to be properly included in the synthetic dataset. The analysis of experimental spectra using these methods will be pursued in a future publication.

V. Conclusion

We explore the use of ML methods for rapid spectroscopic analysis of emission spectra in the visible regime. Our goal is to predict density and temperature in a pulsed-power-driven aluminum plasma generated by an exploding wire array. In contrast to previous work, which has typically focused on low-density homogenous plasmas, we aim to diagnose a high-energy density nonoptically thin plasma, which necessitates the use of radiation transport calculations to accurately model the recorded intensity spectrum. These radiation transport calculations use spectral emissivity and opacity values computed using the collisional-radiative code

PrismSPECT. Consistent with previous observations in exploding wire arrays, we assume a constant temperature and a Gaussian density variation (peak density n_0 and standard deviation σ) along the diagnostic LOS.

The radiation transport solver is first used to directly solve a single-objective optimization problem, which varies the values of n_0, T_0, and σ to minimize the MSD between a generated spectrum and the target spectrum. This approach provides reliable fits to the target spectrum, robust to noise and stray line radiation; however, it can be time intensive, which limits its usefulness for rapid or real-time analysis of large datasets.

We then use the radiation transport solver to generate a [2500 $\times$ 400] dataset of synthetic emission spectra, and we compare the performance of different ML models on their ability to predict density, temperature, and σ from the given intensity spectra. The AutoGluon model performs best, with an R^2-score of roughly 98% for density and temperature predictions. Simpler models (RF, KNN, and DNN) also exhibit high R^2-scores (>90%) for density and temperature predictions, showing the potential of ML models in providing rapid and accurate analysis of spectral data. However, the prediction of σ is relatively poor, which is typical of radiation transport problems, and is related to the relatively smaller sensitivity of the optical depth on this parameter, when compared with the peak density n_0. The performance of the AutoGluon model was also compared with single-objective optimization by making predictions on a smaller subset of the test set. AutoGluon demonstrated a performance comparable to GA-based curve fitting, showing similar MSD and R^2-scores, but requiring a fraction of the time for evaluation.

REFERENCES

[1] H. R. Griem, *Principles of Plasma Spectroscopy*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[2] U. Fantz, "Basics of plasma spectroscopy," *Plasma Sources Sci. Technol.*, vol. 15, no. 4, pp. 137–147, Nov. 2006.

[3] X.-M. Zhu and Y.-K. Pu, "Optical emission spectroscopy in low-temperature plasmas containing argon and nitrogen: Determination of the electron temperature and density by the line-ratio method," *J. Phys. D: Appl. Phys.*, vol. 43, no. 40, Oct. 2010, Art. no. 403001.

[4] I. H. Hutchinson, "Principles of plasma diagnostics," *Plasma Phys. Control. Fusion*, vol. 44, no. 12, p. 2603, 2002.

[5] G. A. Rochau, J. E. Bailey, and Y. Maron, "Applied spectroscopy in pulsed power plasmas," *Phys. Plasmas*, vol. 17, no. 5, May 2010, Art. no. 055501.

[6] M. R. Gomez et al., "Magnetic field measurements via visible spectroscopy on the Z machine," *Rev. Sci. Instrum.*, vol. 85, no. 11, Nov. 2014, Art. no. 11E609.

[7] H. R. Griem, "Plasma spectroscopy in inertial confinement fusion and soft X-ray laser research," *Phys. Fluids B: Plasma Phys.*, vol. 4, no. 7, pp. 2346–2361, Jul. 1992.

[8] S. Sudkewer, "Spectroscopic diagnostics (don't short) of tokamak plasmas," *Phys. Scripta*, vol. 23, no. 2, pp. 72–86, Feb. 1981.

[9] D. Astic, A. Canosa, P. Vigliano, and M. Autric, "Visible spectroscopy of laser produced aluminum plasma. Electron temperature and density determinations," in *Proc. Int. Congr. Appl. Lasers Electro-Opt.*, 1991, pp. 94–103.

[10] J. T. Knudtson, W. B. Green, and D. G. Sutton, "The UV-visible spectroscopy of laser-produced aluminum plasmas," *J. Appl. Phys.*, vol. 61, no. 10, pp. 4771–4780, May 1987.

[11] J. E. Bailey et al., "Diagnosis of X-ray heated Mg/Fe opacity research plasmas," *Rev. Sci. Instrum.*, vol. 79, no. 11, Nov. 2008, Art. no. 113104.

[12] J. MacFarlane, I. Golovkin, P. Woodruff, S. Kulkarni, and I. Hall, "Simulation of plasma ionization and spectral properties with prismspect," in *Proc. Abstr. IEEE Int. Conf. Plasma Sci. (ICOPS)*, Jun. 2013, p. 1.

