

Review: Machine learning for advancing low-temperature plasma modeling and simulation

Jan Trieschmann^{Ⓜ, a, b, *}, Luca Vialetto^{Ⓜ, a, c} and Tobias Gergs^{Ⓜ, a, d}

^aKiel University, Theoretical Electrical Engineering, Department of Electrical and Information Engineering, Kiel, Germany

^bKiel University, Kiel Nano, Surface and Interface Science KiNSIS, Kiel, Germany

^cStanford University, Department of Aeronautics and Astronautics, Stanford, California, United States

^dRuhr University Bochum, Chair of Applied Electrodynamics and Plasma Technology, Department of Electrical Engineering and Information Science, Bochum, Germany

ABSTRACT. Machine learning has had an enormous impact in many scientific disciplines. It has also attracted significant interest in the field of low-temperature plasma (LTP) modeling and simulation in past years. Its application should be carefully assessed in general, but many aspects of plasma modeling and simulation have benefited substantially from recent developments within the field of machine learning and data-driven modeling. In this survey, we approach two main objectives: (a) we review the state-of-the-art, focusing on approaches to LTP modeling and simulation. By dividing our survey into plasma physics, plasma chemistry, plasma–surface interactions, and plasma process control, we aim to extensively discuss relevant examples from literature. (b) We provide a perspective of potential advances to plasma science and technology. We specifically elaborate on advances possibly enabled by adaptation from other scientific disciplines. We argue that not only the known unknowns but also unknown unknowns may be discovered due to the inherent propensity of data-driven methods to spotlight hidden patterns in data.

© 2023 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JMM.22.4.041504](https://doi.org/10.1117/1.JMM.22.4.041504)]

Keywords: artificial intelligence; simulations; neural networks; plasmas

Paper 23050SSV received Jun. 30, 2023; revised Dec. 14, 2023; accepted Dec. 14, 2023; published Dec. 28, 2023.

1 Introduction

Low-temperature plasmas (LTPs) consist of a mixture of neutral and charged species in thermal non-equilibrium. The governing discharge physics and chemistry lead to a zoo of (excited) species interacting with bounding surfaces. Their modification (e.g., activation, etching, functionalization) is exploited in the frame of plasma processing (e.g., plasma-enhanced etching, deposition, catalysis). LTPs also render a dominant part of the many steps required in semiconductor device manufacturing. This may involve a diverse portfolio of plasma discharges, chemical precursors, surface materials, and process recipe steps to achieve desired deposit/etch patterns.^{1,2} It may as well include indirect plasma discharges, for example, in the generation of extreme ultraviolet light for state of the art 13.5 nm wavelength lithography, required for device feature scales of the order of 10 nm and beyond (for instance, termed 3 nm nodes).³ Modeling and numerical simulation contribute an integral component in the design and optimization of related plasma processes. In the recent past decade, data-driven methods such as machine learning (ML), deep learning (DL), and artificial neural networks (ANNs), among other

*Address all correspondence to Jan Trieschmann, jt@tf.uni-kiel.de

computational statistics methods, have (re-)gained enormous interest—which may be subsumed under the umbrella term artificial intelligence (AI). Although these fields have seen a renaissance and have continued to rapidly develop in the past decade, many fundamental concepts have been well established in computational statistics and related disciplines. The past and present interest in the context of plasma processing may be attributed to the potential for hidden pattern detection in correlated data and, particularly, the efficiency in related optimization tasks.^{4–12}

Although we limit our scope to modeling and simulation in the following, it should be stressed that ML methods and ANNs have been widely employed in the experimental studies (e.g., process regression and plasma diagnostics). For instance, as early as in 1992, they have been explored for parameter estimation in plasma etching based on optical emission spectroscopy (OES) and mass spectroscopy^{13,14} or global process parameters paired with measured etch characteristics.¹⁵ Plasma virtual metrology (VM) was conducted based on multivariate sensor data,¹⁶ with regard to real-time fault detection in reactive ion etching,¹⁷ batch process characterization in semiconductor fabrication,¹⁸ and a deep learning VM framework based on OES data¹⁹ and “plasma information” descriptors.²⁰ Further examples have devised an inverse reconstruction of intrinsic plasma properties, such as the electron energy distribution function from OES diagnostics data²¹ as well as an active learning guided scheme based on Fourier transform infrared spectroscopy data for parameter space exploration.²²

In contrast to these previous examples, which rely on the experimental and diagnostics data, the focus of this paper will be on data-driven aspects of plasma modeling and simulation. We consider how these methods can complement and/or replace theoretical approaches, including low and atmospheric pressure LTPs. Specifically, in Sec. 2, we initially attempt to provide a state-of-the-art review of data-driven LTP modeling. This specifically covers plasma physics, plasma chemistry, plasma–surface interactions (PSIs), as well as process control and design aspects. In Sec. 3, we further relate this to the closely linked disciplines of plasma fusion research, and to more general connections in data science and fluid dynamics research. In this context, an assessment and perspective of future developments and opportunities in LTP research is devised. Overall, the main focus is not only on DL methods but also computational statistics methods, such as Gaussian process regression (GPR), are outlined. Finally in Sec. 4, concluding remarks with a perspective of the current state of the art in adjacent disciplines and the foreseen requirements, potential applications, and future developments of ML in plasma modeling and simulation in the next years 5–10 years is drawn.

2 Review

A general perspective on the current state of data-driven LTP science—among many other aspects—has been laid out in Ref. 9. Despite the ample momentum that data science and ML have recently gained in LTP, the potential of a community effort for systematic data harvesting and exploration to initiate a form of “plasma informatics” has been identified therein. Kambara et al.¹² further detailed this paradigm with a comprehensive summary of science-based, data-driven developments. A wide range of plasma processing technologies has been addressed, whereas a number of exemplary use cases of data-driven plasma modeling and simulation will be revisited later. Moreover, in an overarching effort, the broad scope of the field of data-driven plasma science has been reviewed.¹⁰ Among other aspects, the fundamentals of data science and its application to space and astronomical plasmas, fusion plasmas, and LTPs have been elaborated. Owing to its broad review scope, LTP processing and its applications have only been generally assessed. In contrast, Bonzanini et al.¹¹ elaborated in more detail on the fundamentals of ML for LTPs. Therein, the authors not only provide a comprehensive introduction into the concepts and a collection of ML methods with relevant links to the literature but also underpin their theoretical foundations with an instructive set of applied LTP case studies.

In an attempt to bridge the gap between the mentioned fundamentals of ML—in general and specifically in the context of LTPs—and a rather broad overview level, this work focuses on the large variety of physical and chemical aspects of LTP processing that have so far been addressed by data-driven methods (although probably not complete). While we target the topic of modeling and simulation, we specifically consider (i) how data-driven plasma science can complement conventional theoretical approaches, as well as (ii) when data-driven plasma science approaches

may replace classical concepts in the field of LTP modeling and simulation. Further, the interested reader is referred to fundamental literature for a general introduction to data science and ML.⁴⁻⁸

In the following, a review on data-driven methods for plasma modeling and simulation is presented, divided into the major physicochemical characteristics. In Sec. 2.1, approaches to plasma discharge physics phenomena are addressed. Related aspects on plasma chemistry follow in Sec. 2.2. Section 2.3 deals with plasma-surface phenomena on the mesoscopic and atomistic scale. Finally, Sec. 2.4 focuses on plasma process control and optimization.

2.1 Plasma Physics

LTPs for plasma processing entail a diverse zoo of physical and chemical phenomena with complex interactions. Modeling of such systems must, therefore, comprise at least the governing mechanisms and their intrinsic coupling. For example, consider the spatiotemporal dynamics of charged species coupled to electromagnetic fields. Data-driven modeling of LTP discharge characteristics has followed three prevalent routes: (i) data-driven modeling of a component of a physical model, based on data sets from numerical discharge simulations and finally coupled to the overall numerical simulation. (ii) Surrogate modeling that pursues representing the complete discharge dynamics. (iii) Direct embedding of the physical discharge model equations to be solved into a physics-informed ML procedure.

2.1.1 Data-driven surrogate sub-modeling

Contributions to this topic have prevalently considered separable submodels as part of a superordinate model consisting of several components. Plasma chemistry presents a specific case detailed later in Sec. 2.2. PSI is similarly addressed separately in Sec. 2.3. A major aspect to be considered in the following concerns the description of electromagnetic fields, in particular in electrostatics description. In this case, the Poisson equation presents an elliptic partial differential equation (PDE) subject to time-dependent boundary conditions (BCs). This problem stands out because of its computational effort and difficulties in an efficient parallel solution in multiple spatial dimensions (as the information of the whole domain is required at every iteration). Early work by Zhang et al.²³ on two-dimensional (2D) Monte Carlo (MC) simulations of a PN junction (conceptually related to methods applied in LTP) has considered reference data from the solution of the inhomogeneous Poisson equation subject to Dirichlet BCs using a finite-difference (FD) method. A multi-scale ANN was trained using the internal and boundary values of the electric potential as ground truth. Good agreement paired with a more than 10 times speedup has been reported. In an attempt to transfer a similar concept to LTPs, Trieschmann et al.²⁴ implemented a one-dimensional (1D) self-supervised learning scheme using an FD discretization to establish a physics-based loss, as, for instance, proposed by Dissanayake and Phan-Thien.²⁵ The data-driven solver was further assessed within a particle-in-cell (PIC) simulation of two collisionless counterpropagating electron streams. A relative error of the order of 10^{-4} was reported for the electric potential. However, as the force acting on each particle requires the electric field, evaluation of the first derivative implied significant distortion of the long-term spatiotemporal dynamics. This finding has been corroborated by Aguilar et al.²⁶ who have reported similar findings using fully connected ANNs [a.k.a. multilayer perceptrons (MLPs)²⁷] and convolutional neural networks (CNNs).²⁸ A 2D solution to Poisson's equation coupled to a particle balance equation for a double headed streamer simulation (in drift-diffusion approximation for electrons and otherwise no transport) has been pursued by Cheng et al.²⁹ (closely related to the work Özbay et al.³⁰) Therein, a comparison of different ANN architectures was presented, specifically a multi-scale network and a U-Net configuration.³¹ Whereas a significant speedup has been consistently reported, the accuracy of the predictions was observed to agree only for short term prediction (<2 ns). Note that a simple rectangular domain was considered with the addition of a standardization scheme to extend to variable domain sizes. Further, the loss was composed of the inner domain and Dirichlet boundary values, paired with a physics-based loss representing the Laplacian. Zeng et al.³² followed a different paradigm using physics-informed neural networks (PINNs).^{33,34} Therein, physical coordinates are input to an ANN and the underlying PDE is implemented as a physics-informed loss function. In general, this loss

function may include a set of possibly coupled PDEs to be solved within the domain, the BCs, and the initial conditions in case of transient phenomena.^{33,34} Automatic differentiation is exploited to obtain the respective local error gradients with respect to the input variables.³⁵ It resembles a paradigm closely linked to the backpropagation algorithm²⁷ for training ANNs and enables a mesh-less representation of the PDE solution. In the referenced work, the inhomogeneous Poisson equation with rectangular and circular dielectric objects was solved in two and three dimensions, subject to time-dependent Dirichlet and/or Neumann BCs. Note that the authors have initially included time as a variable to consider time-dependent BCs (but later reverted to a time-invariant formulation), although Poisson's equation essentially poses a static PDE problem. Hence, the error metric was composed of a domain loss representing the Laplacian and the loss corresponding to Dirichlet BCs. Whereas the application to dielectric barrier discharges (DBDs) was proposed as use case, an integration into a complete LTP model remains due. An approach dedicated to the requirements of complex LTP simulations has been put forward by Siffa et al.³⁶ based on a combination of U-Net,³¹ DenseNet,³⁷ and transformer^{38,39} network architectures, called TransDenseUNet. On a 2D Cartesian mesh, the non-overlapping domain information was encoded into types (plasma, dielectric, BCs), corresponding physical values (dielectric constant, charge density, and boundary potential/normal derivative), and spatial coordinates. A feature scaling of the physical equation was proposed and all input/output quantities were normalized in prior training. The data set was composed of a few physical reference cases for intended applications and a substantial amount (10 k to 50 k) of randomly generated data samples. The training loss was combined with various contributions [including a gradient loss and a structural similarity index measure (SSIM)⁴⁰] to increase the accuracy and the smoothness of the electric potential prediction (and the associated electric field). The proposed data-driven model was demonstrated to successfully generalize on arbitrary geometries, BCs, and charge distributions—including realistic cases encountered in LTP processing. Also here, an implementation within a complete plasma model remains due.

