PAPER

Data-driven prediction of the output composition of an atmospheric pressure plasma jet

IOP Publishing Journal of Physics D: Applied Physics
J. Phys. D: Appl. Phys. **57** (2024) 015203 (17pp) https://doi.org/10.1088/1361-6463/acfcc7

Data-driven prediction of the output composition of an atmospheric pressure plasma jet

Li Lin[1,*], Sophia Gershman[2], Yevgeny Raitses[2] and Michael Keidar[1,*]

[1] School of Engineering and Applied Science, The George Washington University, Washington, DC, United States of America
[2] Princeton Plasma Physics Laboratory, Princeton, NJ, United States of America

E-mail: lilin@gwu.edu and keidar@gwu.edu

Received 10 May 2023, revised 6 September 2023
Accepted for publication 25 September 2023
Published 9 October 2023

Abstract

Cold atmospheric plasma (CAP) in open air hosts numerous chemical species engaged in thousands of chemical reactions. Comprehensive diagnosis of its chemical composition is important across various fields from medicine, where reactive oxygen and nitrogen play key roles, to surface modification. In applications, a centimeter-scale helium–air jet operates for minutes, featuring micrometer-sized streamers and an atmospheric pressure-induced collision frequency in the hundreds of GHz range. To address this intricate multi-scale issue, we introduce a machine learning approach: using a physics-informed neural network (PINN) to tackle the multi-scale complexities inherent in predicting the complete list of species concentrations, gas temperature, and electron temperature of a CAP jet supplied with a mixture of helium and air. Experimental measurements of O_3, N_2O, and NO_2 concentrations downstream of the plasma jet, combined with fundamental physics laws, the conservation of mass and charge, constrain the PINN, enabling it to predict the concentrations of all species that are not available from the experiment, along with gas and electron temperatures. The results, therefore, obey all the physical laws we provided and can have a chemical balance with the measured concentrations. This methodology holds promise for describing and potentially regulating complex systems with limited experimental datasets.

Supplementary material for this article is available online

Keywords: physics-informed neural network, multi-scale problem, plasma chemistry, cold atmospheric plasma, physics-informed data-driven modeling, deep learning, evolutionary algorithm

1. Introduction

Cold atmospheric plasma (CAP) jet produces a low-temperature plasma (LTP) at atmospheric pressure, characterized by nonequilibrium between the gas temperature and the electron temperature [1, 2]. Gas is at a low enough temperature to avoid any thermal damage to objects in direct contact with the jet, while the electrons have high enough energy to cause electron-impact processes. Electron-impact reactions trigger complex LTP chemistry leading to the production of reactive oxygen and nitrogen species (RONS). Exposure to plasma RONS can lead to outcomes as diverse as cell apoptosis, improved tissue regeneration, inactivation of microorganisms, and improved seed germination [1, 3]. Therefore, CAP is a versatile plasma source currently used for

* Authors to whom any correspondence should be addressed.

agricultural, biomedical, and many other applications broadly based on the interaction of the plasma with a variety of substrates. Understanding the interactions of the CAP plume with physical and biological substrates starts with the characterization of the composition and plasma parameters in the plume. CAP jet effluent composition depends on the plasma gas, the surrounding gas composition (e.g. the humidity of the ambient air or the type and flow of a shielding gas), the type of substrate, etc. A detailed description of kINPen, a well-studied example of a CAP jet, can be found in a topical review by Reuter *et al* [4]. The importance of the understanding and standardization of jet performance for consistent results in surface modification and biomedical applications is addressed in [5, 6].

The plasma parameters of a free CAP jet depend on the input power, the input gas flow, and composition along with the nozzle geometry to determine the gas flow field [7]. Several well-developed plasma diagnostics have been used to quantify the species concentrations in CAP. For example, laser-induced fluorescence (LIF) provides the concentration of selected species with a very high resolution of nanoseconds and millimeters [8]. However, the implementation of LIF is usually limited to determining the concentrations of one species at a time, and an extension of the method to several species requires expensive equipment. Imaging optical emission spectroscopy (OES) can reach millimeter and tens of ns resolution under certain circumstances. It is a passive method that does not disturb the plasma but OES provides information on the excitation of certain species and under certain conditions, plasma parameters, but it does not yield concentrations of species in ground states. Fourier-transform infrared absorption spectroscopy (FTIR-AS) can provide absolute concentrations of ground-state species that have dipole active vibrational states for solids, liquids, and gases [9]. Many of the species in air plasmas, such as O_3, H_2O, CO_2, CO, NO_2, N_2O, NO, HNO_3, HNO_2, and N_2O_5, are detectable using FTIR-AS [10, 11]. The advantage of FTIR in addition to measuring a broad spectrum with high resolution in a single sweep is low sensitivity to plasma emission compared to the direct absorption techniques. The drawback is that most measurements are conducted ex-situ for atmospheric pressure plasmas often due to a lack of optical access, such as for dielectric barrier discharge (DBD), or due to a lack of a well-defined optical path as for open jets. Ex-situ FTIR-AS analysis of gas composition is used broadly from studies of plasma jets to the applications of DBD in plasma agriculture [12, 13], while in-situ studies are rare and require plasma reactors with optical access [14, 15].

These diagnostic methods cannot offer a complete picture of the required species concentrations. Therefore, they may not provide sufficient information for understanding biochemical reactions or quantifying the effective dose of CAP therapy and its associated side effects [16–18]. Kinetic simulations can often give access to the parameters that are not measurable in experiments. However, the inelastic collisions in chemical interactions are microscopic and occur on a nanosecond scale in a simulation, while the gas flows are macroscopic on a millisecond time scale for flow rates of meters per second in a centimeter-size CAP. A CAP axial streamer propagation and DBD contains multiple filaments developed between the plate electrodes [19–22]. On the other hand, the streamer propagation time is 1 μs corresponding to the time it takes the pulse of high electron temperature region to sweep along the plasma plume; while the remaining 10–100 μs in the discharge period correspond to the electron energy dissipation stage [23, 24]. Therefore, the simulation of CAP also requires a fine time step length to describe the streamer pulse smoothly while completing the discharge period and repeating thousands of the discharge period for several milliseconds. A simulation covering several discharge periods was completed overcoming such a multi-scale problem by keeping the simulation running up to the 30th discharge period, with a nano or even pico-second time step for plasma chemistry and a microsecond time step for the macroscopic fluid dynamics [25]. This work gave highly informative results such as the accumulation of species when the discharge frequency is high, and the gas flow rate is low [25]. However, it is not clear if the simulation reached the steady state since experimentally CAP jets have been observed to take seconds and orders of magnitude more individual discharge events to reach the steady state [26]. The steady-state concentrations are important in real-world applications, but it takes several seconds or even minutes with thousands of discharge periods for the CAP jet to reach its steady state [27]. This means a significant time cost to simulate the entire time needed to reach the steady state numerically unachievable.

Machine learning (ML) techniques with neural networks can help to solve such a problem. Recently, an ML analysis was developed that used the data from OES to generate computationally the concentrations of the relevant species in CAP in real time [28]. This ML method relied on a 0D chemical simulation to acquire the error of each iteration to update the neural network weights during the training. This means that the output concentrations correspond to the plasma plume and are consistent with the plasma optical emission spectrum [29, 30]. Also, ML methods, especially those using neural networks, can control and automatically optimize the plasma generator performance [31–33]. In the last decade a new technique namely ‘physics-informed data-driven modeling’ was developed especially the one using a physics-informed neural network (PINN) [34]. PINN is a deep neural network that contains multiple hidden layers, and the fitting capability of a fully connected neural network is proportional to its number of layers, in other words, how deep it is. Therefore, to solve physics problems, PINN with multiple layers is designed to mimic highly non-linear partial differential equation systems [35], and even the notoriously complicated ones, such as the fluid dynamics systems with the Navier–Stokes equations [36, 37]. The advantage of PINN is to handle multi-physics problems in complex geometry without meshing and reduce the computational cost [38]. Currently, PINN has been developed to solve physical problems not only with the constraints of physical

laws but also with the users feeding experimental data [38–40]. This enables the PINNs to give more realistic results and mimic the actual engineering systems because pure constraints of physical laws cannot fully represent unexpected flaws and disturbances in real hardware.

The goal of this work is to predict the overall gas output composition of the CAP jet using a PINN, therefore, providing information about the chemical composition associated with plasma exposure. We hypothesize that a scheme based on ML using PINN will be able to predict the species concentrations in the jet effluent based on quantitative analysis of operando and in situ FTIR measurements. In addition to the concentrations, the ML model will return reasonable values for the electron temperature and density. This approach is particularly useful when experimental data is difficult or expensive to obtain and traditional simulation-based methods are too computationally intensive.