[13] R. Drake, *High-Energy-Density Physics*. Berlin, Germany: Springer, 2013.

[14] S. V. Lebedev, A. Frank, and D. D. Ryutov, "Exploring astrophysics-relevant magnetohydrodynamics with pulsed-power laboratory facilities," *Rev. Modern Phys.*, vol. 91, no. 2, Apr. 2019, Art. no. 025002.

[15] K. Bell et al., "Optical spectroscopy experiments on the 500 Ka XP pulsed-power generator," in *Proc. AIP Conf.*, vol. 1, 2009, pp. 161–163.

[16] R. Datta et al., "The structure of 3-D collisional magnetized bow shocks in pulsed-power-driven plasma flows," *J. Plasma Phys.*, vol. 88, no. 6, Dec. 2022, Art. no. 905880604.

[17] D. R. Russell et al., "Perpendicular subcritical shock structure in a collisional plasma experiment," *Phys. Rev. Lett.*, vol. 129, no. 22, Nov. 2022, Art. no. 225001.

[18] R. Datta et al., "Plasma flows during the ablation stage of an over-massed pulsed-power-driven exploding planar wire array," *Phys. Plasmas*, vol. 30, no. 9, Sep. 2023, Art. no. 092104.

[19] R. Datta et al., "Plasmoid formation and strong radiative cooling in a driven magnetic reconnection experiment," 2024, *arXiv:2401.04643*.

[20] R. Datta et al., "Simulations of radiatively cooled magnetic reconnection driven by pulsed power," 2024, *arXiv:2401.01795*.

[21] J.-H. Park, J.-H. Cho, J.-S. Yoon, and J.-H. Song, "Machine learning prediction of electron density and temperature from optical emission spectroscopy in nitrogen plasma," *Coatings*, vol. 11, no. 10, p. 1221, Oct. 2021.

[22] T. Shah Mansouri, H. Wang, D. Mariotti, and P. Maguire, "Methane detection to 1 ppm using machine learning analysis of atmospheric pressure plasma optical emission spectra," *J. Phys. D: Appl. Phys.*, vol. 55, no. 22, Jun. 2022, Art. no. 225205.

[23] K. Shojaei and L. Mangolini, "Application of machine learning for the estimation of electron energy distribution from optical emission spectra," *J. Phys. D: Appl. Phys.*, vol. 54, no. 26, Jul. 2021, Art. no. 265202.

[24] D. J. Clayton et al., "Electron temperature profile reconstructions from multi-energy SXR measurements using neural networks," *Plasma Phys. Controlled Fusion*, vol. 55, no. 9, Sep. 2013, Art. no. 095015.

[25] S. H. Lee, J. H. Lee, I. Yamada, and J. S. Park, "Development of a neural network technique for KSTAR Thomson scattering diagnostics," *Rev. Scientific Instrum.*, vol. 87, no. 11, p. 11, Nov. 2016.

[26] C. M. Samuell, A. G. Mclean, C. A. Johnson, F. Glass, and A. E. Jaervinen, "Measuring the electron temperature and identifying plasma detachment using machine learning and spectroscopy," *Rev. Sci. Instrum.*, vol. 92, no. 4, Apr. 2021, Art. no. 043520.

[27] A. Ciardi et al., "The evolution of magnetic tower jets in the laboratory," *Phys. Plasmas*, vol. 14, no. 5, May 2007, Art. no. 056501.

[28] G. F. Swadling et al., "Oblique shock structures formed during the ablation phase of aluminium wire array Z-pinches," *Phys. Plasmas*, vol. 20, no. 2, Feb. 2013, Art. no. 022705.

[29] S. Mirjalili, "Genetic algorithm," in *Evolutionary Algorithms and Neural Networks* (Studies in Computational Intelligence), vol. 780. Cham, Switzerland: Springer, 2019, doi: 10.1007/978-3-319-93025-1_4.

[30] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*, vol. 112. Berlin, Germany: Springer, 2013.

[31] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Nov. 2011.

[32] M. Abadi et al., "TensorFlow: A system for large-scale machine learning," in *Proc. USENIX OSDI*, 2016, pp. 265–283.

[33] J. A. Anderson, *An Introduction to Neural Networks*. Cambridge, MA, USA: MIT Press, 1995.

[34] N. Erickson et al., "AutoGluon-tabular: Robust and accurate AutoML for structured data," 2020, *arXiv:2003.06505*.

[35] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Disc. Data Mining*, 2019, pp. 2623–2631.

[36] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 618–626.