With the goal of studying the complex interaction of electromagnetic waves with plasma, Desai et al.⁴¹ suggested a U-Net architecture to predict the 2D distribution of the root mean square electric field based on plasma density distributions. The ground truth data set was obtained from finite-difference time-domain simulations. Interestingly, while a simple regularized L2 loss was used, the quality of the prediction was assessed using the SSIM and the average percentage error. Whereas overall good agreement with reference simulations was achieved, an issue generally noticeable with many ANN surrogate models was observed. Specifically, the convergence toward a reliable prediction is subject to a spectral bias. Lower frequency features converge significantly faster than higher frequency features. Consequently, it may be computationally intractable to train until convergence is achieved for all relevant frequencies.⁴²

2.1.2 Data-driven discharge surrogate modeling

An early example by Verma et al.⁴³ has considered data sets from 1D time-resolved fluid simulations of capacitively coupled radio-frequency discharges to establish surrogate models for integrated quantities, such as the plasma sheath voltages/currents, that capture reduced temporal modes of operation. In contrast to physical model based solutions, the authors have highlighted the equation-free representation of their long short-term memory [LSTM, cf. Sec. 3.1] surrogate model. The corresponding reduced order description and its latent space representation was suggested for parameter space exploration. This may otherwise be infeasible due to computational requirements. Another approach to capacitively coupled plasma surrogate modeling by Hamaguchi and co-workers has been devised based on 1D fluid and particle-in-cell with Monte Carlo collisions (PIC/MCC) simulations of an Ar discharge.⁴⁴ Instead of a grid-based approach for training and prediction (where the whole spatial domain is regressed on at once), MLPs were trained on phase-average plasma quantities sampled at varying positions within the discharge. The dynamics of an inductively coupled plasma with additional radio frequency substrate bias has been implemented based on hybrid plasma equipment model⁴⁵ simulations by Ko et al.⁴⁶ 2D axisymmetric simulations of an Ar etch plasma were performed with a fixed geometry (rectangular mesh) and varying process conditions (pressure, source power, bias power at 1 and 13.56 MHz). The devised data set was further used to train a multi-encoder/decoder ANN,

which encodes the process variables and geometry into a latent representation to finally decode the corresponding phase-averaged discharge quantities. Note that although the geometry was fixed in the presented study, the proposed ANN architecture with dedicated geometry inputs (sources, material, and properties) may also be capable of taking varying geometrical features (e.g., chamber aspect ratio, wafer diameter) into account. Due to its low computational costs, the ML surrogate model allowed for a rapid exploration and optimization of a uniform and sufficient etch rate using the multi-objective particle swarm optimization algorithm.

Data-driven surrogate modeling applied to different atmospheric discharges has been proposed by Zhang et al.^{47–49} A drift-diffusion model taking into account particle balance coupled to Poisson's equation, and the electron energy conservation equation was used to simulate pulsed He and CO₂ discharges. Corresponding data sets were established for varying source properties and subsequently used for setting up an MLP. A reduced set of input descriptors was selected, such as time and source properties (e.g., voltage, pulse rise rate). As output descriptors the discharge current density or the spatial distributions of species were chosen. With reference to these simulations (and experiments), it was demonstrated that a rather simplistic ANN even in the absence of complex learning schemes is capable of capturing the spatiotemporal discharge dynamics as well as the chemical reaction kinetics.

2.1.3 Physics-informed machine learning

PINNs were further used to directly consider the equations governing the physics of LTP discharges. Similar to the previously mentioned PINN solution to Poisson's equation, corresponding PDEs may be implemented in the error metric during training.^{33,34} Automatic differentiation is again exploited to obtain desired partial derivatives with respect to the input variables.³⁵ Early work in LTP has been proposed by Kawaguchi et al.⁵⁰ who have considered a PINN solution to the stationary Boltzmann equation subject to ionization and electron attachment processes. An MLP has been trained to predict the electron velocity distribution function (EVDF) following an exponential ansatz. The solution was validated for two reference gas mixtures, whereas the dependence on a variation of the reduced electric field was used as a case study. Excellent agreement between the data-driven approach and MC simulations was reported. Several improvements to their initial scheme have been suggested by the authors. Specifically, a nondimensionalization of the equation combined with an improved ANN architecture was used.⁵¹ Last is a fully connected ANN, which includes two transformer layers for an improved training of the PINN.^{38,52} Furthermore, a scheme was proposed to take into account variable reduced electric field values E/N .⁵³ In both studies, overall good agreement with MC simulations was demonstrated. For Ar gas eight processes were included: momentum transfer, six electronic excitations, and ionization. For SF₆ gas nine processes were included: momentum transfer, vibrational and electronic excitations, electron attachment, and ionization. An extension to the time-dependent Boltzmann equation and assessment of the memory capacity of the ANN of 0.1% compared to mesh-based direct numerical simulation (DNS) calculations of EVDFs in three-dimensional (3D) velocity space was further argued.

In the frame of PINNs for LTP, Zhong et al.⁵⁴ initially investigated stationary and transient 1D arc discharges. A fully connected ANN was used to implement the stationary Elenbaas–Heller equation (with decay term) as well as the time-dependent Elenbaas–Heller equation (considering mass and energy conservation). Whereas the PDE was generally solved using PINNs, the thermodynamic, transport, and radiation properties of the plasma were (partially) taken into account via so-called coefficient subnets. These resemble small regression networks to interpolate the corresponding plasma properties. In a follow-up study,⁵⁵ the authors have introduced this framework more generally in the frame of coefficient-subnet physics-informed neural network and Runge–Kutta physics-informed neural network. For the first approach, a comparison with cubic spline interpolation was included. The last approach embeds PINNs in an implicit Runge–Kutta formalism. In a case study for three examples of Boltzmann equation (with ionization, electron attachment, elastic collision, and excitation), Poisson-drift-diffusion equations (for electrons and ions), and the previously addressed 1D arc discharge using the time-dependent Elenbaas–Heller equation, excellent agreement was noted. Zhong et al.⁵⁶ lastly revisited the data-driven 1D arc discharge simulation, introducing the concept of meta-learning.⁵⁷

Therein, a meta loss was aggregated over multiple tasks (e.g., parameters governing discharge conditions, time-dependence) to enable an efficient two-step optimization and, therefore, reduce the training time. For each training task, a regular PINN was trained. In contrast, the meta network was trained to provide good initial weights for a PINN solving a new task. Based on a variation of arc discharge parameters and corresponding meta-training, a substantial acceleration with a factor in the range of 1.1 to 6.9 was reported.

A different approach to PINNs following Raissi et al.³³ involves the discovery of PDEs from observations. A single PDE or a set of PDEs may be formulated as prototype equations. Therein, unknown coefficients λ_i determine the exact form of the equations. The problem can be phrased such that training a PINN not only yields the desired solution to the PDE but also includes the coefficients λ_i as the parameters of the PINN. This concept has been used by Kawaguchi et al.⁵⁸ to “discover” electron transport coefficients from electron swarm maps measured by scanning drift-tube experiments. The temporal and spatial evolution of the electron swarm was formulated in two PDEs with corresponding transport coefficients. A fully connected ANN with transformer layers was utilized similar to previous works by the authors.⁵¹ The tunable parameters are optimized along with the ANN, without assuming any analytical form of the electron swarm map. It was concluded that more accurate transport coefficients could be obtained, including higher-order dependencies if required.

Xiao et al.^{59,60} considered a data-driven surrogate model based on recurrent neural networks (RNNs) and a reduced-order model (ROM) for model predictive control (MPC) of a plasma etch process. It is, therefore, addressed more specifically in Sec. 2.4.

2.2 Plasma Chemistry

LTP models are based on a set of chemical reactions that describe volumetric interactions between all the species tracked in the model, such as electrons, ions, and neutrals.⁶¹ An accurate description of such interactions is required in order to understand and optimize the selective generation of chemically active species and their properties, such as mean energy and velocity distribution. For plasma modeling applications, chemistry sets may include up to hundred species and several thousands of reactions.⁶² Such chemical complexity can hinder the process of understanding and optimizing plasma applications. Furthermore, the plasma modeling community is sometimes facing scarce availability of fundamental data, such as cross sections and reaction rate coefficients, that are used as input in the models.⁶³ With significant advances of ML, new tools are available to the community to improve plasma chemistry models. There are three major roles of ML with respect to plasma chemistry developments, that are: (i) extraction of cross sections and reaction rate coefficients, (ii) derivation of reduced reaction mechanisms, and (iii) chemical engineering. In particular, the following questions are addressed in the present review:

- How can electron-impact cross sections with atoms and molecules of interest be rapidly computed?
- How can ML enable developments of reduced reaction mechanisms for plasma chemistry?
- How to efficiently engineer chemical reactions systems with ML?

2.2.1 Extraction of cross sections and reaction rate coefficients

The use of an ANN as an optimization technique for treating the inverse problem of obtaining electron collision cross section from electron transport coefficients has been proposed by Morgan.⁶⁴ In that pioneering work, it was assessed that neural networks are indeed useful to determine the elastic momentum transfer cross sections of electrons in real gases, but the accuracy level is highly dependent on the quality of input cross section data that are used for training of the model. Stokes et al.^{65,66} improved the idea by Morgan and implemented deep neural network (DNN) using cross sections from the LXCat database⁶⁷ associated with electron transport coefficients found by numerical solutions of the electron Boltzmann equation or by experimental measurements. In their works, the authors demonstrated that their automatic solution using ANN had an accuracy comparable to that of a human expert in determining cross sections of electrons in He or in the biomolecule tetrahydrofuran (THF). The authors also showed that the use of large amount of synthetic training data generated by using the real cross sections available from LXCat provide good results for the prediction of elastic momentum transfer and ionization cross sections

of He and Ar.⁶⁶ However, the architecture of the ANN used in Refs. 65 and 66 had minor improvements over the architecture proposed by Morgan.⁶⁴ Jetly and Chaudhury⁶⁸ implemented an ANN, specifically CNN and densely connected convolutional network (DenseNet)³⁷ to derive electron impact cross sections from a solution of the inverse electron swarm problem. The networks were trained using elastic momentum transfer, ionization, and excitation cross sections for different gases available in the LXCat database⁶⁷ and their corresponding electron transport coefficients, which were calculated using the BOLSIG+ solver.⁶⁹ For the first time, the performance and accuracy in cross sections determination from different network architectures was assessed, and it was found that DenseNet predicts cross sections with significantly higher accuracy compared to other ANN models due to its ability to extract both long and short term features from the electron transport parameters. Figure 1 shows that the DenseNet model, as opposed to other ANN-based models (e.g., CNN), is able to predict specific cross sections features in energy,

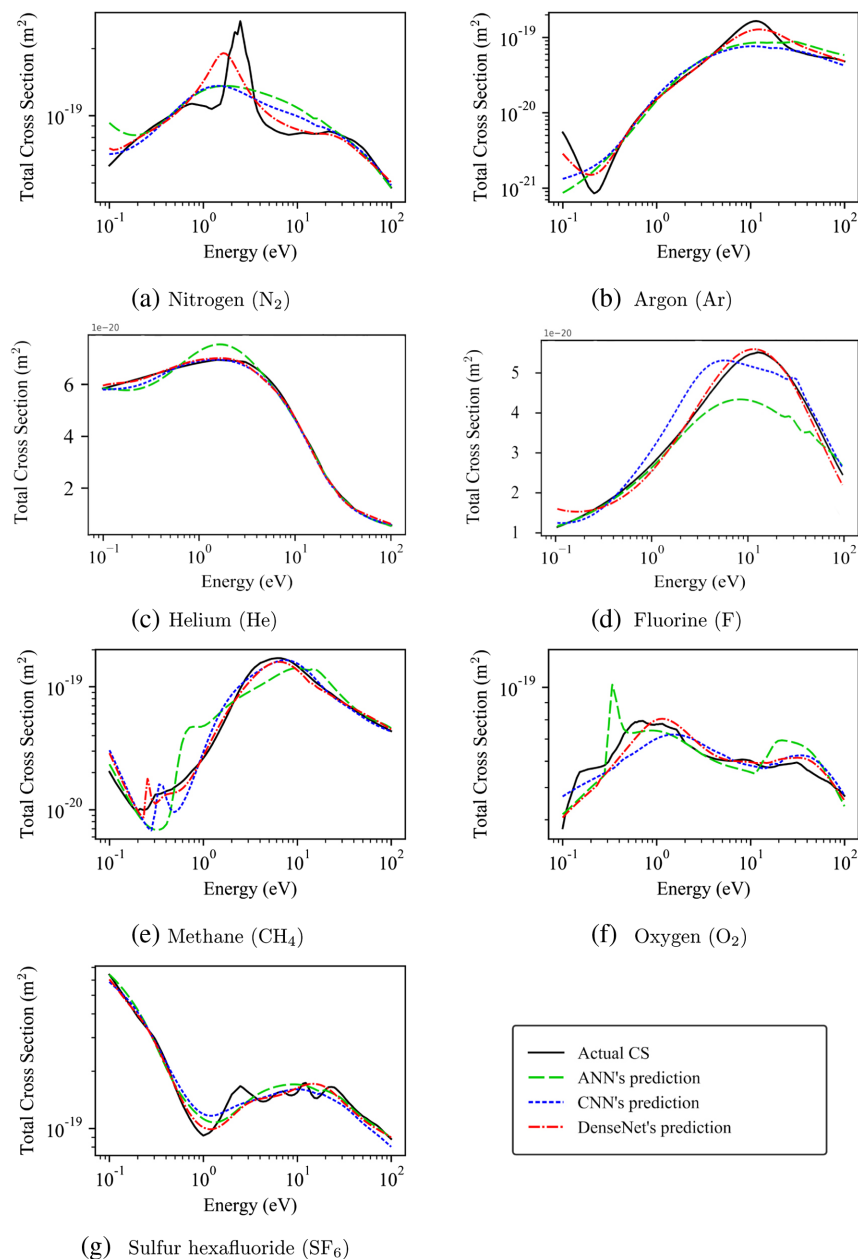


Fig. 1 Predicted total momentum transfer cross sections of electrons with various gas species. Figure reproduced from Ref. 68 licensed under CC BY 4.0. (a) Nitrogen (N_2), (b) argon (Ar), (c) helium (He), (d) fluorine (F), (e) methane, (f) oxygen (O_2), and (g) sulfur hexafluoride (SF_6).

such as the resonance of the elastic momentum transfer cross sections of electrons in N_2 . Improvements of such deep learning models, such as the ability to extract multiple excitation cross sections and/or the use of sophisticated synthetic data generation schemes, may facilitate the use of ML-derived electron-impact cross sections in plasma chemistry models.