2. Methods

2.1. The main concept of using PINN to solve a multi-scale plasma chemistry

The idea of this work is to skip the time-consuming simulation and let the artificial neural network guess the parameters of CAP at the steady state condition, including the electron temperature and the species concentrations of all the species (as defined below). Several constraints are established to quantify the error of the neural network output. These constraints are:

(1) All possible chemical reactions should occur among these species, but the concentrations should not change over time because the entire system is in a steady state with a chemical balance. The concentration of species should agree with the experimentally measured (FTIR) concentrations.
(2) All the concentrations predicted by PINN must agree with conservation laws: conservation of chemical elements and charge.
(3) If multiple solutions (combinations of species concentration) exist to meet all of the above constraints, the one with the highest total entropy should be the most stable state while others are metastable states.

The error is used to update the network weights and to train the network until the error is converged to an acceptable low value. The initial guess of a neural network with its initial weights is usually inaccurate, but after the training, the guesswork is expected to be accurate and to agree with the experimental observations and with the other constraints. Therefore, this provides a method to determine the species concentrations in the CAP jet effluent.

2.2. The experimental setup

The first step is to experimentally collect the physical information on the CAP jet system, including electrical characteristics and FTIR absorption spectra. We attach a helium–air plasma jet to an FTIR gas cell positioned in the FTIR sample compartment as shown in figure 1. The details of the plasma jet generator design can be found in multiple previous publications [26, 27, 41]. The plasma jet nozzle is 4 mm in diameter with a pair of electrodes installed to create streamer propagation along the helium–air flow field. The pin electrode is in the nozzle and 5 cm upstream from the lower ring electrode which is 0.5 cm wide and located 0.5 cm above the nozzle exit. The voltage applied between the electrodes is sinusoidal, 12.5 kHz with an amplitude of 8 kV, 9.5 kV, 11 kV, or 12.5 kV. The gas is a mixture of helium and air at a combined flow rate of 3 LPM with the helium–air ratio selected from 80%, 85%, 90%, and 95%. Therefore, FTIR spectra are collected for 20 experimental conditions. The FTIR spectrometer is Jasco 6800. A 10 cm optical path, flow-through gas cell with KBr windows confines the jet effluent to a fixed path length but does not prevent the gas flow. Therefore a steady state species concentration is established in the cell at each experimental condition. FTIR spectra allow the detection of multiple species, including CO_2, CO, H_2O, N_2O, NO_2, and O_3, to name the species corresponding to the strongest peaks. Three gaseous species are selected for training the PINN from the measured output from the jet: N_2O, O_3, and NO_2, because the concentrations of these species are least dependent on the environmental conditions of the measurements and depend only on the jet effluent. FTIR results provide absolute concentrations of these species under 20 conditions.

Therefore, each species has a 2D result: one dimension for the Helium–air ratio, and the other for the discharge voltage. In the view of ML, these are 20 pieces of data for the training. However, 20 pieces of data are insufficient. Therefore, a data set preparation procedure is applied, which contains the following steps: (1) applying a 2D linear interpolation on both dimensions to make fine data. (2) saving the maximum and the minimum concentration for each species; (3) applying a moving average with an averaging width at two data points in both dimensions for each matrix to smooth the curvature; (4) scaling the resulting matrices with the maximum and minimum concentration values saved in step (2) to recover the absolute value loss during the moving average. The resulting data contains 19 discharge voltage values between 8 kV and 12.5 kV and 21 helium–air ratio values between 80% and 95%, resulting in a data space containing 304 pieces of data, each data piece is a 1D array of six concentration values of six species (concentration constraints) under one of the 304 conditions. The 304 data are then randomly shuffled and 54 data (about 17.76%) are saved as the validation data set, 50 data (about 16.45%) are saved as the testing data set, and 200 data (about 65.79%) will be used as the training data set.

2.3. The architecture of the neural network

Next, as shown in figure 2, an artificial neural network is designed to map the experimental information space and the species concentration space that is the concentrations of all the species under all the possible conditions. Such a mathematical

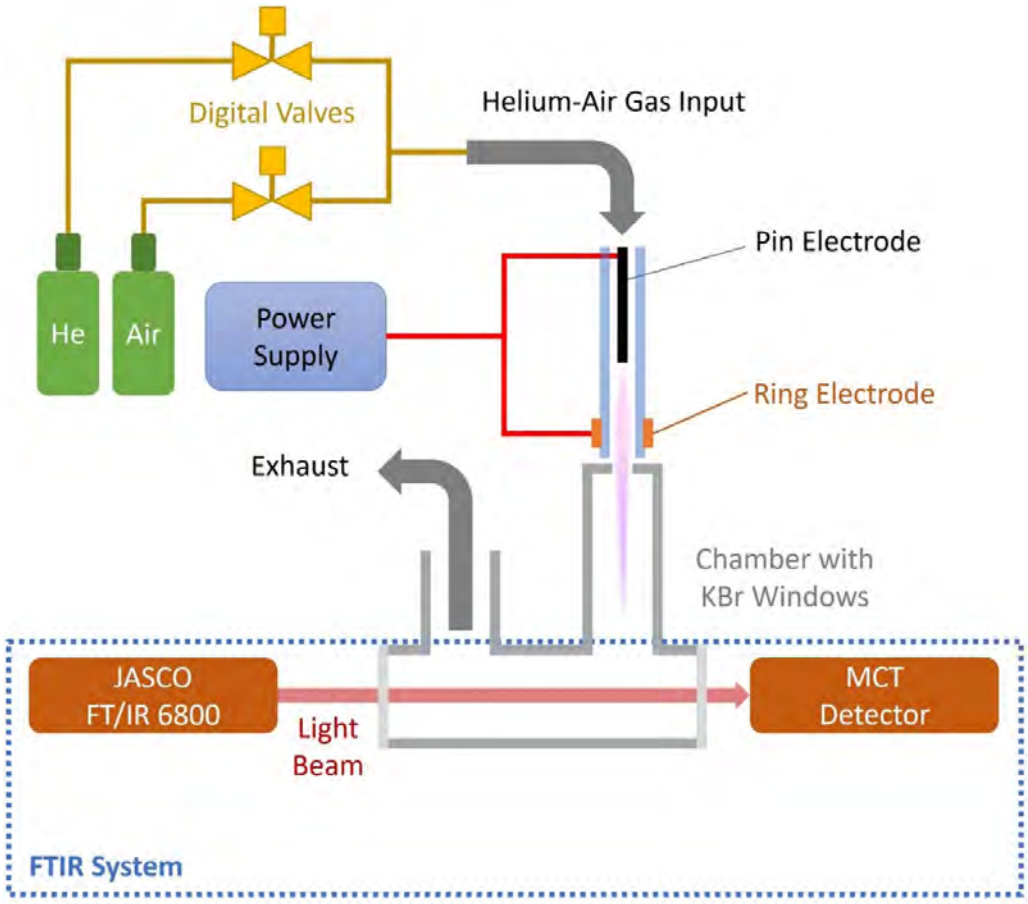

Figure 1. The schematic of FTIR measurement of species concentrations. This is the experiment providing the raw FTIR results for the data set preparation.

mapping between the two vector spaces is thus a data-driven modeled PINN that can solve the multi-scale problem: output the desired species concentration list in a steady-state CAP jet, when inputting the CAP jet setup.

The PINN is a fully connected network containing five hidden layers between the input and output layers. The activation function for all the neurons is Leaky ReLU with the leak constant set to 0.01 for negative inputs. The neural number in each layer is 6, 50, 200, 300, 300, 200, and 149. The input layer is a 1D array randomly selected from the data set which is a working condition as a specific pair of discharge voltage and the helium–air ratio. These two values are represented by the first two neurons. The next three neuron values are the three species concentrations as the concentration constraints. The last neuron is a dummy which is always −10. The dummy neuron must be added to improve the neural network performance as a negative value source. Also, since the input values are not on the same scale, the discharge voltage is around 10 kV, the helium–air ratio is less than 1, and the species concentrations can vary from 0.1 ppm to hundreds of ppm, these numbers are multiplied by scaling constants before the input as shown in purple text in figure 2. The scaling constants will ensure that all the input neurons can have equal initial effects on the next layer and thus increase the training speed of the network, although the entire algorithm still works without the scaling constants.