Fast prediction of collision cross sections based on simulated cross sections has been initially suggested by Zhong.⁷⁰ While the author used density functional theory (DFT) calculations to devise a data set for electron impact ionization cross sections for molecules, it has been argued that the approach would be similarly applicable to experimentally measured cross sections. Based on a support vector machine (SVM) model with radial basis function kernel, the prediction of accurate cross sections has been verified. Moreover, the capability to generalize from small to large molecules has been argued, without the need for computationally costly DFT calculations. Following a similar paradigm,⁷¹ Harris and Nepomuceno recently developed a conceptually simple but effective ANN model to predict electron-molecule ionization cross sections from a database of measured cross sections. In particular, by exploiting the chemical configuration (i.e., number of atoms in a molecule) as input to the network, the authors have evaluated the capability of a generalized prediction of cross sections for additional molecules.

Direct determination of reaction rate coefficients and their temperature dependence is also important in plasma modeling. In fact, each reaction needs to be assigned with its own kinetic parameters and these are sometimes not available or subject of large uncertainties. Reiser et al.⁷² used ML to derive rate coefficients for radio-frequency atmospheric pressure plasmas operated with various gas mixtures of O_2 , CH_4 , and He. In their work, the authors used a genetic algorithm to derive rate coefficient values that bring the results of the integrated set of particle balance equations for different species in the plasma into good agreement with the experimental data. Even if this approach could be prone to biases induced by user experience and be limited by the availability of experimental data, it has been demonstrated that results from the genetic algorithm open up the possibility of learning more about the significance and correlation of individual reactions. Hence, this method is potentially interesting for post-processing and evaluating experimental data using simplified reaction schemes. Hanciniec et al.⁷³ explored the use of ML to supply unknown reaction rates in plasma chemistries, thus allowing complete chemistry sets to be generated without resorting to estimations by analogy or an educated guess. In their work, ML was used to provide rate coefficients of binary chemical reactions using three distinct optimized regression models: an SVM model, a random forest model, and a gradient-boosted trees model. The models were trained on kinetic data for binary heavy-species collisions at or near room-temperature extracted from the QBD,⁷⁴ KIDA,⁷⁵ NFRI,⁷⁶ and UfDA⁷⁷ databases. After removing duplicate reactions, the final data set consisted of 9470 reactions involving 1080 distinct species. As a sample use case, the ML results were used to augment the chemistry of a BCl_3/H_2 gas mixture. In this example, the reactions between the various BCl_x species and H were missing. Results of their work showed the importance of taking such reactions into consideration for which the ML algorithm likely gives better estimates than intuitive guesses while being faster than measurements or *ab initio* calculations.⁷³

2.2.2 Derivation of reduced reaction mechanisms

Several methods have been developed to provide interpretation and obtain proper reduction of complex chemical kinetic sets. There are two main methods for reducing reaction mechanisms: reparametrization of the chemical state space and graph theory related algorithms. For the first category, we can refer to principle component analysis (PCA)⁷⁸ and intrinsic low-dimensional manifold (ILDM)⁷⁹ as two methodologies that are used for LTP kinetics reduction. In PCA, a manifold is generated from linearization of the chemical source terms. Reduction is obtained since the chemical state space that is described by a small number of parameters, the principal components (PCs), and balance equations for the PCs are solved instead of the full set of equations. Peerenboom et al.⁸⁰ applied PCA to a 0D state-to-state kinetic model of CO_2 including vibrational levels of CO_2 and CO. The *a priori* analysis showed that log-transformation and scaling are important for the manifold identification. Moreover, a significant speedup in central processing unit (CPU) time has been obtained due to the fact that PCA reduced not only the number of variables but also the stiffness of the equations. The ILDM method has been applied

by Rehman et al.⁸¹ to Ar and H₂ plasma kinetic schemes. It has been shown that the ILDM method automatically extracts the relevant information about fast and slow time scales in a chemical reaction system. After a short interval of time, the fast time scale processes will quickly move toward this low-dimensional manifold and the slow time scale processes will move tangential or along the manifold to finally reach the equilibrium point.⁸¹ Despite successful application in plasma chemical systems, PCA and ILDM also present some drawbacks. For example, in PCA, the manifolds are constructed based on the hypotheses of linearity between original and transformed variables that cannot be always satisfied for complex chemical networks. Moreover, ILDM can be problematic for chemical kinetics that have more than one equilibrium point or for more complex systems where the increase in the dimensionality of the manifolds and subsequently of the generated look-up tables can require higher computational resources.

Graph theory has been applied to chemical kinetics for visualization, topological statistics calculation of a chemical network, and determination of pathways from a sequence of reactions. Sakai et al.⁸² have been the first to apply it to complex plasma chemistry. In their work, they have constructed a network structure for methane and silane plasmas by representing the species as nodes and reactions as directed unweighted edges and they demonstrated that this network analysis allows one to estimate substantial effects of given species on other ones existing in the network. Then, the work was extended with betweenness and centrality indices to highlight not only the influentiality of each species on/to the others but also global structures in chemical networks.⁸³ In this respect, these works have opened up possibilities for simpler prediction of chemical reactions that one would like to understand and/or control. A similar approach for graph construction has been adopted by Sun et al.,⁸⁴ with a different choice of undirected edges weighted based on steady-state reaction rates. The reduction method has been applied to CO₂ chemical kinetics for a gliding arc plasma. Murakami and Sakai⁸⁵ have used a graph-based approach for extracting a low-dimensional reaction set based on centrality indices. In particular, their approach was applied to He – O₂ and He-humid air plasma chemistry, and they have shown that a reduced chemical reaction set for these systems can be achieved with the help of topological centrality (closeness and betweenness) and scale-freeness. In particular, in Fig. 2, it is shown that a plasma-induced reaction network can be described with a scale-free network, where the number of degrees originating from a particular species follows a power-law distribution. The advantage of their approach is that it can be used as preprocessing to reduce complex plasma chemistry before multi-dimensional simulations. Mizui et al.⁸⁶ extended the analysis in Ref. 85 to transient states by constructing temporal graphs. Their classification method was used to deduce information from silane and methane plasma chemistries. In particular, quantitative insights into chemical mechanisms have been derived with axes of multi-centrality indices. Holmes et al.⁸⁷ used a directed graph representation of plasma mechanisms in which both species and reactions are identified as nodes, with the edges weighted based on the corresponding reaction rate coefficients. Furthermore, they have used a connectivity matrix to get qualitative information on fast and slow reaction pathways. With this approach, they have shown that the new approach can produce a clearer image for the interpretation of complex chemistry. Venturi et al.⁸⁸ presented

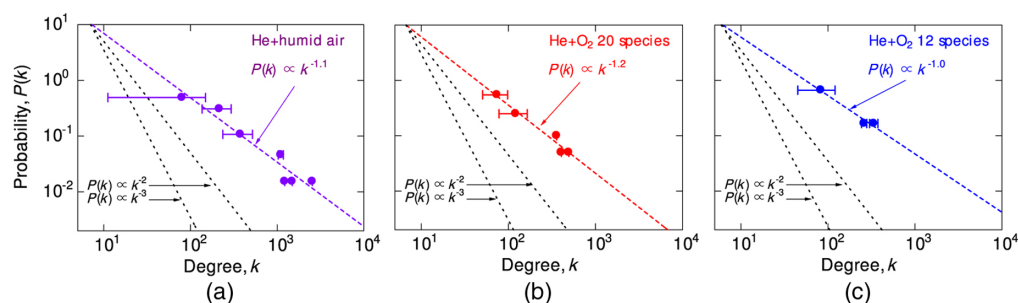


Fig. 2 Power-law distributions of the number of degrees for (a) He + humid air plasma, (b) He + O₂ plasma, 20-species model, and (c) He + O₂ plasma, 12-species model. Figure reproduced from Ref. 85 licensed under CC BY 4.0.

a novel method for reducing plasma chemistry mechanisms by taking into account the uncertainties affecting the reaction parameters. As a difference with respect to previous works, their approach relies on an ensemble of weighted directed graphs having both species and reactions as nodes. Edges are weighted based on the reaction contribution to the particle balance equation. The methods were applied to a plasma mixture of 23 species including 15 vibrational states of the H_2 molecule, three electronic states of the H atom, four ions, and electrons. As a result of their work, it has been shown that a graph ensemble significantly improves the methodology's robustness while preserving the predictive accuracy. In general, graph theory related algorithms have been successfully applied to find patterns in large chemical data set. These algorithms could be used to identify and analyze patterns in chemical kinetic data prior running experiments or simulations.⁸⁸

2.2.3 Chemical engineering

In many applications, the plasma community is interested in knowing “how we can” generate a specific concentration of species. In this respect, modeling and simulations are powerful tools to allow one to actively modulate plasmas and/or discharges for production of chemically reactive species. In this section, advances in ML for chemical engineering of plasma systems are explored. Zhu et al.⁸⁹ developed a DL framework called DeePlasKin for non-equilibrium plasma systems that can be described by a global chemistry model. The model is based on a predictor-corrector approach detailed in Sec. 2.4. For a given kinetic scheme and pre-defined temporal evolution of target species densities, it allows one to extract the temporal profile of the reduced electric field (E/N) and all other species densities. Pan et al.⁹⁰ investigated a multi-layer feed-forward neural network for plasma kinetic modeling. The deep learning model has been applied to kinetic schemes for $CH_4 - Ar$ pulsed discharge and a $N_2 - H_2$ plasma catalysis model for ammonia production. The magnitude of E/N and time were chosen as the input data of the DNN, whereas the output was selected as the target densities of each species calculated at every instance of time by inquiring the kinetic model. The results showed that the network can replace complex kinetic models with large reactions set while significantly improving the computational efficiency. Wang et al.⁴⁹ proposed an MLP to describe the discharge characteristics and plasma chemistry of CO_2 pulsed DBD discharges at atmospheric pressure. The MLP was trained using data obtained from fluid simulations. As a result of their work, the trained model takes a few seconds to predict various features of CO_2 pulsed discharges, which significantly improves the computational efficiency with respect to the 30 h CPU time that are required for the fluid model. Overall, these works have shown the great potential of ML for optimization and design of plasma sources in practical applications. In fact, the proposed models have been able not only to automatically extract plasma parameters that are difficult to measure in experiments but also possess high computational efficiency, which provides a promising approach for multi-dimensional plasma modeling.

2.3 Plasma–Surface Interaction

Particle fluxes onto LTP bounding surfaces contain a variety of species (e.g., electrons, neutrals, ions, and radicals), which trigger different kinds of PSIs upon impact, for instance, surface chemical reactions or sputtering. Each of which is more or less relevant for certain plasma applications, which are correspondingly engineered to lessen or strengthen one or the other mechanism. For modeling such processes and in particular the involved PSIs, a wide range of methodologies has been established, such as transport of ions in matter⁹¹ or molecular dynamics (MD). Each method is limited by a combination and trade off between accuracy and computational feasibility. Recent advances in both regards utilizing ML will be outlined in the following for processes that either are driven by (i) surface chemical reactions or (ii) physical sputtering.

2.3.1 Surface chemical reactions

In addition to the species and surface chemical reactions inherent to ordinary gas-phase processing (e.g., chemical vapor deposition, atomic layer deposition, catalysis), plasmas introduce radicals, excited species, and ions and hereby enable a manifold of new reactions at the surface.

This circumstance is employed for instance for the plasma enhanced atomic layer deposition (PEALD) of metal oxides. It consists of two self-limiting half cycles with purging cycles in between, out of which the second half cycle, the plasma cycle, is dedicated to the deposition of O atoms, molecules, and radicals. Ding et al.⁹² modeled PEALD of HfO₂ thin films with a kinetic MC scheme. Reaction path ways and energy barriers were obtained by combining an intensive literature research with DFT calculations. Such simulations were used to assemble a data set to train a Bayesian regularized artificial neural network (BRANN).⁹³ The trained network was used for the prediction of the half cycle process times, which are essential for industrial considerations, as a function of total precursor pressure and surface temperature. The BRANN was constructed by combining an MLP network architecture with a Bayesian regularization, which penalizes large weights to achieve a better generalization. In a follow-up work,⁹⁴ the authors pair their kinetic MC model with a computational fluid dynamics (CFD) process simulation. An RNN was then correspondingly trained to predict the evolution of the surface profile as well as gas-phase kinetics. Recurrent links establish internal states, which allow for the prediction of the temporal dynamics. The issue of vanishing as well as exploding gradients during the training of the RNN was addressed by including an LSTM (cf. Sec. 3.1).

Plasma enhanced atomic layer etching (PEALE) of copper has been suggested by Sheil et al.⁹⁵ to include a half cycle of plasma induced oxidation followed by a half cycle of exposing the surface to etchant molecules (e.g., HCOOH), which solely removes the just formed oxide layer. While the latter step is naturally self-limiting, the former is not. This finding has been seconded by Xia and Sautet,⁹⁶ who performed MD simulations for a corresponding plasma oxidation of copper. A high-dimensional neural network potential was trained by considering the ReaxFF potential^{97,98} and DFT calculations for the initial and on-the-fly data generation, respectively. The ML potential was demonstrated to overcome the limitations of classical as well as reactive interaction potentials imposed by their functional form, which are argued to not allow for accurate descriptions of very short range interatomic distances.

For the particle fluxes from the plasma, not only radicals but molecules with high vibrational excitation are assumed to play an integral role for plasma catalysis. This is assessed although their contribution and cannot be experimentally clarified due to an overlay of effects, for instance, as induced by high electric fields. In contrast, simulations are well suited to provide experimentally unattainable insight. Wan et al.⁹⁹ conducted DFT simulations to investigate the role of electric field-dipole interactions for the CH₃ synthesis catalyzed with Ru surfaces. The resource intensive computations were accelerated by approximately five orders of magnitudes by generalizing the aggregated data with ANNs. Relevant atom configurations were processed with graph neural networks (GNNs),^{100,101} specifically graph convolutional networks,¹⁰² whose network structure consists of nodes (i.e., atoms) and edges (i.e., configurational similarities). Multiple shared MLPs were used to combine the output with information on the electric field, adsorbate energy, and surface energy to predict all quantities of interest (i.e., adsorption energies, dipole moments, and polarizabilities of reactive intermediates). Kedalo et al.¹⁰³ performed MD simulations to show that the activation of highly excited N₂ molecules on Ru surfaces during NH₃ synthesis can be approximated by the Fridman–Macheret α model.¹⁰⁴ The interatomic interactions were described by an ML potential, utilizing smooth overlap of the atomic positions (SOAP) descriptors,¹⁰⁵ to enable *ab initio* accuracy but with increased computational efficiency. Bal and Neyts¹⁰⁶ demonstrated that the Fridman–Macheret α model is insufficient to describe the vibrational nonequilibrium (i.e., lowering of the free energy barrier) during the catalytic dissociation of H₂ and CH₄ on Ni surfaces at low temperatures of the order of 750 K. The multitudes of potential reaction pathways in between two metastable states were simplified by applying a harmonic linear discriminant analysis as a supervised dimensionality reduction algorithm.¹⁰⁷ The states themselves were studied by conducting short MD simulations with enhanced sampling methods (i.e., variationally enhanced sampling,¹⁰⁸ metadynamics^{109,110}) to introduce the vibrational nonequilibrium.

Siron et al.¹¹¹ proposed a high-throughput workflow for the study of amorphous material (i.e., a-C, a-Si, a-SiO₂, a-Al₂O₃) surface reactions relevant to dry plasma etching. Surface sites were characterized with SOAP descriptors and clustered with a Bayesian Gaussian mixture model to effectively reduce sites to be studied for reaction kinetics, specifically etching.

2.3.2 Physical sputtering

The bombardment of surfaces with energetic ions spawns collision cascades in the subsurface region, which eventually lead to persistent radiation damage and the emission of surface material atoms from the surface. The latter is referred to as sputtering and is of major importance for the sputter deposition of thin films as well as etching of high aspect ratio features.

Hamedani et al.¹¹² used an ML interaction potential to conduct an MD study on the radiation damage and sputtering of Si. The utilized ML potential, specifically the Gaussian approximation potential (GAP),¹¹³ was originally trained to equilibrium properties and, hence, complemented with a DFT repulsive potential (i.e., DMol¹¹⁴) for short range interactions to resolve the evolution of the collision cascade accurately. The simulations provided sputtering yields that were in good agreement with experimental reference values. A greater number of isolated small defect clusters was observed than provided by MD simulations based on traditional interaction potentials (i.e., Stillinger–Weber,¹¹⁵ Tersoff¹¹⁶). Moreover, a new defect structure was revealed that is three split interstitials surrounding a central vacancy. This structure and the likelihood for its formation was further analyzed in a follow-up study.¹¹⁷

Sputtering yields can also be predicted by utilizing the Yamamura model.¹¹⁸ Phadke et al.¹¹⁹ shown that its model parameter Q , which relates to linear scaling, has a simple positive correlation with the sputter threshold energy. This result was obtained by applying a Bayesian Markov chain MC algorithm to TRIDYN¹²⁰ simulations of Ar⁺ and Ne⁺ ions bombarding 70 elemental targets. Eckstein and Preuss¹²¹ applied Bayesian probability theory to improve the accuracy of the Yamamura model's sputtering yield predictions by empirical means. Preuss et al.¹²² conducted SDTrimSP^{120,123} simulations of D⁺ ions sputtering Fe targets to revise the model's internal parameter (i.e., surface binding energy). This was achieved by to applying a reduced-order spectral expansion¹²⁴ uncertainty quantification based on the Bayesian framework.

Krüger et al.¹²⁵ conceptualized a PSI ML surrogate model for the Ar⁺ ion bombardment of a TiAl composite target as indicated in Fig. 3. The ML model, specifically an MLP, was meant to map the incident ions' energy distributions to the energy angular distributions (EADs) of sputtered particles. It was trained with noisy data from TRIDYN simulations (poor statistics), but demonstrated to be capable of predicting smooth EADs, which were almost congruent to reference TRIDYN simulations with high statistical quality. Gergs et al.¹²⁶ continued those efforts by considering TiAl composite targets with varying compositions, introducing the stoichiometry as a basic system state to the model. A similar data aggregation scheme based on TRIDYN simulations was executed. The ML model was replaced by a combination of a convolutional β -variational autoencoder (β -VAE)^{127–129} and CNN, defining a reduced dimensional (latent) space for the sputtered particles' EADs and mapping of the incident ion energy distributions as well as TiAl target composition to this particular latent space, respectively. Predictions and reference TRIDYN simulations agreed with each other, even though the assembled PSI ML surrogate model had merely 0.39% of the MLP's¹²⁵ degrees of freedom and, hence, a significantly reduced likelihood for overfitting. Kim et al.¹³⁰ pursued a similar but more advanced approach for the

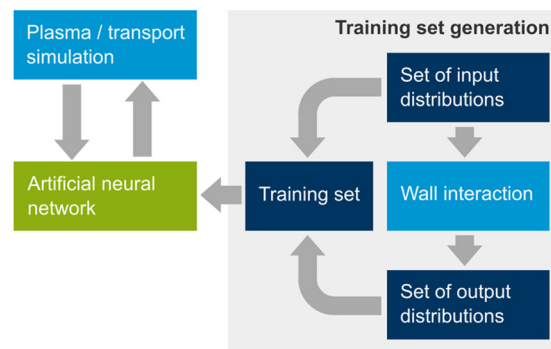


Fig. 3 Conceptual diagram of the ANN plasma-surface interface model and the data flow for training set generation and run-time model evaluation. Figure reproduced from Ref. 125 licensed under CC BY 4.0.

sputtering of fluorinated Si surfaces due to Ar⁺ ion bombardment. MD instead of TRIDYN simulations were conducted to obtain information on the emitted (i.e., sputtered, reflected) particles' EADs as a function of the surface F coverage as well as Ar⁺ ions' incident angle (but not kinetic energy). The aggregated EADs were used for a deep learning-based ROM (DL-ROM), comparing the performance of different ML architectures [i.e., autoencoder²⁷ (AE), denoising AE,¹³¹ VAE, conditional VAE (CVAE)¹³²] with each other. The CVAE was found to best its competitors. The PSI ML surrogate model was completed by introducing an additional MLP, which mapped the input information (i.e., surface F coverage, Ar⁺ ions' incident angle) to the latent space established by the DL-ROM, enabling resource efficient predictions with atomic, MD fidelity. Gergs et al.¹³³ performed hybrid MD/time-stamped force-bias Monte Carlo (tfMC)^{134–136} simulations to setup a PSI ML surrogate model but treated sputtering of Al targets due to Ar⁺ ion bombardment and growth of Al thin films at the substrate as generalized wall interactions. The incident particle flux composition as well as Ar⁺ ion energy was varied to setup the data set.¹³⁷ Two CVAEs were stacked to setup a physics-separating artificial neural network (PSNN), separating the ion bombardment induced damage from the damage to surface state (e.g., composition, mass density) translation. This network architecture has been argued to allow for training with experimental and simulation data in spite of an inherent discrepancy regarding the information accessibility (damage cannot be measured by similar means). The predictions mitigated the noise but maintained key physical features (e.g., threshold for Ar clustering, reflection probability) for small particle doses. In a follow-up work,¹³⁸ the authors have shared a methodology that enabled predictions for particle emission as well as growth during the reactive sputter deposition of AlN thin films in Ar/N₂ plasmas on experimental time-scales, that is, minutes to hours. First, the MD interaction potential was revised with an evolution strategy implemented in the genetic algorithm-based reactive force field optimizer method (GARFField)¹³⁹ to guarantee accurate PSI simulations.¹⁴⁰ Second, a high-throughput, randomized data generation scheme was pursued to efficiently populate the relevant parameter space with hybrid MD/tfMC simulations.¹³⁸ Third, two CVAEs were trained to predict either the PSIs or the diffusion processes. They were combined to form a PSNN¹³³ as PSI ML surrogate model, which can readily be used to complement either plasma simulations or experimental diagnostics. The latter scenario was considered to successfully validate the model. Process times of 45 min were predicted within merely 34 graphics processing unit (GPU) hours. In contrast, hybrid MD/tfMC simulations were argued to take more than ~8 million CPU years to finish on a comparable case study.

2.4 Plasma Process Control and Optimization

Many applications in LTP processing require precise control of the process and discharge parameters governing the deposition and/or removal of material from the plasma-facing surfaces (e.g., plasma etching). Whereas a detailed understanding through a virtual process simulation ideally provides all relevant intrinsic information in real-time, this pathway is often limited by a mismatch between the real process and its virtual replica, along with the typical run-time requirements of a process model. Therefore, several data-driven alternatives have been explored: (i) MPC based on ROMs. (ii) Plasma state estimation from incomplete observable information. (iii) Data-driven process recipe design and optimization. (iv) Data-driven parameter space exploration and optimization. Corresponding examples for these categories will be addressed in the following.

2.4.1 Model predictive control

In a series of contributions, Woelfel et al. developed MPC schemes for radio-frequency reactive magnetron sputtering fundamentally based on the Berg model.¹⁴¹ Following the identification of an ROM taking into account the nonlinear process dynamics due to surface poisoning,¹⁴² they have devised a controller for process operation at the transition between poisoned and clean surface modes using the previous reduced model, combined with measurements from multipole resonance probe experiments.^{143,144} An improved reduced model and novel control scheme (also based on the Berg model) has been introduced later, taking into account the plasma state and its dynamical behavior.¹⁴⁵

The control problem of delivering a specific thermal dose during application of a kHz atmospheric pressure plasma jet (APPJ) operated in helium has been detailed by Mesbah et al. in a sequence of publications discussed in the following. Subsequent to model identification of an ROM (from experimental measurements), a feedback control strategy for 2D spatial thermal dose delivery was proposed.^{146,147} The substrate temperature as obtained from infrared thermal imaging was controlled based on the He flow and lateral shuttle position over a surface. The hierarchical feedback control strategy enabled reliable dose delivery despite abrupt disturbances that were introduced. An extension of the approach was suggested as treatment protocol, where the control of the surface temperature included the treatment time in an optimal control problem. One aspect was the adaptation of the control scheme to converge to the optimal solution for the real APPJ system.¹⁴⁸

The optimization of a plasma etch process has been investigated by Xiao et al. using a MPC scheme based on a physical multi-scale model pairing a 2D fluid description of the discharge and a kinetic MC simulation of the surface kinetics. Initially, based on the simulated fluxes at the substrate surface, an RNN surrogate model was set up. Another one was implemented for the surface description, taking into account the stochasticity of the MC model.⁵⁹ Subsequently, a similar physical modeling approach has been proposed for a hierarchical surrogate model strategy.⁶⁰ Using similar 2D plasma fluid simulation results as for the previous approach, an ROM was devised via proper orthogonal decomposition and Galerkin's method.¹⁴⁹ This reduced model was argued to capture the intrinsic system dynamics on a low-dimensional representation. It thereafter was used for training another RNN surrogate model for rapid prediction of the plasma discharge dynamics during argon-chlorine etching of silicon. In both works by Xiao et al.,^{59,60} MPC schemes have further been developed to control the etching depth and the feature bottom roughness of the process.

2.4.2 Plasma state estimation

With the goal to estimate the internal state of an LTP system in real time despite limited access to observable quantities, a physically constrained extended Kalman filter (PC-EKF) has been devised by Greve et al.¹⁵⁰ First, the general principle including physical constraints (e.g., prohibiting negative quantities) was assessed at the example of the Lorenz system. The PC-EKF has been shown to closely resemble its chaotic system dynamics. Thereafter, the dynamics of a Hall effect thruster described by a global ionization model were considered. Based on particle balance equations for electrons and ions, the electron temperature was dynamically estimated based on the measured discharge currents. Successful recovery of the electron temperature was demonstrated by the investigation of four different discharge modes subject to varying types of ionization oscillations. The proposed scheme has been argued to benefit from improved robustness with the ability to estimate low-frequency dynamics in the presence of possibly not well resolved high-frequency dynamics (in the model or experiments). In a follow-up study also by Greve et al.,¹⁵¹ the PC-EKF approach was generalized to more complex Ar and Ar/O₂ chemical reaction sets and global models of ICP sources. Particle balance equations as well as a power balance equation for electrons were taken into account. In an extensive case study, the robustness of the scheme has been demonstrated. Among other cases, a real time estimate of one or multiple unknown state variables based on corresponding limited measurements (e.g., electron absorbed input power based on argon ion density) has been shown, including pulsed operation.

2.4.3 Data-driven process recipe design and optimization

In contrast to real time process control and estimation, a study by Kanarik et al.¹⁵² has investigated ML aided process recipe design. A virtual process of a radio-frequency etch plasma with fluorocarbon and O₂ chemistry coupled to a feature profile simulator to estimate etching and critical dimension (CD) has been used for reference. A schematic of the virtual process is depicted in Fig. 4. A comparison of junior and senior human experts against three computer algorithms to design a process recipe was evaluated, focusing on a minimization of the accumulated cost-to-target for achieving design goals. Their GPR model was found to greatly outperform the other algorithms (MCMC sampling, tree-structured Parzen estimator).