The top 147 neurons in the output layer are the concentration values of all the species other than He, N_2, and O_2, which are determined by the helium-air ratio, and the three measured concentrations, N_2O, O_3, and NO_2. The 148th neuron is the electron temperature and the last one is the gas temperature under the input working condition. However, the output of the neural network is purely a mathematical process and the output layer neurons are all at the same scale rather than the concentration values and the temperatures with direct physical meanings and at different scales. For example, one can have an electron density of about 10^{12} (cm^{-3}), but the electron temperature is less than 10 (eV). In the view of a neural network, 10^{12} and 10 are numbers too different to handle. The neural network, therefore, outputs a set of numbers around 0.05 instead of directly outputting the values at very different scales. These numbers around 0.05 are translated into the power values of 10 to represent the species' concentrations using equations 1(a)–(c). Also, we have equations 1(d) and (e) to translate the last two neurons into the electron temperature in (eV) and the gas temperature in (K)

$$A_i' = A_i - \frac{\sum_i A_i}{145} \tag{1a}$$

$$A_i'' = \frac{5\,(A_i' - 0.02)}{0.08 - 0.02} + 9 \tag{1b}$$

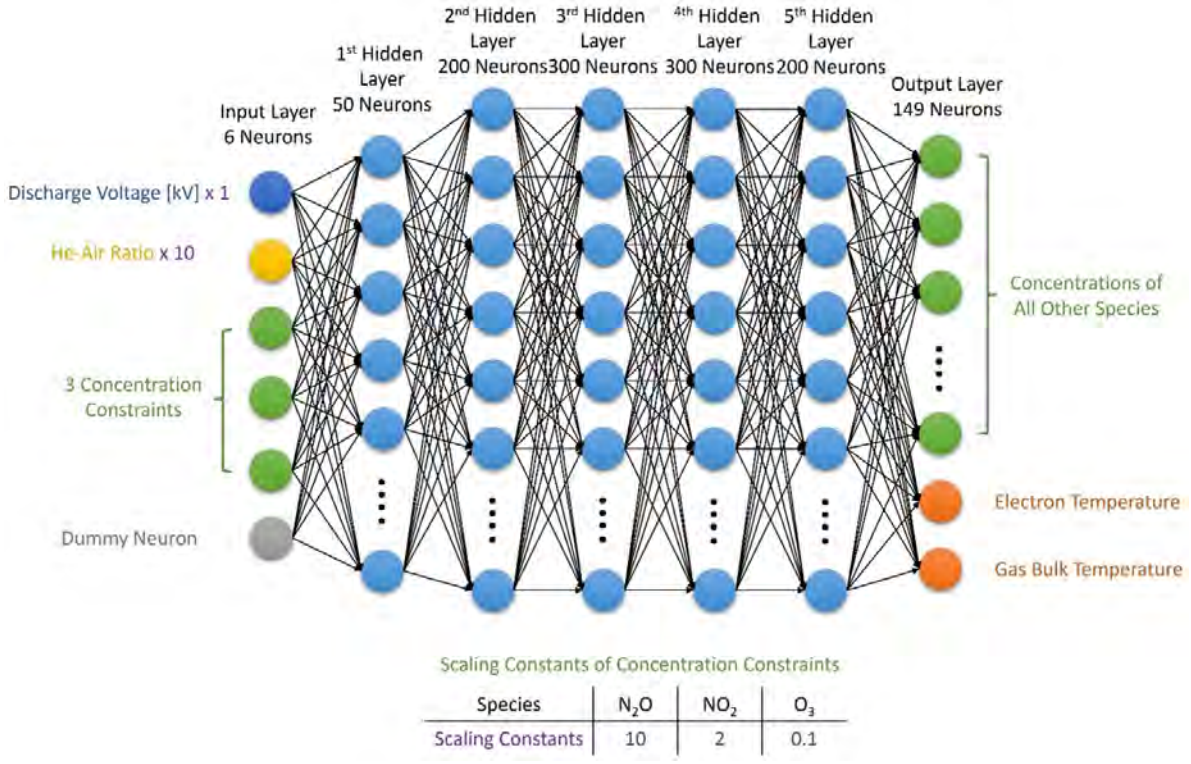

Figure 2. The fully connected PINN includes five hidden layers. The input layer is data randomly selected from the input space which is the physics-informed data set. The network is thus a data-driven model to provide the electron temperature and the concentrations of all the species other than those measured ones.

$$n_i = 10^{A_i''} \ [\mathrm{cm}^{-3}] \ \ i \leqslant 147 \tag{1c}$$

$$T_\mathrm{e} = \max\left[\frac{(A''_{148} - 9)}{1.5}, \frac{T_\mathrm{g} k_\mathrm{B}}{e}\right] [\mathrm{eV}] \tag{1d}$$

$$T_\mathrm{g} = \max\left(30A''_{149}, 273 + 20\right) [\mathrm{K}] \tag{1e}$$

where A_i is the ith neuron value in the output layer, A_i' is the first intermediate parameter representing the data distribution of all the output neurons around their average value, A_i'' is the second intermediate parameter normalized and shifted to a range between 9 and 15. The ith species concentration n_i in cm^{-3} can thus take the A_i'' as a power of 10. The electron temperature T_e is determined by the 148th value of A'' but no lower than the gas temperature T_g. The gas temperature is determined by the 149th parameter and is not lower than 293 K. In equation (1*d*), k_B and e are Boltzmann constant and unit charge respectively.

2.4. The training technique for the neural network

The training flow chart of the neural network is summarized in figure 3. The training technique is an evolutionary algorithm. First, a group of data is randomly picked from the training data set. The total amount of data in the group is a constant, namely the 'minibatch size' N_{mb}. In this work, $N_{mb} = 50$. The minibatch loop is thus a loop from $n_{mb} = 1$ to $n_{mb} = N_{mb}$. At the beginning of each minibatch loop, the n_{mb}^{th} data is selected and fed to a group of networks. Each group contains X original versions and additional $N-1$ mutated version. A mutated version is a network with weights equal to the weights from the original version adding a set of small random numbers. The mutation is mathematically defined as

$$w_{i,j,k} = w_{i,j,k} + \zeta\left(\delta - 0.5\right) w_{i,j,k} \tag{2a}$$

$$\lambda = \left[\frac{\varepsilon_2 - \varepsilon_1}{\max\left(\varepsilon_1, \varepsilon_2\right)}\right] - 0.01 \tag{2b}$$

$$\chi = \begin{cases} 0 \ if \ \lambda \leqslant 0 \\ 1 \ if \ \lambda > 0 \end{cases} \tag{2c}$$

$$\zeta = \zeta - 10^{-5}\chi \tag{2d}$$

where $w_{i,j,k}$ is the weight connecting the ith neuron in the kth layer and the jth neuron in the $(k+1)$th layer, ζ is the mutation range which is similar to the learning rate that controls the speed of the error convergence and stability, δ is a random number between 0 and 1, ε_1 and ε_2 are the average error of the last 2000th to the last 1000th iteration and the last 1000th to the latest iteration. The mutation range ζ is set to 0.01 initially and will decrease by 10^{-5} in each iteration after the 2000th iteration when the error decrement is slow.

In each training iteration, a total number of $X \times N$ networks are tested by feeding them N_{mb} number of data. This means that each network will provide N_{mb} number of prediction results and also N_{mb} number of errors. The user will evaluate the