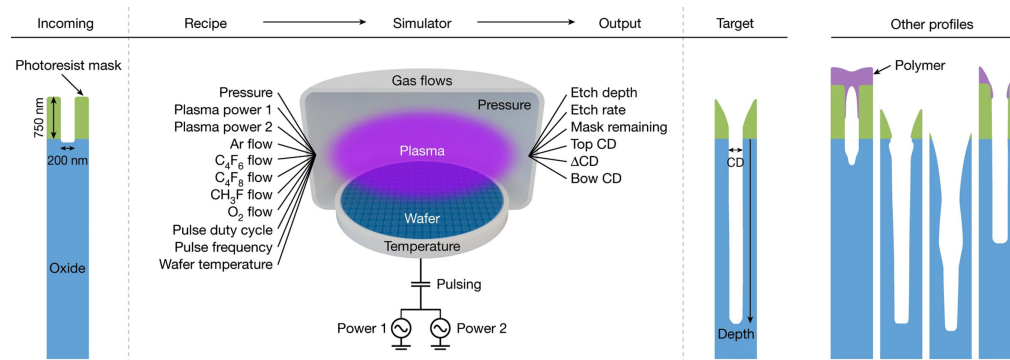


Fig. 4 Schematic of the virtual process used for the process recipe design challenge. Figure reproduced from Ref. 152 licensed under CC BY 4.0.

Notably, a threshold was observed, determining the ideal handover from human to computer algorithm design for the specific challenge. This is linked to the number of trials (and hence data samples) required for achieving the specified target. The proposed human first–computer last strategy has been argued to reliably reduce the cost-to-target, whereas human experts and data-driven computer algorithms excel at different tasks (high versus low-dimensional parameter space exploration).

2.4.4 Data-driven parameter space exploration and optimization

Data-driven parameter space exploration in LTP has been predominantly pursued for atmospheric plasmas. As highlighted in Sec. 2.2, Zhu et al.⁸⁹ investigated a global chemistry model (with the possibility to interface with a Boltzmann solver) for a DBD as a function of the electric field. A discrete time-stepping predictor/corrector scheme has been suggested: (a) a reconstruction step to obtain a guess of the next reduced electric field E/N value for a selected target density (at the next time step). (b) A prediction step using an ordinary differential equation (ODE) solver of the physical model and the estimated E/N to obtain the future system state. (c) A correction step where the procedure is reevaluated if the difference between the outcome and the target exceeds a given threshold. The model enabled tailoring of the LTP system state evolution, allowing for efficient parameter space exploration and control of the plasma chemistry.

A more general approach to parameter space exploration has been suggested by Shao et al.²² based on GPR. The authors have trained an MLP as a surrogate model based on experimental data. Following the paradigm of Bayesian optimization (BO), this data-driven model is then exploited to set up a GPR model with an expected improvement acquisition function. For a considered use case of nitrogen fixation experiments of an atmospheric DC pin-to-pin glow discharge, for example, the energy cost of NO_x generation was targeted for minimization. An active learning scheme based on this BO was established for exploration of a multi-dimensional parameter space.

3 Advances

The need for LTP and plasma chemistry databases has been previously highlighted (see Sec. 2 and referenced literature).^{9,10,63} Many data-driven advances in modeling and simulation of LTP science and technology may be enabled only by a strategic data generation and acquisition initiative. This is particularly the case since data-driven techniques intrinsically require significant amounts of training data, which need to efficiently explore the relevant parameter space. While there is enormous potential in such data generation and exploitation, we reason that this is a matter of scientific practice and awareness but not a technical obstacle to a substantial extent. In the following, we focus on technical aspects that are hypothesized to have the potential to advance the field, proceeding from ML methods to LTP disciplines. We initially discuss ML concepts and methodology of general relevance to LTP modeling and simulation. Therein, we limit the discussion mainly to DL concepts. Thereafter, approaches and concepts specific to individual LTP problems in plasma physics, plasma chemistry, and PSIs are addressed.

3.1 ML Methodology

Data-driven LTP modeling and simulation inherently addresses the description of complex physicochemical systems, on multiple dynamical time and length scales. Whereas some aspects of data-driven modeling targeting dynamical systems have been investigated in plasma science, a vast number of concepts from related scientific disciplines appears to be under-explored in LTP. Therefore, in what follows, we target widely applicable ML approaches to dynamical system modeling and their interpretation: (i) artificial neural network topologies. (ii) Reduced order modeling and model identification. (iii) Learning schemes for optimal data exploitation. (iv) Explainable AI to support physics understanding. We aim to highlight the general concepts and draw the connection to specific aspects of plasma science in general and LTP processing in particular where applicable.

3.1.1 Artificial neural network topologies

Several approaches to the modeling of dynamical systems using ANNs have been proposed. Early concepts have evolved around the class of RNNs.¹⁵³ Recurrent network connections establish feedback loops between neurons, whereas the previous output of a neuron subject to a time delay is fed back to the same or another neuron. Feed-forward RNNs only maintain local recurrent connections per layer; fully recurrent network architectures establish a complete coupling between all neurons. Similar to ordinary fully connected ANNs, recurrent connections may be multiplied with learnable weight parameters and passed through nonlinear activation functions. Training of an RNN involves the backpropagation of errors through time, passing them on through the evolution of the system.¹⁵³ This might be challenging due to vanishing or exploding gradients and a large computational effort. The RNN may finally resemble the system dynamics, through its (possibly many) fitted weight parameters. Examples in LTP processing have already been addressed.^{59,60}

An important variant of RNNs are LSTM networks suggested by Schmidhuber et al.^{154,155} While similar in terms of recurrent feedback, their success is largely enabled by memory cells and gate units instead of basic neurons. Each memory cell with a self-connection includes gate units to control the information flow into and out of the cell, as well as of the recurrent feedback. As such, this may be attributed to the concepts of learning and forgetting¹⁵⁵ and allows for a substantially improved learning efficiency and robustness. LSTM networks have been applied in fluid dynamics for the prediction of the temporal dynamics of turbulent flow through channels based on DNS data.¹⁵⁶ They have further been used for the prediction of disruption in magnetic confinement fusion plasmas.^{157,158}

Moreover, in a study on a hierarchical approach to multiscale data-driven modeling, LSTM networks have been included for comparison.¹⁵⁹ Another fundamentally related variant of RNNs that was included in this study are echo-state networks (ESN).^{153,160} Therein, an RNN network reservoir is initially set up. A number of inputs are fed into the RNN reservoir, but a number of outputs are extracted from the reservoir. In contrast to ordinary RNN networks, which are comprised of learnable weight parameters at every neural node, only a subset of output weight parameters is adjusted during training. Consequently, a manifold of complex nonlinear system dynamics may inherently be evolved by the network, whereas only a small number of output weight parameters capture the desired system dynamics to be described by the data-driven model. To this end, a physics-informed ESN architecture has been investigated for the description of chaotic systems.¹⁶¹ For the chaotic dynamics of a Lorenz system and a Charney–DeVore system, robust training was achieved with an extended prediction time horizon of approximately two Lyapunov times.

Following the physics-informed learning paradigm,³³ the number and variants of approaches exceeds the frame of what can be covered here. The interested reader is referred to an extensive review on PINNs in scientific ML by Cuomo et al. (which also partially covers RNNs and its variants).¹⁶² It may be noted that the technical implementation of PINNs is subject to ongoing research. This is partially due to the limited performance and convergence characteristics of PINNs in general and of ordinary ANNs such as an MLP in particular.^{52,163} Another approach to physics-informed ML extending this idea is based on deep neural operator networks (DeepONets).¹⁶⁴ With the reasoning that ANNs can be considered universal operator

approximators (similar to the property that ANNs are universal function approximators), these networks comprise a trunk network and a (single or stack of) branch network(s). The branch part takes as input discrete sensor samples $u(x_i)$ of a function $u(x)$. The trunk part takes as input the quantity y . The output of both sub-networks is combined to provide a functional representation $G(u)(y)$. Following an offline training for a variety of examples, Lu et al.¹⁶⁴ demonstrated online inference performance significantly improving on other reference ANNs (among others LSTM networks). Different variants and extension taking advantage of Fourier and Laplace transformations as nonlinear kernels have further been proposed.^{165–167}

Transformer architectures have been originally developed in NLP.³⁸ The examples previously addressed have fundamentally explored similar architectures without substantial conceptual modifications (e.g., either by using standalone transformer layers^{51,53,58} or similar to an image-to-image translation task³⁶). Classical transformer networks process input information through embeddings that transform the input data into a corresponding vector space representation. Recently, an extension to the concept of transformer networks has been suggested by Geneva and Zabaras for surrogate modeling of physical systems.¹⁶⁸ Therein, alternative embeddings have been suggested based on Koopman dynamics in a two step learning procedure. Initially, embedding ANNs are trained in an encoder–decoder structure to approximate the Koopman based dynamics in a latent space representation. Thereafter, a transformer network is trained using the pretrained Koopman embeddings. The obtained data-driven surrogate models have been demonstrated to effectively capture the chaotic system dynamics of a Lorenz system, the 2D fluid dynamics of a transient flow, and 3D reaction-diffusion dynamics. In comparison, the Koopman based transformer models have been demonstrated to outperform alternative surrogate models of such dynamical physical systems (including other embeddings and LSTM networks).

3.1.2 Reduced order modeling and model identification

ROMs have previously been used in LTP modeling and simulation as exemplified above. The particular notion of sparse identification of nonlinear dynamics (SINDy)^{169,170} has been demonstrated to provide interpretable and generalizable nonlinear differential equations of reduced order, which effectively captures the dynamics of the full model. This is achieved by identifying the relevant system dynamics from a library of potential functional forms by means of a sparse regression. It has been applied to space and astrophysical plasmas based on PIC simulations by Alves and Fiuza demonstrating the potential of recovering the fundamental dynamics of the high-fidelity kinetic plasma simulations.¹⁷¹ It may be straightforwardly applicable to LTP modeling and simulation data to capture their nonlinear dynamics, albeit a complex chemistry may necessitate further adaptation. Sparse regression has similarly been applied by Kaptanoglu et al. in the context of magnetic confinement fusion plasmas for the optimization of magnetic fields in stellarator fusion plasmas.^{172,173} Moreover, MPC based on such ROMs has been suggested with limited availability of training data.¹⁷⁴ An aspect that has been applied using time-dependent simulations of the tokamak plasma boundary.^{173,175} It may be specifically relevant also in the context of LTP processing with its diverse requirements for precise process control.

3.1.3 Learning schemes

In contrast to classical supervised learning schemes, alternative approaches may be relevant also due to limited availability of training data. In a Bayesian setting, provided with an estimate of the prediction uncertainty, active learning has been exploited for efficient parameter space exploration.¹⁷⁶ While such an uncertainty metric is straightforward in the context of GPR modeling,²² similar techniques have been adapted to MD simulations and active learning of corresponding ML surrogate models.¹⁷⁷

Transfer learning enables an alternative route to ML in the small-data limit.¹⁷⁸ The fundamental idea relates to the property of generalization that data-driven models pursue. Given an initial problem task with abundant data, an ML model may be robustly trained. The pretrained model may be subsequently retrained on a related, yet different small-data task, whereas just a minor subset of weight parameters is allowed to be adjusted. Hence, the robustness and

generalized prediction of the initial model may be successfully transferred to the more challenging learning task. This approach has been successfully applied for instance in the design optimization in inertial confinement fusion.¹⁷⁹

3.1.4 Explainable AI

As surveyed by Gilpin et al.,¹⁸⁰ a multitude of explainable AI (XAI) concepts have been proposed for the interpretation and explanation of data-driven models. As an example, local interpretable model-agnostic explanations (LIME) have been suggested for identifying interpretable models that locally approximate the nonlinear behavior of any underlying model.¹⁸¹ While this is possibly an underexplored aspect in the context of LTP, an example where LIME relates to the data-driven classification and interpretation of OES experiments of plasma in aqueous solution has been proposed.¹⁸² Further works by Lundberg et al.¹⁸³ have included LIME and other methods in an extension Shapley additive explanations. It is similarly based on local explanation models but defines additive feature attribution methods to unify several of the included interpretable methods. There is arguably no mathematical proof that physics-based and physics-informed ML and corresponding models capture exactly the desired dynamics (and only those). LTP modeling and simulation has got its inherent foundation in a rigorous mathematical description and numerical solution. XAI methods may not only foster advances in data-driven approaches to the field but also strongly contribute to the general acceptance in LTP science and technology.

3.2 Plasma Physics

Modeling and simulation of LTP physics is traditionally facing two substantial challenges. First, their complex system dynamics typically involve multiple physical phenomena and, second, these dynamics spread across many orders of magnitude in length and time. As previously outlined, both challenges may be addressed and partially solved by (i) data-driven surrogate modeling of sub-problems and (ii) data-driven surrogate modeling of the complete discharge dynamics. Despite intrinsic progress in the LTP field alone, these efforts may be largely supported through consideration of advances in adjacent research disciplines, as discussed in the following.

3.2.1 Data-driven surrogate sub-modeling

Describing the evolution of dynamical systems (or sub-systems) accurately is an omnipresent task in scientific computing. The time integration schemes of differential operators have been correspondingly considered in different fields (e.g., pairing Runge–Kutta integration schemes with ANNs^{55,159,184,185}). For the time integration of multiscale physical systems, Liu et al.¹⁵⁹ suggested a hierarchical time-steppers (HiTSs) framework. The approach builds on the decomposition of the system dynamics into a hierarchy of prediction steps that can be arranged and interpolated in time. It has been shown to outperform other methods based on LSTM, ESN, and clockwork RNNs for the Lorenz system, 2D fluid flow past a cylinder, the Kuramoto–Sivashinsky (KS) equation, as well as music and video sequences. An illustrative example comparison of the prediction output is shown in Fig. 5. LTP modeling and simulation is facing essentially alike problems, suggesting it may similarly benefit from application of the HiTSs method (and/or other methods referenced) in the description of their multiscale dynamics.

An approach using the concept of dynamic mode decomposition (DMD) to obtain low-dimensional linear dynamics of given data has been applied to magnetic confinement fusion plasma simulations by De Pascuale et al.¹⁸⁶ The obtained dynamic modes are evolved individually solving a linear ODE. The constructed linear time advance operator based on DMD has been successfully applied for the analysis of several 1D diffusion transport problems, as well as for the dynamics of 1D profiles at the divertor targets of a tokamak plasma. This scheme could be similarly applied to complex LTP processes.

The solution to Newton's equations for particle based simulations has been explored by Kadupitiya et al. using LSTM networks.¹⁸⁷ The accurate applicability has been investigated for few-particle systems subject to various interaction potentials (e.g., Lennard–Jones). Energy-conserving dynamics up to 4000× larger timesteps compared to a Verlet integration have been

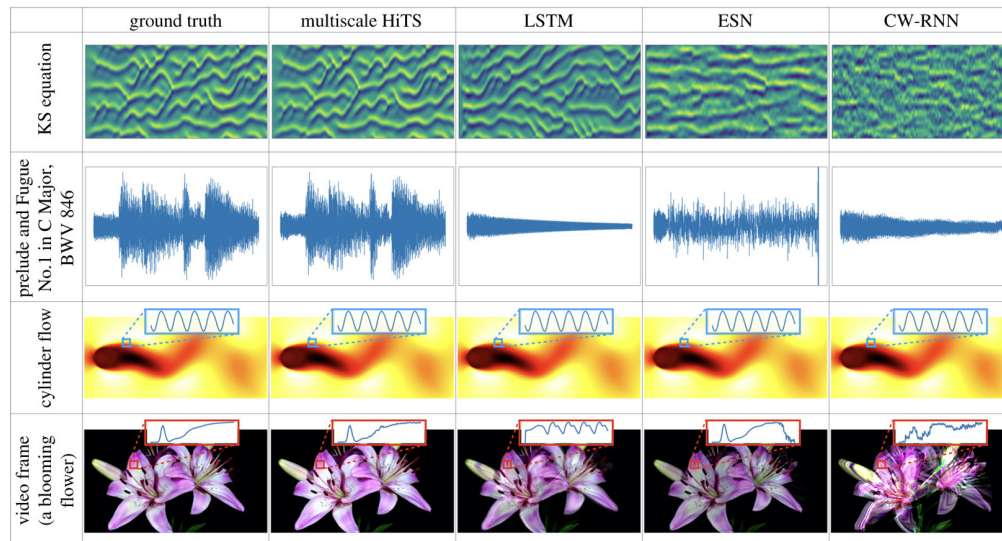


Fig. 5 Output of different ANN architectures for the prediction of different training sequences (Kuramoto–Sivashinsky equation, music, flow past a cylinder, video). Figure reproduced from Ref. 159 licensed under CC BY 4.0.

demonstrated. Following the concept of GNNs,^{100,101} which embed the interaction dynamics of nodes (e.g., particles) in their learnable edge weights (e.g., interaction), the integration of ODEs in a Hamiltonian framework have been discussed by Sanchez-Gonzalez et al.^{188,189} The proposed Hamiltonian ODE graph network trained on Runge–Kutta scheme data generalizes the integrators to accurately represent the Hamiltonian dynamics. In the context of many-particle simulations, Li et al.¹⁹⁰ proposed a GNN accelerated MD framework that predicts the interaction forces for the system, given its momentary state (e.g., atom positions and types). It has been argued that the scheme can be easily integrated into existing numerical integrators. Due to the consideration of multi-particle interactions, this scheme may be specifically relevant in PIC simulations should Coulomb interactions be important.

3.2.2 Data-driven discharge surrogate modeling

Many of the ML methods that may be transferred to data driven surrogate modeling of LTPs were discussed in Sec. 3.1. This relates specifically to the representation of dynamical systems through recurrent ANN models. A few examples from plasma science and adjacent fields were covered. However, it goes beyond the scope of this paper to cover related works from fluid dynamics. The reader is referred to corresponding comprehensive surveys by Brunton and Vinuesa et al.^{191,192}

Concerning data-driven process control and MPC, it should be noted that the plasma information based VM methodology for LTP etching processes previously introduced may be adapted based on physical or data-driven model data.^{19,20,193} Despite a possible discrepancy between models and reality, in particular systematic deviations may be straightforwardly taken into account in a data-driven approach by means of discrepancy learning. A learnable translation layer may suffice to capture and combine the relevant information by means of data-fusion. This may enable data-driven MPC with limited observables (cf. also PC-EKF^{150,151}). In addition, MPC schemes based on LSTM networks that have been successfully demonstrated in magnetic confinement fusion plasmas subject to inherent complex nonlinear dynamics by Degraeve et al.¹⁹⁴ may be promising candidates for adaptation in LTP processing.

3.3 Plasma Chemistry

Several challenges arise in modeling and simulations of plasma chemistry. These are linked to the setup of accurate reaction mechanisms and high computational cost of plasma models with complex chemistries. This section describes advances in data-driven models and applications for such systems. Examples are mainly taken from the field of combustion, reactive gas flow modeling,

and astrophysics. In particular, two different topics are discussed for (i) dimensionality reduction of complex chemical networks and (ii) data-driven approaches for fast and accurate chemical models. Such topics offer promising solutions for future investigations that are focused on bridging the gap between low cost and high accuracy in plasma chemistry models.

3.3.1 Reduction of complex chemical networks

The adoption of detailed mechanisms for chemical kinetics often poses two types of challenges. First, the number of degrees of freedom is large; and second, the dynamics is characterized by a wide range of time scales. This has motivated the development of several techniques for reducing the complexity of such kinetic models where only a few variables are considered in the simplified model. Grassi et al.¹⁹⁵ used AEs to reduce complex chemical data sets and decrease the computational time associated with numerical solutions of ODEs for a large number of species and reactions. In particular, an encoder is used to reduce the dimensionality of the chemical state space into a latent (lower-dimensional) space. Analogously to the chemical network in the original space, which is represented by the system of ODEs, the chemical network in the latent space can be evolved in time with a different set of ODEs, which has got a significantly smaller number of dimensions. Hence, a decoder is used to transform the variable from the latent space back to the original space. The proposed model enables a compression of chemical networks composed of 224 reactions and 29 species into a network with 12 reactions and 5 species that are evolved in time using a standard ODE solver. Their work has demonstrated that a 65-times speed-up can be achieved. Tang and Turk¹⁹⁶ proposed a combination of AE and neural ODE to model the temporal evolution of chemical kinetics in a reduced subspace. The architecture of the network is similar to the one by Grassi et al.¹⁹⁵ In their work, a 10-fold speed-up compared to commonly used astro-chemistry solver for a 9-species primordial network has been achieved while maintaining 1% accuracy across a wide range of density and temperature. As described in Sec. 2.2, it is well known that, due to the presence of fast and slow dynamics, the chemical systems are characterized by low-dimensional manifolds in the concentration space. Chiavazzo et al.¹⁹⁷ applied the diffusion Map (DMAP) approach for constructing reduced kinetics models for combustion applications. The DMAP approach can be considered as a non-linear counterpart of PCA that can be used for searching a low-dimensional embedding of a high-dimensional chemical data set.

3.3.2 Fast and accurate models of complex chemistries

Stiff equations for chemical kinetics are computationally expensive to solve numerically. The efficacy of ML for applications that need to solve systems of ODEs or PDEs has been explored in many areas, such as CFD and combustion. Goswami et al.¹⁹⁸ developed a neural operator-based surrogate model (DeepONet) to efficiently solve stiff chemical kinetics. The model was applied to a chemical reaction system for CO/H₂ burning of syngas, which contains 11 species and 21 reactions, and a temporally developing planar CO/H₂ jet flame (turbulent flame) using the same syngas mechanism. It has been shown that the DeepONet, once trained, can accurately integrate the thermochemical state for arbitrarily large time advancements, leading to significant computational gains compared to stiff integration schemes. Moreover, both results obtained with DeepONet and AE-based DeepONet are computationally very efficient compared to conventional CFD solvers. One of the advantages of this approach is the possibility to be included in a CFD solver for solving such stiff chemically reacting problems, which can drastically reduce the computational cost. Campoli et al.¹⁹⁹ performed state-to-state numerical simulations of high speed reacting gas flow using data-driven ML regression models. The models were applied to a system of equations for a 1D reacting flow of a five-component air mixture including a total of 122 excited states of N₂, O₂, NO, N, and O species with detailed vibrational kinetics. Several methods have been compared for estimation of the state-to-state relaxation rate terms, such as kernel ridge (KR), SVM, k-nearest neighbor, Gaussian processes, decision tree (DT), random forest, extremely randomized trees, gradient boosting, histogram-based gradient boosting, and MLP. From their comparison, it has been found that KR reports the lowest error levels while the SVM reports the highest. Comparable error levels are reported by the remaining algorithms;

however, there are noticeable differences in the prediction time, where DT is found to be the fastest algorithm. Finally, the solution of the state-to-state Euler system of equations was inferred by means of a DNN by-passing the use of the solver. As a result of this study, it has been demonstrated that embedding of ML algorithms into ODE solvers offers a speed-up of several orders of magnitude. Chemical reactor networks (CRNs) have been investigated by Savarese et al.²⁰⁰ for performing faster CFD simulations of combustion systems. In their work, a novel automatic data-driven method for the design of CRN models has been proposed. The method is based on a combination of unsupervised clustering and graph scanning algorithms. Results of their work show that the novel methodology is capable of extracting equivalent CRN models of realistic combustion devices from CFD data. Furthermore, the unsupervised clustering algorithm was able to automatically detect important regions within the combustor domain, such as inlets, inflame, and post-flame reactors. Overall, the approach is promising as it has the potential to drive toward the development of ROMs without a huge amount of user-based knowledge. Mao et al.²⁰¹ developed a novel CFD solver that combines the multi-functionalities of the OpenFOAM library²⁰² with the ML framework Torch²⁰³ and chemical kinetics program Cantera.²⁰⁴ In their solver, DNNs are used to accelerate chemistry solvers and improve the simulation efficiency. As a result, a speed-up of two orders of magnitude is achieved in a simple hydrogen ignition case when performed on a medium-end GPU. This work highlights the potential of the integration of ML techniques within current libraries and computing architecture for detailed CFD simulations of complex hydrocarbon fuels.

3.4 Plasma–Surface Interaction

The review of ML utilized for PSIs was subdivided by means of physics (i.e., surface chemical reactions, sputtering) as well as related processes (e.g., PEALD/PEALE), highlighting differences and how pioneering works utilized ML to advance the research in such fields. The consideration of ML methods established and proposed for similar processes without plasmas (e.g., thermal atomic layer deposition/etching) may allow one to adopt novel techniques as well as inspire new studies. This opportunity is particularly valuable for research on plasma catalysis, where ML has already been applied in manifold ways and promising advances are foreshadowed. An excellent overview on ML assisted modeling and simulations of heterogeneous catalysis has been put together by Mou et al.,²⁰⁵ whereas Chen et al.²⁰⁶ assembled a more detailed review for the most recent accomplishments due to ML potential-based atomic simulation.

In the following, overarching challenges relevant to all fields of LTP processing are discussed. The intrinsic time and length scales of the two states of matter (i.e., plasma, solid-state) differ in orders of magnitude. From a material scientist point of view, who resolves atomic time and length scales, PSIs have to be perceived as individual or a sequence of impinging particles, neglecting the true spatiotemporal correlations. Multiscale modeling approaches that repeatedly substitute short time and length scales are, however, always accompanied with an information loss of the fundamental atomic phenomena. ML is assumed to provide an alternative.

For electrons, information on their dynamics, specifically, on the secondary electron emissions are essential for LTP modeling and simulation.²⁰⁷ Although experimental as well as theoretical approaches have been pursued in the past, the availability of such coefficients is scarce. More recent efforts include the combination of both (i.e., PIC/MCC simulations, phase resolved OES) and was framed as γ -CAST.^{208,209} Moreover, a review of MC methods to simulate the electron scattering in solids was put together by Chang et al.,²¹⁰ who studied the low-energy electron deposition in W surfaces. Electron force field simulations have been demonstrated to predict electron emissions due to the Auger process.²¹¹ Auger neutralization was also described by Pamperin et al.,²¹² who formulated a generic quantum-kinetic approach that is based on Anderson-Newns-type effective Hamiltonians. Bronold and Fehske²¹³ derived an equation for the electron emission yield in case of low energy electrons impinging onto metal surfaces. An invariant embedding principle was employed. It is argued that such approaches or combinations thereof could synergistically complement each other, in particular when paired with corresponding ML methods.

Information on the atomic scale have been limited for a long time by a compromise of accuracy and computational efficiency, effectively restricting the physical fidelity of related case

studies. Recent advances by means of ML interaction potentials have indicated to eventually overcome this dilemma. However, only a small number of such works have been published so far in the context of LTP modeling and simulation (cf. Sec. 2.3). It is therefore essential to further encourage their usage, enabling more likely developments that meet demands specific to plasma processing. The burden of setting up suitable ML interaction potentials could be bypassed at least to some extent by joining available ones with more traditional methods. For example, the Ziegler–Biersack–Littmark potential²¹⁴ and all-electron DFT repulsive potential, DMol, have been merged with ML interaction potentials (e.g., GAP), which were trained to equilibrium properties, to account for screened nuclear repulsion during high-energy collisions when radiating Si or Al.^{112,117,215,216} Caro et al.^{217,218} conducted MD simulations with the GAP to study the ion beam deposition of C thin films. Studying a complementary sputter deposition can be readily achieved by introducing Ar⁺ ions to the process, whose interactions are sufficiently well described by traditional means. An overview of and introduction to ML interaction potentials for the popular large-scale atomic/molecular massively parallel simulator has been presented by Thompson et al.²¹⁹

The increased efficiency of such approaches allows for a more thorough but still accurate parameter space exploration. This is of major importance since most PSIs (e.g., surface chemical reactions, ion radiation) are subject to excessive degrees of freedom and cannot straightforwardly be evaluated (e.g., reaction path ways, 0D to 3D defect structures). Remedies combine ML algorithms for clustering (e.g., Bayesian Gaussian mixture model), dimensionality reduction (e.g., harmonic linear discriminant analysis), or surrogate modeling (e.g., PSNN) with high-throughput simulations (e.g., randomized trajectories,¹³⁸ amorphous surface site melt-quenching¹¹¹) or enhanced sampling methods (e.g., variationally enhanced sampling, metadynamics).^{106,111,138,220} Notably, GNNs have also been demonstrated to significantly accelerate atomic simulations.⁹⁹ An active learning scheme as applied by Diaw et al.¹⁷⁷ when conducting MD simulations of warm dense matter to setup a corresponding surrogate model may be considered to secure spending available resources most effectively.