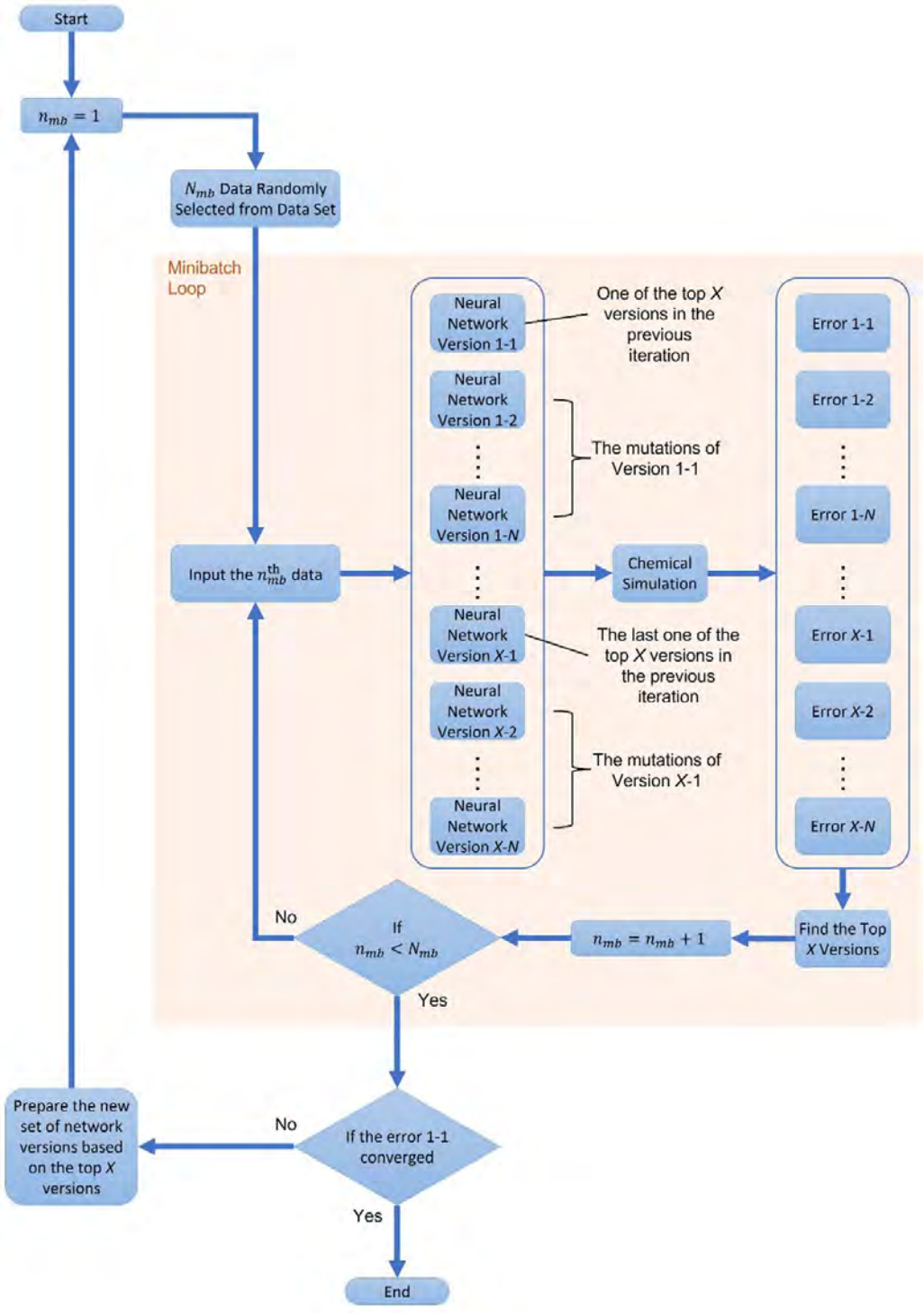

Figure 3. Using the evolutionary algorithm to train the PINN using the training data set with a minibatch internal loop.

performance of each network by considering the average error of each network. At the end of each iteration, the top X best (that can make lower errors) network versions will be saved as the original versions for the next iteration, and all other networks will be deleted. This is an evolutionary algorithm in that only the best versions are allowed to generate offspring in the next iteration. The user will thus see the network evolve during the training and finally acquire a high-quality neural network that can make an accurate prediction. Finally, once the error of the top 1 best version has converged, the entire training is ended. At this point, a (pre-selected) testing set will be applied to verify the training quality as shown in figure 4. When the PINN can always provide a low error, no matter what data (CAP working condition and the concentration constraints from the data sets) is fed, we can consider the PINN on the concentrations of all other species and the electron

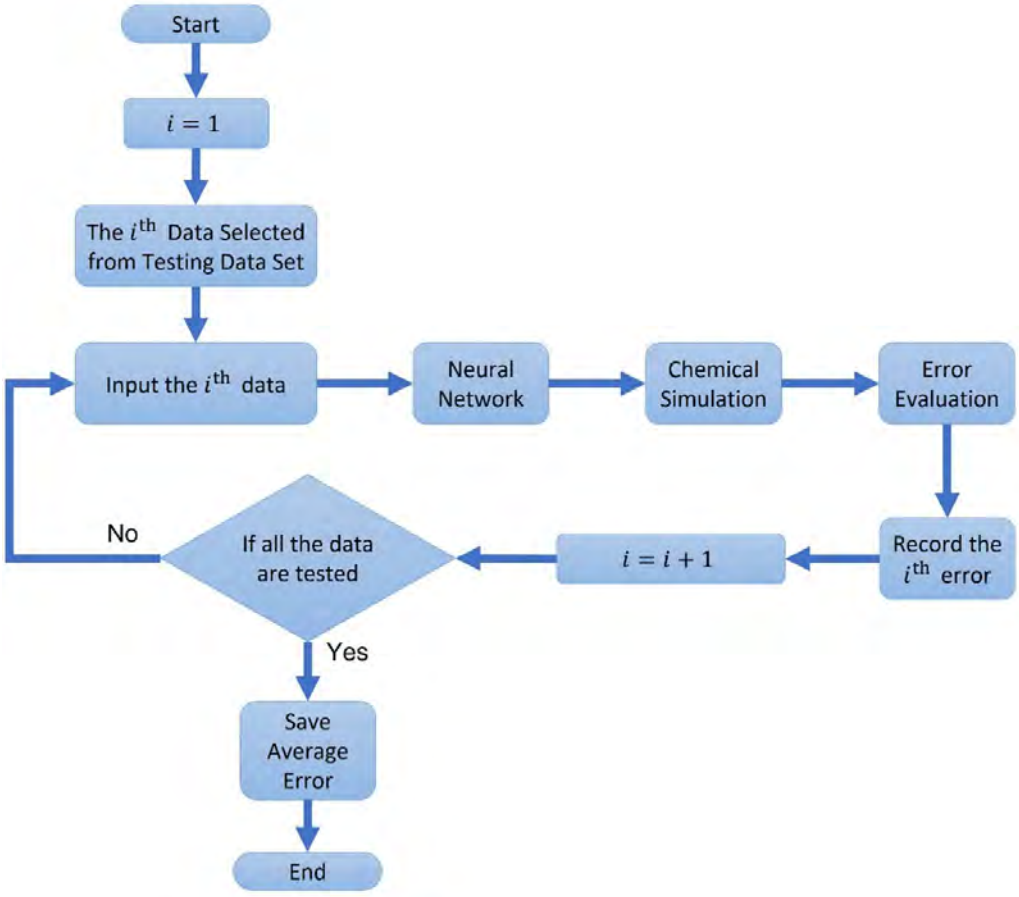

Figure 4. Using the 'Testing Data Set' to verify the training error.

temperature are an accurate solution of the problem that can output all the species with no significant variation over time (steady state) in the CAP plume. The results will be discussed in the next chapter. The MATLAB codes of this work are available at https://mpnl.seas.gwu.edu/open-codes/.

2.5. The loss function of the training

To quantify the error of each prediction of each version of the network after an input is fed, a loss function must be defined as the rule for error evaluation. The loss function of this work is designed as

$$J_{total} = J + \frac{W_S}{\sum_i S_i n_i} + W_o J_o \tag{3}$$

where J_{total} is the total error value of each network output (a 1D matrix of species concentrations and temperatures), $W_S = 10^{-7}$ as the weight to balance the priority of ML to handle the entropy along with the chemical imbalance error J, S_i is the entropy value of the ith species in ($JK^{-1}\,cm^{-3}$), n_i is the concentration of the ith species, J_o is the penalty for other bad behaviors of ML such as the violation of conservation of chemical element (mass) and the conservation of charge.

The first term on the right-hand side of equation (3) is the error of species concentration variations at steady state. This is to test if a PINN prediction together with the experimentally observed species concentrations can achieve a chemical balance. To do so, an output along with the experimental results will be sent to a 0D chemical simulation.

The simulation is designed to solve a set of rate equations:

$$\frac{\partial n_i}{\partial t} = \sum_p \left(k_p \prod_q n_{p,q}^{\Delta} \right) \tag{4}$$

where n_i is the concentration of the ith species, t is the time, k_p is the rate coefficient of the pth chemical reaction, and $n_{p,q}^{\Delta}$ is the concentration of the qth reactant in the pth reaction with a reaction order Δ. The simulation will run for 100 ns with $dt = 10$ ns. The kinetic scheme of all the chemical reactions and their rate coefficients can be found in the supplemental information. The simulation will thus provide the temporal development of all the species concentrations. An error evaluation process will check each temporally resolved species concentration acquired from the simulation for each $X \times N$ prediction. The first part is defined as

$$J = \sum_i \overline{n_i(t)}^{-1} \left[\max(n_i) - \min(n_i) \right] \tag{5}$$

where J is the error value of the chemical unbalance, $\overline{n_i(t)}$ is the average value of the ith temporally resolved species

concentration, and t is the time. The value J represents the scale of the variation of a species concatenation $\max(n_i) - \min(n_i)$ compared with its average. A balanced chemical system should have all the concentrations of species at a steady state, in other words, all the species concentrations should have a small variation that $\max(n_i) - \min(n_i)$ compared with its average. Once the minibatch loop is finished, the top X versions with the lowest total error values are collected, and they will be the X original versions to make the mutation.

However, a chemical system usually has multiple steady states or even an infinite number of steady states. The existing state of CAP is merely one of all the possible steady states. Therefore, the algorithm must ensure that the PINN can give a real steady state. In reality, this depends on the 2nd law of thermodynamics. A system without external interferences always trends to increase its total entropy. Therefore, unless the CAP users try to decrease the entropy in the CAP plume, the actual steady state of the CAP must have the highest entropy compared with other possible steady states. The mathematically possible solutions with lower entropy are merely metastable states.

In equation (3), the entropy of formation for most of the ions and molecules can be found in the NIST database. However, the entropy of free electrons, excited species, and photons are not in the database and must be computed. For example, let us consider the reaction $A^* \rightarrow A + h\upsilon$, where h is the Planck's constant and υ is the frequency of the photon. The system loses $h\upsilon$ because of the photon leaving. Therefore, the net energy change is:

$$E = -h\upsilon = -\frac{hc}{\lambda} \tag{6}$$

where c is the light speed and λ is the wavelength of the photon. Assuming the system is a canonical ensemble, we have the partition function:

$$Z = \exp\left(-\frac{E}{k_{\mathrm{B}}T_{\mathrm{e}}}\right) = \exp\left(\frac{hc}{\lambda k_{\mathrm{B}}T_{\mathrm{e}}}\right). \tag{7}$$

The temperature used in the thermodynamic beta term $(k_{\mathrm{B}}T_{\mathrm{e}})^{-1}$ is electron temperature because we assume that all the excitations of molecules are due to the electron impacts. These electron impacts are the energy source to produce photons. Next, the relationship between entropy and partition function, yields:

$$F = -k_{\mathrm{B}}T_{\mathrm{g}}\ln(Z) \tag{8a}$$

$$S_{h\upsilon} = -\frac{\partial F}{\partial T_{\mathrm{g}}} \tag{8b}$$

where F is the Helmholtz free energy, and $S_{h\upsilon}$ is the entropy of the photon. Using the Avogadro number $\aleph_A$ to convert the unit from (J K^{-1} per particle) to (J $(\mathrm{K\ mol})^{-1}$), we finally have:

$$S_{h\nu} = \frac{ch}{\lambda T_{\mathrm{e}}}\aleph_A. \tag{9}$$

Similarly, one can also derive the entropy of the excited species as

$$S_{EX} = \frac{cz_{EX}h}{T_{\mathrm{e}}}\aleph_A \tag{10}$$

where z_{EX} is the energy difference between two energy levels, the unit is (cm^{-1}), representing the excitation state. For the electrons, note that their energy source is the electric field, and their temperature is higher than the gas temperature. This gives the net energy change, and using the similar derivation above, we finally have:

$$S_{\mathrm{e}} = \frac{k_{\mathrm{B}}T_{\mathrm{e}} - k_{\mathrm{B}}T_{\mathrm{g}}}{T_{\mathrm{e}}}\aleph_A \tag{11}$$

where S_e is the entropy of electrons as a function of electron temperature.

The 3rd term on the right-hand side of equation (3) represents the error of breaking the conservation laws. Conservation of chemical elements (equivalent to mass conservation) and charge conservation are considered in this work. The mathematical expression for J_o is

$$J_0 = \sum_i \sum_j \mathcal{N}_{i,j}\left|n_j' - n_j\right| + \sum_j Z_j n_j' \tag{12}$$

where i is the index of the ith chemical element, j is the index of the jth chemical species, $\mathcal{N}_{i,j}$ is thus the number of the ith chemical element in the j$^{\mathrm{th}}$ species, n_j' is the concentration of jth species provided by the PINN, n_j is the concentration of the jth species in the initial gas supply, and Z_j is the charge of the jth species. For example, if the PINN suggests that the concentration of atomic O in the results is higher than the initial O_2 we provided in the air input, or if the resulting composition is significantly charged, a penalty error value in the loss function will be added. The weight W_o is 1000 for J_o to ensure the balance among error terms.

3. Results and discussions

Figure 5 shows the results of FTIR measurements of three species: N_2O, NO_2, and O_3. The original FTIR results are shown on the left, and the smoothed fine data sets used for ML are on the right. The measured concentrations vary with the discharge voltage and helium-air ratio as seen in figure 5. The concentration of N_2O is in the range of 1–3 ppm (±15% and signal/noise ≈5), and roughly proportional to the discharge voltage as shown in color. However, there is no significant relation between the NO_2 concentration and the discharge voltage, but it is proportional to the helium–air ratio. The highest measured NO_2 concentration is about 10 ppm (±20% and signal/noise ≈10) when the helium is 95%, although NO_2 concentration is expected to decrease when the helium-air ratio is close to 100% due to the lack of air supply. The FTIR short cell–MCT detector combination has a detection limit in the single ppm range making it difficult to identify significant trends in the N_2O and NO_2 concentrations. On the other

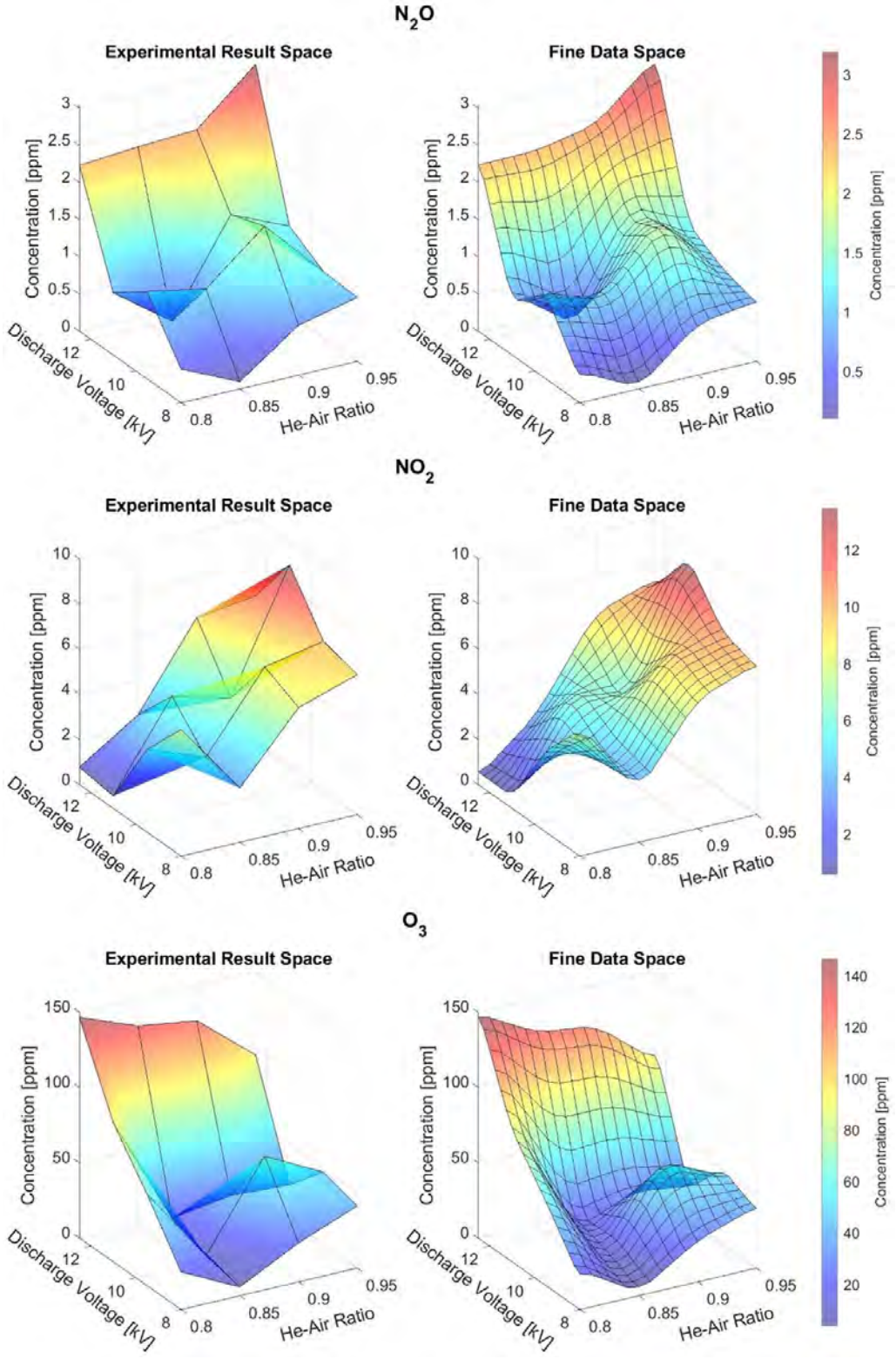

Figure 5. The experimental result space and fine data space of N_2O, NO_2, and O_3.

hand, O_3 has a clear trend of concentration proportional to the discharge voltage of the plasma. As mentioned in the previous section, to increase the data set available for training, the resulting space is smoothed using a moving average and linear interpolation that lead to a fine data space shown by the graphs on the right in figure 5. These smoothed curves enable network training with more detailed data space and avoid overfitting. The training results are discussed below.