Whereas samples in the parameter space are naturally related to different process conditions, they may also be viewed as instances or states of a system (surface patch), which may evolve during operation. Lim et al.²²⁰ formulated an according transition state modeling approach^{221–223} for the long-time scale (i.e., of the order of microseconds) restructuring of Pd deposited on Ag. States were found by detecting relevant positional changes of adatoms during MD simulations, clustering them (unsupervised), and eventually sorting them (supervised). Gergs et al.¹³⁸ used a PSNN to generalize on previously sampled but disjointed system state trajectories. The system alternately underwent PSIs and diffusion processes for given experimental discharge conditions, continuously evolving the system state (e.g., composition, point defect populations). Following a similar line of thought, single instances or states could also be understood as surface patches, which may be assembled to setup a PSI ML surrogate model on the feature scale. Kim et al.¹³⁰ suggest a similar usage for etching high-aspect ratio feature patterns. However, such an approach may also be of significant interest for research on plasma catalysis, taking complex large-scale surfaces into account. Joining such ML models, which are created in a bottom-up style (starting with simulations on the microscopic scale), with a complementary top-down approach (starting with experiments on the macroscopic scale) may ultimately enable a fidelity beyond DFT.

4 Concluding Remarks

Scientific ML has had a significant impact in many research disciplines. In modeling and simulation of LTPs, it has experienced increased interest and important developments primarily within the past few years. In this time period, however, already a substantial interdisciplinary exchange could be witnessed. Within this review we had two main objectives: (a) review the state of the art in ML and data-driven approaches to LTP modeling and simulation. With the goal to highlight specific use cases and applications, we divided our survey by physicochemical and modeling aspects relevant for LTP processing (rather than methodologies). (b) Provide a perspective of potential advances. These may be enabled through transfer and inspiration of current methodologies from other scientific disciplines as well as future ML developments in general.

It certainly demands adaptation to the requirements faced in LTP processing, e.g., including the system dynamics for a precise control of deposition/etching process at the atomic level.

Open access to literature, databases, research data, and codes, as well as data-driven models are foreseen to be an integral part in these efforts.^{63,224–226} Sharing information following the findable, accessible, interoperable, and reusable principles²²⁷ would allow researchers independent of status and financial budget to develop the field. We believe this may enable vital new pathways for research on LTP processing.

It is important to stress that taking advantage of these developments should be assessed on a rational basis. A (more or less) critical evaluation should therefore ask which advances are truly enabled by corresponding ML methods. An answer may relate simply to the run-time performance of a corresponding numerical simulation. Moreover, it may pertain novel insights previously hidden within data, e.g., results from numerical simulation. ML methods may ultimately uncover unprecedented insights and interpretation otherwise not attainable. An associated aspect further relates to the net computational budget required to obtain certain data sets, the training procedure, the costs of run-time evaluation, and the capability of generalization of the data-driven models. This may be contrasted with the effort of evaluating individual numerical simulations.

Despite being generally true for science and technology, it may be envisioned that scientific ML and data-driven approaches to LTP modeling and simulation may not only support the analysis of known unknowns. In contrast, due to its inherent propensity to spotlight hidden patterns in data, it may encourage discovery of unknown unknowns. This aspect is argued to offer unprecedented potential for advances in plasma science and technology and beyond.

Code and Data Availability

Data sharing is not applicable to this article, as no new data were created or analyzed.

Acknowledgments

The article was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project-ID 138690629 (TRR 87), Project-ID 434434223 (SFB 1461), and Project-ID 445072286. L. V. was supported by the Stanford Energy Postdoctoral Fellowship through contributions from the Precourt Institute for Energy, Bits & Watts Initiative, StorageX Initiative, and TomKat Center for Sustainable Energy.

References

1. R. A. Gottscho, C. W. Jurgensen, and D. J. Vitkavage, "Microscopic uniformity in plasma etching," *J. Vac. Sci. Technol. B: Microelectron. Nanometer Struct. Process. Meas. Phenom.* **10**, 2133–2147 (1992).
2. G. S. Oehrlein and S. Hamaguchi, "Foundations of low-temperature plasma enhanced materials synthesis and etching," *Plasma Sources Sci. Technol.* **27**(2), 023001 (2018).
3. T. Tomie, "Tin laser-produced plasma as the light source for extreme ultraviolet lithography high-volume manufacturing: history, ideal plasma, present status, and prospects," *J. Micro/Nanolithogr. MEMS MOEMS* **11**, 021109 (2012).
4. C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford, United Kingdom (1996).
5. C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, MIT Press (2005).
6. S. Haykin, *Neural Networks and Learning Machines: A Comprehensive Foundation*, 3rd ed., Prentice Hall International, New York (2008).
7. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, Cambridge (2016).
8. A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, 3rd ed., O'Reilly Media (2022).
9. I. Adamovich et al., "The 2022 plasma roadmap: low temperature plasma science and technology," *J. Phys. D: Appl. Phys.* **55**(37), 373001 (2022).
10. R. Anirudh et al., "2022 review of data-driven plasma science," *IEEE Trans. Plasma Sci.* **51**, 1750–1838 (2023).
11. A. D. Bonzanini et al., "Foundations of machine learning for low-temperature plasmas: methods and case studies," *Plasma Sources Sci. Technol.* **32**, 024003 (2023).
12. M. Kambara et al., "Science-based, data-driven developments in plasma processing for material synthesis and device-integration technologies," *Jpn. J. Appl. Phys.* **62**(SA), SA0803 (2023).

13. R. Shadmehr et al., "Principal component analysis of optical emission spectroscopy and mass spectrometry: application to reactive ion etch process parameter estimation using neural networks," *J. Electrochem. Soc.* **139**, 907–914 (1992).
14. S. J. Hong, G. S. May, and D.-C. Park, "Neural network modeling of reactive ion etching using optical emission spectroscopy data," *IEEE Trans. Semicond. Manuf.* **16**, 598–608 (2003).
15. C. Himmel and G. May, "Advantages of plasma etch modeling using neural networks over statistical techniques," *IEEE Trans. Semicond. Manuf.* **6**, 103–111 (1993).
16. A. Bleakie and D. Djurdjanovic, "Growing structure multiple model system for quality estimation in manufacturing processes," *IEEE Trans. Semicond. Manuf.* **29**, 79–97 (2016).
17. S. J. Hong and G. S. May, "Neural-network-based sensor fusion of optical emission and mass spectroscopy data for real-time fault detection in reactive ion etching," *IEEE Trans. Ind. Electron.* **52**, 1063–1072 (2005).
18. K. Suthar et al., "Next-generation virtual metrology for semiconductor manufacturing: a feature-based framework," *Comput. Chem. Eng.* **127**, 140–149 (2019).
19. M. Maggipinto et al., "DeepVM: a deep learning-based approach with automatic feature extraction for 2D input data virtual metrology," *J. Process Control* **84**, 24–34 (2019).
20. S. Park et al., "Micro-range uniformity control of the etching profile in the OLED display mass production referring to the PI-VM model," *Phys. Plasmas* **28**(10), 103505 (2021).
21. T. van der Gaag, H. Onishi, and H. Akatsuka, "Arbitrary EEDF determination of atmospheric-pressure plasma by applying machine learning to OES measurement," *Phys. Plasmas* **28**, 033511 (2021).
22. K. Shao et al., "Active learning-guided exploration of parameter space of air plasmas to enhance the energy efficiency of NOx production," *Plasma Sources Sci. Technol.* **31**, 055018 (2022).
23. Z. Zhang et al., "Solving Poisson's equation using deep learning in particle simulation of PN junction," in *Joint Int. Symp. Electromagn. Compatibility, Sapporo and Asia-Pac. Int. Symp. on Electromagn. Compatibility (EMC Sapporo/APEMC)*, pp. 305–308 (2019).
24. J. Trieschmann et al., "A machine learning approach to the solution of Poisson's equations for plasma simulations," in *Bull. Am. Phys. Soc., APS, (Virt. Conf.)*, p. MW3.002 (2020).
25. M. W. M. G. Dissanayake and N. Phan-Thien, "Neural-network-based approximations for solving partial differential equations," *Commun. Numer. Methods Eng.* **10**(3), 195–201 (1994).
26. X. Aguilar and S. Markidis, "A deep learning-based particle-in-cell method for plasma simulations," in *IEEE Int. Conf. Cluster Comput. (CLUSTER)*, pp. 692–697 (2021).
27. D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," ICS Report 8506, University of California, San Diego, California (1985).
28. K. Fukushima, "Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biol. Cybern.* **36**, 193–202 (1980).
29. L. Cheng et al., "Using neural networks to solve the 2D Poisson equation for electric field computation in plasma fluid simulations," arXiv:2109.13076 (2021).
30. A. G. Özbay et al., "Poisson CNN: convolutional neural networks for the solution of the Poisson equation on a Cartesian mesh," *Data-Centr. Eng.* **2**, e6 (2021).
31. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," *Lect. Notes Comput. Sci.* **9351**, 234–241 (2015).
32. X. Zeng et al., "Physics informed neural networks for electric field distribution characteristics analysis," *J. Phys. D: Appl. Phys.* **56**, 165202 (2023).
33. M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *J. Comput. Phys.* **378**, 686–707 (2019).
34. L. Lu et al., "DeepXDE: a deep learning library for solving differential equations," *SIAM Rev.* **63**(1), 208–228 (2021).
35. A. G. Baydin et al., "Automatic differentiation in machine learning: a survey," *J. Mach. Learn. Res.* **18**(1), 5595–5637 (2018).
36. I. Siffa et al., "Towards a machine-learned Poisson solver for low-temperature plasma simulations in complex geometries," arXiv:2306.07604 (2023).
37. G. Huang et al., "Densely connected convolutional networks," in *IEEE Conf. Comput. Vision and Pattern Recognit. (CVPR)*, Honolulu, Hawaii, pp. 2261–2269 (2017).
38. A. Vaswani et al., "Attention is all you need," in *Adv. Neural Inf. Process. Syst. 30*, Curran Associates, Inc. (2017).
39. J. Chen et al., "TransUNet: transformers make strong encoders for medical image segmentation," arXiv:2102.04306 (2021).
40. Z. Wang et al., "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.* **13**, 600–612 (2004).
41. M. Desai et al., "Deep-learning architecture-based approach for 2-D-simulation of microwave plasma interaction," *IEEE Trans. Microwave Theory Tech.* **70**(12), 5359–5368 (2022).