The total error value (loss function) J_{total} during the training is shown in black in figure 6(a), where the red curve is the

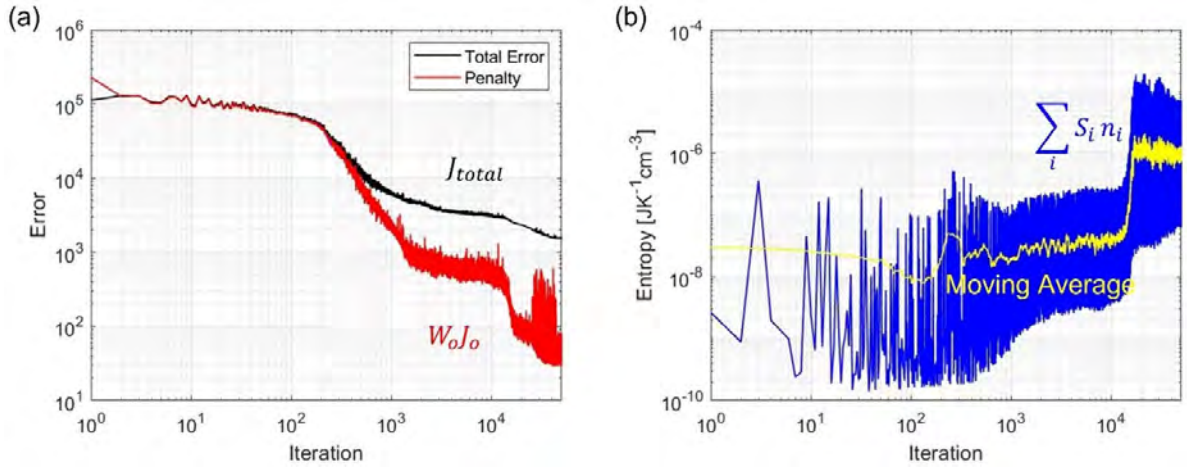

Figure 6. The training results. (a) The error convergence of the 1st version of the neural network during the training. In each iteration, the error value (black) shown is the sum of the neural network's error of the entire minibatch loop including the penalty value (red). (b) The entropy values (blue) during the training and its moving average (yellow).

value of W_oJ_o which represents the error penalty due to the violation of physics laws. Note that each of the values shown in figure 6(a) is a sum of 50 (the minibatch size) cases randomly selected from the training data set. The training stopped at the 50 000th iteration because the error converged. The total error is about 1550 (the end of the black curve in figure 6(a)) and the penalty is about 20 (the end of the red curve in figure 6(a)). This means that the error of breaking the steady state is about $\frac{(1550-50)}{50} = 30$ for each case. For a better understanding of this quantity of error, let us assume that each species has its own error, and all the species contribute to the error equally. This means that for a total of 154 species, each species concentration contributes an average error of $\frac{30}{154} = 0.1948$ After 100 time steps of the chemical simulation, each species concentration provided by the PINN has a temporal variation such that the difference between the maximum and the minimum is about 19.48% of its average value. However, most of the steady-state error is due to the species of low concentration which are thus less important to the system, but making significant errors due to the low value of the denominator in equation (2). After removing the species less than 0.01 (ppm) from the computation of error, the error of breaking the steady state will be about 11.89, meaning that the difference between the maximum and minimum is about 7.72% of its average value. Since the total entropy of the species of each case is about 1×10^{-6} (J K^{-1} cm^{-3}) shown in figure 6(b), the penalty of violation of element conservation is about $\frac{50}{50} - \frac{10^{-7}}{1\times10^{-6}} = 0.9$ for each case. Therefore, the total violation of the conservation law is about $J_0 = \frac{0.9}{W_0} = 9 \times 10^{-4}$ (ppm) which is an acceptable low error value considering that most of the species in the jet effluent have concentrations of around 1 ppm, such as H_2O_2 and OH that are important for cancer therapy [42]. The final weight matrices are also tested using the testing data set as discussed previously (figure 4). The average total error is about 2955, close to the value of 1550 at the end of training. This means that the training is not overfitting. In other words,

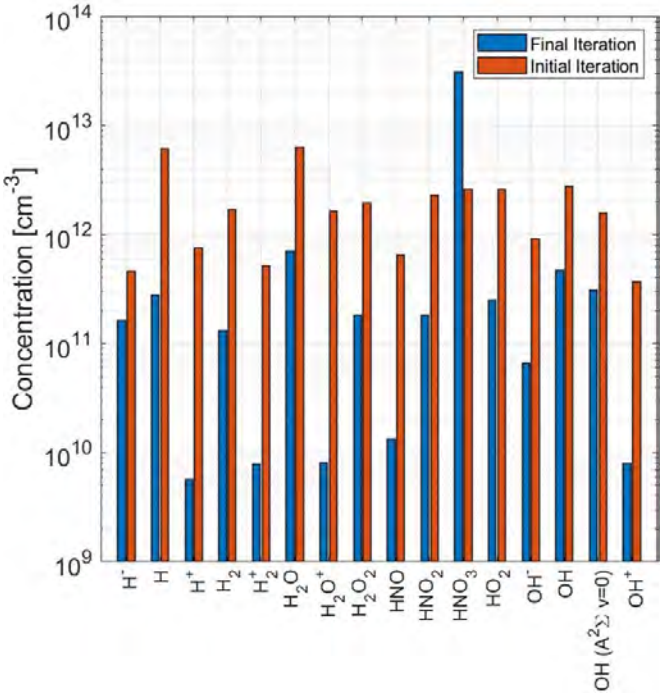

Figure 7. The PINN predicted concentrations of hydrogen species. The bar plots are the average values of all the cases with the variation of discharge voltage and the helium–air ratio and the error bars indicate the distribution of concentrations of these variations.

the ML is not limited by the training set and it can perform well even when the input data is not contained in the training set, as indicated by the testing set results.

Figures 7–11 show the maps of all the species concentrations provided by the PINN for the first iteration and the final iteration. These figures allow us to visualize what the PINN did to match the species concentrations to achieve a steady state within the limits of constraints. The bar value

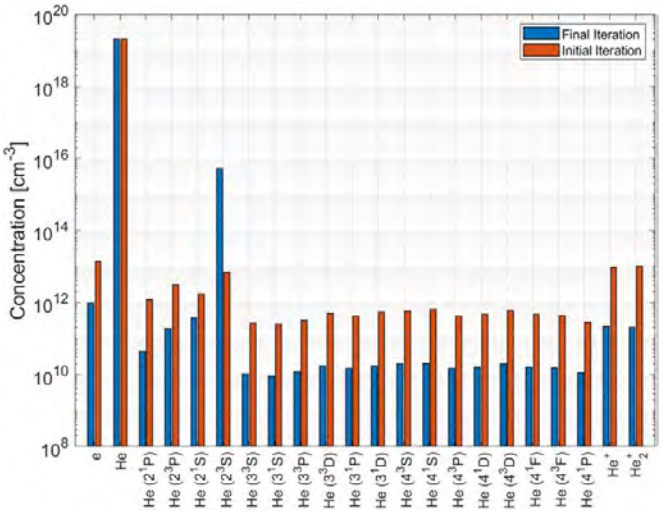

Figure 8. The PINN predicted concentrations of electron and helium species. The bar plots are the average values of all the cases with the variation of discharge voltage and the helium–air ratio and the error bars indicate the distribution of concentrations of these variations.

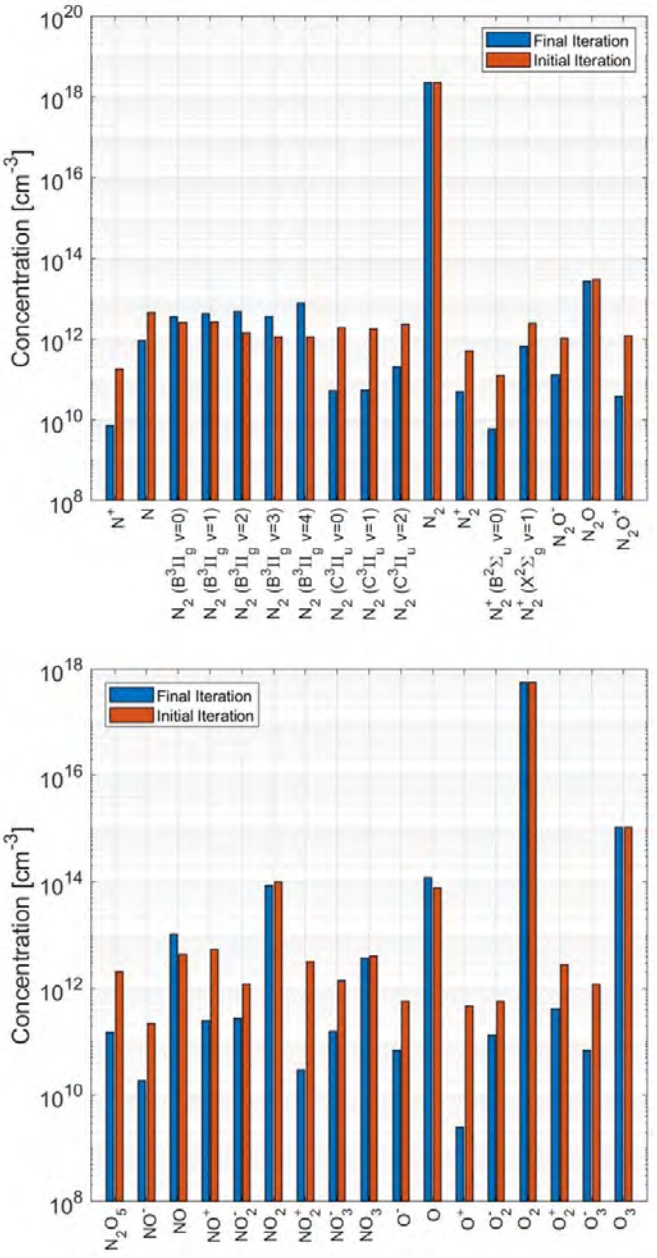

Figure 9. The PINN predicted concentrations of RONS. The bar plots are the average values of all the cases with the variation of discharge voltage and the helium–air ratio and the error bars indicate the distribution of concentrations of these variations.

of each species is the average concentration of all the discharge voltage and helium–air ratio conditions, while the error bars in these figures indicate the variation among these conditions. The concentrations of hydrogen-related species are summarized in figure 7. First, in figure 7, the PINN suggests a high concentration of HNO_3. This is not surprising under such high discharge voltage setups, and HNO_3 has a long lifetime to survive downstream of the plume [43]. Usually, in the gas phase, there are N_xO_y that react with H_2O directly to produce nitric acid but mostly that happens in the liquid phase, not in the gas phase. PINN also suggests another major conversion of the H_2O (the source of the H) is OH and HO_2. These significant species agree with the common sense in the last decade of CAP studies [44–47]. However, the first iteration (the one without training) suggests roughly uniform concentrations as shown in orange bars in figure 7. This means that such a concentration combination cannot reach a low error as suggested in figure 6(a). After the training, the blue bars are the concentration combination found by the ML to make the chemical system obey the constraints listed at the beginning of the Method section. Similarly, the concentrations of the electron and helium species are summarized in figure 8. The PINN suggests relatively high concentrations of both helium metastable states: He(2^1S) and He(2^3S). It is not surprising that the metastable state with a long lifetime can have high concentrations downstream in the plume. PINN model discovered this feature by itself. The PINN also suggests that the electron density is about 10^{12} cm^{-3} downstream of the plume, where the plume should have a slightly higher electron density as measured using Rayleigh microwave scattering [22, 27, 41].

Figure 9 summarizes the concentrations of RONS other than the OH family. While the values of NO_2, N_2O, and O_3 are given by the FTIR, PINN suggests a high concentration of O, NO, and N_2(B) are mostly above around 10^{12} cm^{-3} which agrees with other experimental observations and is common for pulsed plasma oxidation [48–51]. The N_2(B) series are the results of the second positive system emissions which provide the most significant OES peaks such as the well-known 337.13 nm and 357.69 nm [28]. N_2 is not IR active and is thus not available from FTIR measurement, but the concentration of N_2(B) can be quantified using PINN. These photons can be found in figure 10 where the after-training values (blue) of them are all higher than the before ones (orange).

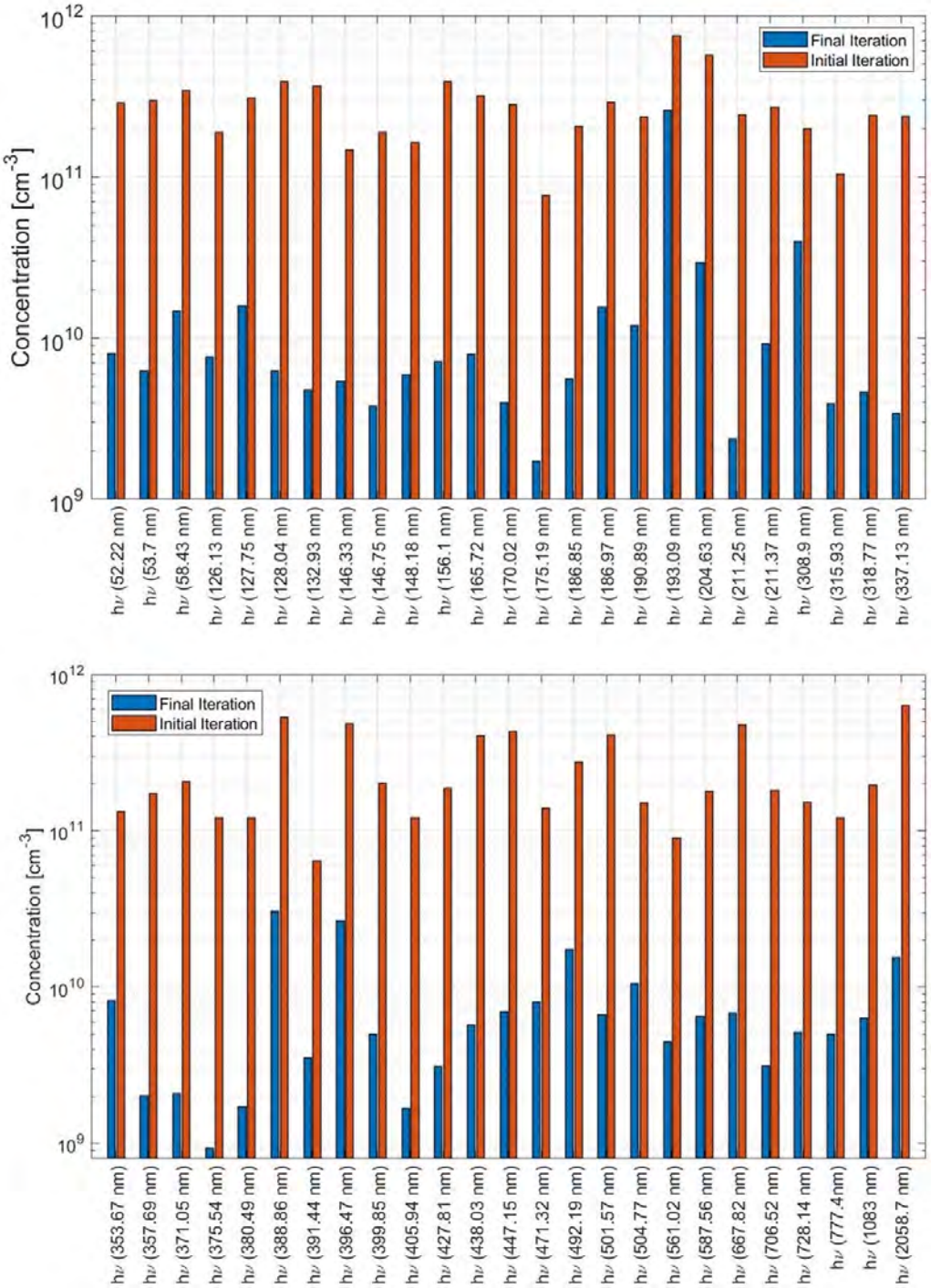

Figure 10. The PINN predicted concentrations of photons. The bar plots are the average values of all the cases with the variation of discharge voltage and the helium-air ratio and the error bars indicate the distribution of concentrations of these variations.

In figure 10, we summarized the photon numbers, and all the wavelengths are at very low concentrations. The highest peak is the 193.09 nm UV at about 10^{11} cm^{-3} (0.01 ppm). This agrees with the common sense that the downstream area below the plume is invisible to the naked eye. Figure 11 summarizes the PINN's suggestions of carbon series species. There are some atomic carbons dissociated from the CO_2 in the air, but the amounts are much lower than CO and CO_2 because the dissociation requires high electron energy. There are about 400 ppm of CO_2 and 0.1 ppm of CO in the ambient air, and these are the carbon sources of the low-temperature atmospheric plasma. In this work, there is a small amount of CO_2 entered into the plasma jet system, providing the carbon element.

Within the range of our experiments, the discharge voltage and helium–air ratio only slightly affect the species concentrations, electron temperature T_e and gas temperature T_g are shown in figure 12. The value of T_e is at about 1.1–1.7 eV

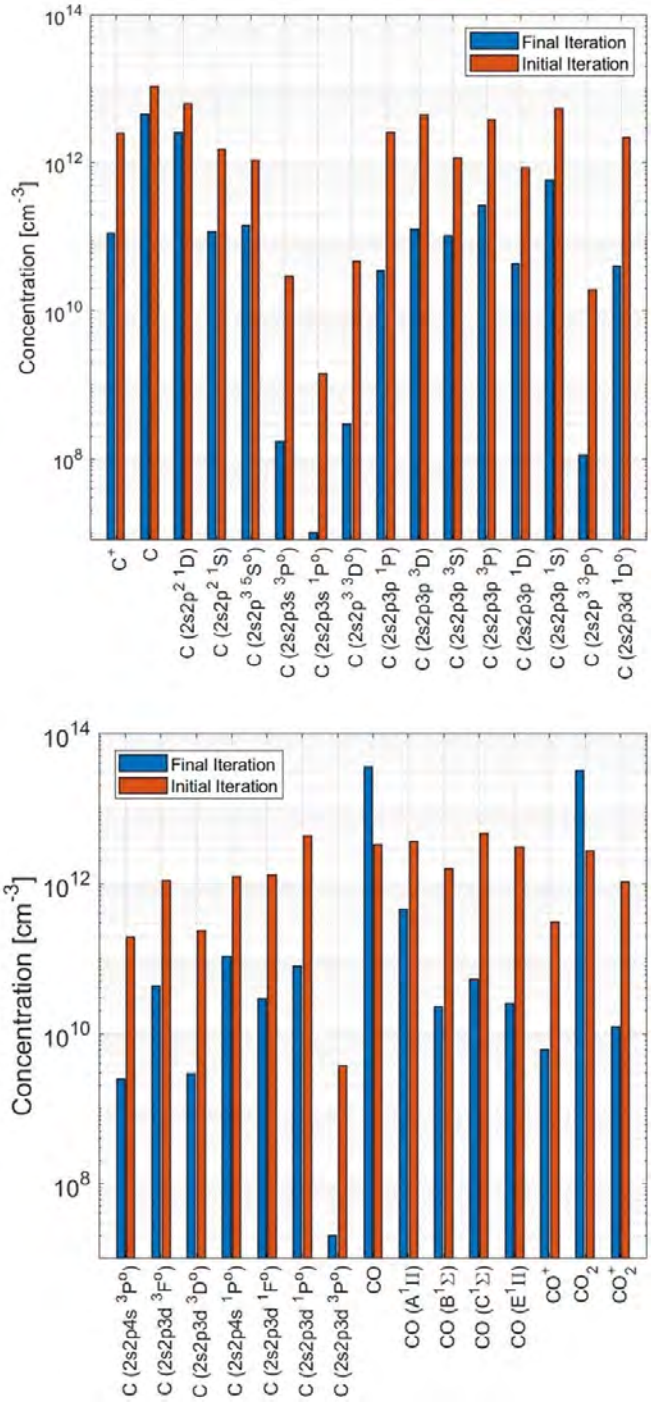

Figure 11. The PINN predicted concentrations of carbon species. The bar plots are the average values of all the cases with the variation of discharge voltage and the helium–air ratio and the error bars indicate the distribution of concentrations of these variations.

while the one of T_g is at about 315–340 K. The T_g is about 15–40 K above room temperature, which is higher than expected in a CAP jet [52]. However, the discharge voltage of this work, above 8 kV amplitude, is much higher than the usual setups that are about 4 kV amplitude or 8 kV pk–pk sinusoidal leading to 300–330 K. Some other works also report such a relatively high gas temperature [53]. In the visible region upstream, the thermal energy of high T_g is acquired from the T_e through the momentum transfer during the electron-impact collisions. The downstream temperature should be lower than the temperature upstream. The nonmonotonic relation of the gas temperature versus the discharge voltage and helium–air ratio can also be found in other publications [54]. Also, more importantly, the downstream plasma chemistry has no external energy input; therefore, its chemical reactions should be different from the plasma chemistry in the visible plume region. For example, we should have very low ionization rates but relatively higher recombination rates in the dark downstream. The

species concentrations we introduced above are thus generated in the visible region and traveled downstream.

In this work, we have 153 species as mentioned above. All these species concentrations can have a contour figure to show their values as functions of the discharge voltage and He–air ratio, while the bars shown from figures 7–11 are their average values. However, due to the limit of paper space, we cannot show all the contour plots but only the most important ones and those with high concentrations as summarized in figure 13. The first row of subplots in figure 13 contains electron, He(2^3S) and He^+. As expected, their concentrations are proportional to both the discharge voltage and the He–air ratio. In figure 13, all other species shown in the 2nd, 3rd, and 4th rows require air to be generated, but they share a general trend that these species concentrations are also proportional to the He–air ratio. This agrees with many publications that these species' concentrations can be increased when adding air in the helium but will then decrease when adding too much because the oxygen can decrease the electron density with the attachment reaction [55–59]. However, since they need air to be generated, when the helium is close to 100%, their concentrations should be decreased. It is also interesting that the concentrations of the short-lifetime O, NO, and the long-lifetime NO_3 have different contour patterns in figure 13. However, the chemical pathway network is complicated, and investigation of what exact combination of reactions is causing this is beyond this work. We should leave this topic open and will be studying it in the future.

It is also interesting to compare the species concentration results with CAP applications, such as the OH production in this work and the CAP-based cancer treatment which relies on the H_2O_2 in the liquid cell-culture medium [52]. The nonmonotonic relations of these RONS production with gas flow rate and discharge voltage are used in many biomedical, environmental, and material processing applications [52, 60–64]. These findings of nonmonotonic relations also agree with other works, however, this work provides the full picture of all the species concentrations [53, 65, 66].

4. Conclusions

The physics-informed data-driven model presented here predicts the full picture of the species concentrations at a location the downstream of CAP plume based on a limited experimental data set. The method makes the PINN guesswork agree with the concentrations of N_2O, NO_2, and O_3, measured by FTIR-AS, obey conservation laws, and the chemical balance at the steady state of the system. The PINN output included the electron density, the electron temperature, and the gas temperature in addition to the concentrations of molecular and atomic species and their electronically excited states. The training was validated using a part of the data set of the three species concentrations mentioned above but the PINN prediction results were compared to literature instead. Since these parameters are experimentally available from OES and other measurements, in the future experimental data should be used to validate this model. Finally, FTIR data could provide concentrations of H_2O_2, CO_2, CO, NHO_3, and other species providing better environmental control and experimental system. Shall we add all the available parameters into the constraints, or use some of them to verify the PINN's prediction? What is the best way of using the PINN technique? These are open topics that warrant future studies.

This new approach overcomes the difficulties and the expense of solving a multiscale problem of plasma chemistry while providing a reasonable guess of the concentrations within the given constraints. It helps in situations when experimental measurements can only provide limited species information, and simulations suffer high computational costs when simulating the macroscopic scale with nanosecond time steps. However, this work addressed the multi-scale problem with respect to time only, focusing on a single location in space. In the future work, we expect to use spatially resolved diagnostics allowing both the spatial and temporal resolution to cross the scale.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Acknowledgments

The work of L L and M K is funded by the Department of Energy of the US, Grant DE-SC0022349. Dr Carles Corbella Roca's initial work on this project is greatly appreciated. The work of SG was supported by the Princeton Collaborative Research Facility (PCRF), which is supported by the Department of Energy under Contract No. DE-AC02-09CH11466. Plasma diagnostic resources used in this work were provided by the PCRF.

ORCID iDs

Li Lin https://orcid.org/0000-0003-0176-8858
Sophia Gershman https://orcid.org/0000-0002-8409-4029
Yevgeny Raitses https://orcid.org/0000-0002-9382-9963
Michael Keidar https://orcid.org/0000-0003-0869-4310

References

[1] Laroussi M *et al* 2022 Low-temperature plasma for biology, hygiene, and medicine: perspective and roadmap *IEEE Trans. Radiat. Plasma Med. Sci.* **6** 127–57
[2] Keidar M, Yan D and Lin L 2023 *Cold Atmospheric Plasma-based Cancer Therapy* 2nd edn (IOP Science) (available at: https://iopscience.iop.org/book/mono/978-0-7503-5537-7)
[3] Yan D, Malyavko A, Qihui W, Li L, Jonathan H S and Michael K 2021 Cold atmospheric plasma cancer treatment, a critical review *Appl. Sci.* **11** 7757
[4] Reuter S, Von Woedtke T and Weltmann K-D 2018 The kINPen—a review on physics and chemistry of the atmospheric pressure plasma jet and its applications *J. Phys. D: Appl. Phys.* **51** 233001