42. N. Rahaman et al., "On the spectral bias of neural networks," in *Proc. 36th Int. Conf. Mach. Learn., PMLR*, pp. 5301–5310 (2019).
43. A. K. Verma et al., "Surrogate models for low temperature plasma simulations with deep learning," in *Bull. Am. Phys. Soc., (Virt. Conf.)*, APS, p. PR24.003 (2021).
44. M. Ichikawa et al., "Construction of a surrogate model for low-temperature plasma simulation using machine learning," in *69th Jpn. Soc. Appl. Phys. (JSAP) Spring Meeting 2022* (2022).
45. M. J. Kushner, "Hybrid modelling of low temperature plasmas for fundamental investigations and equipment design," *J. Phys. D: Appl. Phys.* **42**(19), 194013 (2009).
46. J. Ko et al., "Computational approach for plasma process optimization combined with deep learning model," *J. Phys. D: Appl. Phys.* **56**, 344001 (2023).
47. Y.-T. Zhang, S.-H. Gao, and F. Ai, "Efficient numerical simulation of atmospheric pulsed discharges by introducing deep learning," *Front. Phys.* **11**, 1125548 (2023).
48. Y.-T. Zhang, S.-H. Gao, and Y.-Y. Zhu, "Efficient numerical simulation on dielectric barrier discharges at atmospheric pressure integrated by deep neural network," *J. Appl. Phys.* **133**, 053303 (2023).
49. X.-C. Wang and Y.-T. Zhang, "Modeling of discharge characteristics and plasma chemistry in atmospheric CO₂ pulsed plasmas employing deep neural network," *J. Appl. Phys.* **133**, 143301 (2023).
50. S. Kawaguchi et al., "Deep learning for solving the Boltzmann equation of electrons in weakly ionized plasma," *Plasma Sources Sci. Technol.* **29**, 025021 (2020).
51. S. Kawaguchi and T. Murakami, "Physics-informed neural networks for solving the Boltzmann equation of the electron velocity distribution function in weakly ionized plasmas," *Jpn. J. Appl. Phys.* **61**, 086002 (2022).
52. S. Wang, Y. Teng, and P. Perdikaris, "Understanding and mitigating gradient flow pathologies in physics-informed neural networks," *SIAM J. Sci. Comput.* **43**, A3055–A3081 (2021).
53. J. S. Kim et al., "Numerical strategy for solving the Boltzmann equation with variable E/N using physics-informed neural networks," *J. Phys. D: Appl. Phys.* **56**, 344002 (2023).
54. L. Zhong, Q. Gu, and B. Wu, "Deep learning for thermal plasma simulation: solving 1-D arc model as an example," *Comput. Phys. Commun.* **257**, 107496 (2020).
55. L. Zhong, B. Wu, and Y. Wang, "Low-temperature plasma simulation based on physics-informed neural networks: frameworks and preliminary applications," *Phys. Fluids (1994)* **34**, 087116 (2022).
56. L. Zhong, B. Wu, and Y. Wang, "Accelerating physics-informed neural network based 1D arc simulation by meta learning," *J. Phys. D: Appl. Phys.* **56**, 074006 (2023).
57. A. F. Psaros, K. Kawaguchi, and G. E. Karniadakis, "Meta-learning PINN loss functions," *J. Comput. Phys.* **458**, 111121 (2022).
58. S. Kawaguchi, K. Takahashi, and K. Satoh, "Data-driven discovery of electron continuity equations in electron swarm map for determining electron transport coefficients in argon," *J. Phys. D: Appl. Phys.* **56**, 244003 (2023).
59. T. Xiao and D. Ni, "Multiscale modeling and recurrent neural network based optimization of a plasma etch process," *Processes* **9**, 151 (2021).
60. T. Xiao et al., "Recurrent neural-network-based model predictive control of a plasma etch process," *Ind. Eng. Chem. Res.* **61**, 638–652 (2022).
61. "Report on science challenges and research opportunities for plasma applications in microelectronics," Technical Report, Department of Energy Office of Science Fusion Energy Sciences Workshop (2023).
62. M. M. Turner, "Uncertainty and sensitivity analysis in complex plasma chemistry models," *Plasma Sources Sci. Technol.* **25**(1), 015003 (2016).
63. L. L. Alves et al., "Foundations of plasma standards," *Plasma Sources Sci. Technol.* **32**, 023001 (2023).
64. W. L. Morgan, "Use of numerical optimization algorithms to obtain cross sections from electron swarm data," *Phys. Rev. A* **44**, 1677–1681 (1991).
65. P. W. Stokes et al., "Determining cross sections from transport coefficients using deep neural networks," *Plasma Sources Sci. Technol.* **29**, 055009 (2020).
66. P. W. Stokes et al., "An improved set of electron-THFA cross sections refined through a neural network-based analysis of swarm data," *J. Chem. Phys.* **154**, 084306 (2021).
67. LXCat-Team, <https://nl.lxcat.net/home/> (2023).
68. V. Jetly and B. Chaudhury, "Extracting electron scattering cross sections from swarm data using deep neural networks," *Mach. Learn.: Sci. Technol.* **2**, 035025 (2021).
69. G. J. M. Hagelaar and L. C. Pitchford, "Solving the Boltzmann equation to obtain electron transport coefficients and rate coefficients for fluid models," *Plasma Sources Sci. Technol.* **14**, 722 (2005).
70. L. Zhong, "Fast prediction of electron-impact ionization cross sections of large molecules via machine learning," *J. Appl. Phys.* **125**, 183302 (2019).
71. A. L. Harris and J. Nepomuceno, "A data-driven machine learning approach for electron-molecule ionization cross sections," arXiv:2309.02538 (2023).
72. D. Reiser, A. von Keudell, and T. Urbanietz, "Determining chemical reaction systems in plasma-assisted conversion of methane using genetic algorithms," *Plasma Chem. Plasma Process.* **41**, 793–813 (2021).

73. M. Hanicinec, S. Mohr, and J. Tennyson, "A regression model for plasma reaction kinetics," *J. Phys. D: Appl. Phys.* **56**, 374001 (2023).
74. J. Tennyson et al., "QDB: a new database of plasma chemistries and reactions," *Plasma Sources Sci. Technol.* **26**(5), 055014 (2017).
75. V. Wakelam et al., "A kinetic database for astrochemistry (KIDA)," *Astrophys. J. Suppl. Ser.* **199**, 21 (2012).
76. J.-H. Park et al., "A new version of the plasma database for plasma physics in the data center for plasma properties," *Appl. Sci. Conver. Technol.* **29**, 5–9 (2020).
77. D. McElroy et al., "The UMIST database for astrochemistry 2012," *Astron. Astrophys.* **550**, A36 (2013).
78. I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," *Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci.* **374**, 20150202 (2016).
79. U. Maas and S. B. Pope, "Simplifying chemical kinetics: intrinsic low-dimensional manifolds in composition space," *Combust. Flame* **88**, 239–264 (1992).
80. K. Peerenboom et al., "Dimension reduction of non-equilibrium plasma kinetic models using principal component analysis," *Plasma Sources Sci. Technol.* **24**, 025004 (2015).
81. T. Rehman et al., "Simplifying plasma chemistry via ILDM," *J. Phys. Conf. Ser.* **682**, 012035 (2016).
82. O. Sakai et al., "Analysis of weblike network structures of directed graphs for chemical reactions in methane plasmas," *AIP Adv.* **5**, 107140 (2015).
83. O. Sakai, S. Kawaguchi, and T. Murakami, "Complexity visualization, dataset acquisition, and machine-learning perspectives for low-temperature plasma: a review," *Jpn. J. Appl. Phys.* **61**, 070101 (2022).
84. S. R. Sun, H. X. Wang, and A. Bogaerts, "Chemistry reduction of complex CO₂ chemical kinetics: application to a gliding arc plasma," *Plasma Sources Sci. Technol.* **29**, 025012 (2020).
85. T. Murakami and O. Sakai, "Rescaling the complex network of low-temperature plasma chemistry through graph-theoretical analysis," *Plasma Sources Sci. Technol.* **29**, 115018 (2020).
86. Y. Mizui et al., "Graphical classification in multi-centrality-index diagrams for complex chemical networks," *Symmetry* **9**, 309 (2017).
87. T. D. Holmes, R. H. Rothman, and W. B. Zimmerman, "Graph theory applied to plasma chemical reaction engineering," *Plasma Chem. Plasma Process.* **41**, 531–557 (2021).
88. S. Venturi et al., "An uncertainty-aware strategy for plasma mechanism reduction with directed weighted graphs," *Phys. Plasmas* **30**, 043904 (2023).
89. Y. Zhu et al., "Tailoring electric field signals of nonequilibrium discharges by the deep learning method and physical corrections," *Plasma Processes Polym.* **19**(3), e2100155 (2022).
90. J. Pan et al., "Deep learning-assisted pulsed discharge plasma catalysis modeling," *Energy Convers. Manage.* **277**, 116620 (2023).
91. J. P. Biersack and L. G. Haggmark, "A Monte Carlo computer program for the transport of energetic ions in amorphous targets," *Nucl. Instrum. Methods* **174**, 257 (1980).
92. Y. Ding et al., "Microscopic modeling and optimal operation of plasma enhanced atomic layer deposition," *Chem. Eng. Res. Des.* **159**, 439–454 (2020).
93. F. Burden and D. Winkler, "Bayesian regularization of neural networks," in *Artificial Neural Networks: Methods and Applications*, D. J. Livingstone, Ed., Methods in Molecular Biology, pp. 23–42, Humana Press, Totowa, New Jersey (2009).
94. Y. Ding et al., "Machine learning-based modeling and operation of plasma-enhanced atomic layer deposition of hafnium oxide thin films," *Comput. Chem. Eng.* **144**, 107148 (2021).
95. R. Sheil et al., "Precise control of nanoscale Cu etching via gas-phase oxidation and chemical complexation," *J. Phys. Chem. C* **125**, 1819–1832 (2021).
96. Y. Xia and P. Sautet, "Plasma oxidation of copper: molecular dynamics study with neural network potentials," *ACS Nano* **16**, 20680–20692 (2022).
97. H. M. Aktulga et al., "Parallel reactive molecular dynamics: numerical methods and algorithmic techniques," *Parallel Comput.* **38**, 245–259 (2012).
98. W. Zhu et al., "Development of a reactive force field for simulations on the catalytic conversion of C/H/O molecules on Cu-metal and Cu-oxide surfaces and application to Cu/CuO-based chemical looping," *J. Phys. Chem. C* **124**, 12512–12520 (2020).
99. M. Wan et al., "deep learning-assisted investigation of electric field–dipole effects on catalytic ammonia synthesis," *JACS Au* **2**, 1338–1349 (2022).
100. F. Scarselli et al., "The graph neural network model," *IEEE Trans. Neural Networks* **20**, 61–80 (2009).
101. Z. Wu et al., "A comprehensive survey on graph neural networks," *IEEE Trans. Neural Networks Learn. Syst.* **32**, 4–24 (2021).
102. T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Int. Conf. Learn. Represent.* (2022).
103. Y. M. Kedalo, A. A. Knizhnik, and B. V. Potapkin, "Applicability of the Fridman–Macheret α -model to heterogeneous processes in the case of dissociative adsorption of N₂ on the Ru surface," *J. Phys. Chem. C* **127**, 11536–11541 (2023).

104. A. Fridman, *Plasma Chemistry*, Cambridge University Press, New York (2008).
105. A. P. Bartók, R. Kondor, and G. Csányi, “On representing chemical environments,” *Phys. Rev. B* **87**, 184115 (2013).
106. K. M. Bal and E. C. Neyts, “Quantifying the impact of vibrational nonequilibrium in plasma catalysis: insights from a molecular dynamics model of dissociative chemisorption,” *J. Phys. D: Appl. Phys.* **54**, 394004 (2021).
107. D. Mendels, G. Piccini, and M. Parrinello, “Collective variables from local fluctuations,” *J. Phys. Chem. Lett.* **9**, 2776–2781 (2018).
108. O. Valsson and M. Parrinello, “Variational approach to enhanced sampling and free energy calculations,” *Phys. Rev. Lett.* **113**, 090601 (2014).
109. A. Barducci, G. Bussi, and M. Parrinello, “Well-tempered metadynamics: a smoothly converging and tunable free-energy method,” *Phys. Rev. Lett.* **100**, 020603 (2008).
110. A. Laio and M. Parrinello, “Escaping free-energy minima,” *Proc. Natl. Acad. Sci. U. S. A.* **99**, 12562–12566 (2002).
111. M. Siron, N. Chandrasekhar, and K. A. Persson, “Enabling automated high-throughput density functional theory studies of amorphous material surface reactions,” *Comput. Mater. Sci.* **226**, 112192 (2023).
112. A. Hamedani et al., “Primary radiation damage in silicon from the viewpoint of a machine learning interatomic potential,” *Phys. Rev. Mater.* **5**, 114603 (2021).
113. A. P. Bartók et al., “Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons,” *Phys. Rev. Lett.* **104**, 136403 (2010).
114. K. Nordlund, N. Runeberg, and D. Sundholm, “Repulsive interatomic potentials calculated using Hartree-Fock and density-functional theory methods,” *Nucl. Instrum. Methods Phys. Res., Sect. B* **132**, 45–54 (1997).
115. F. H. Stillinger and T. A. Weber, “Computer simulation of local order in condensed phases of silicon,” *Phys. Rev. B* **31**, 5262–5271 (1985).
116. J. Tersoff, “New empirical approach for the structure and energy of covalent systems,” *Phys. Rev. B* **37**, 6991–7000 (1988).
117. A. Hamedani et al., “Insights into the primary radiation damage of silicon by a machine learning interatomic potential,” *Mater. Res. Lett.* **8**, 364–372 (2020).
118. Y. Yamamura and H. Tawara, “Energy dependence of ion-induced sputtering yields from monatomic solids at normal incidence,” *At. Data Nucl. Data Tables* **62**(2), 149 (1996).
119. P. Phadke et al., “Sputter yields of monoatomic solids by Ar and Ne ions near the threshold: a Bayesian analysis of the Yamamura Model,” *Nucl. Instrum. Methods Phys. Res., Sect. B* **520**, 29–39 (2022).
120. W. Möller and W. Eckstein, “Tridyn: a TRIM simulation code including dynamic composition changes,” *Nucl. Instrum. Methods Phys. Res. Sect. B* **2**, 814 (1984).
121. W. Eckstein and R. Preuss, “New fit formulae for the sputtering yield,” *J. Nucl. Mater.* **320**, 209 (2003).
122. R. Preuss, R. Arredondo, and U. von Toussaint, “Bayesian determination of parameters for plasma-wall interactions,” *Entropy* **21**(12), 1175 (2019).
123. W. Eckstein et al., “SDTrimSP: Ein Monte-Carlo Code zur Berechnung von Stossereignissen in ungeordneten Targets,” Technical Report 12/3, Max-Planck-Institut für Plasmaphysik, Garching (2007).
124. N. Wiener, “The homogeneous chaos,” *Am. J. Math.* **60**(4), 897–936 (1938).
125. F. Krüger, T. Gergs, and J. Trieschmann, “Machine learning plasma-surface interface for coupling sputtering and gas-phase transport simulations,” *Plasma Sources Sci. Technol.* **28**(3), 035002 (2019).
126. T. Gergs, B. Borislavov, and J. Trieschmann, “Efficient plasma-surface interaction surrogate model for sputtering processes based on autoencoder neural networks,” *J. Vac. Sci. Technol. B Nanotechnol. Microelectron.* **40**(1), 012802 (2022).
127. D. P. Kingma and M. Welling, “Auto-encoding variational Bayes,” in *Int. Conf. Learn. Represent. 2014* (2013).
128. D. J. Rezende, S. Mohamed, and D. Wierstra, “Stochastic backpropagation and approximate inference in deep generative models,” in *Proc. Mach. Learn. Res.*, Vol. **32**, pp. 1278–1286 (2014).
129. I. Higgins et al., “ β -VAE: learning basic visual concepts with a constrained variational framework,” in *Int. Conf. Learn. Represent.*, p. 22 (2017).
130. B. Kim et al., “Deep neural network-based reduced-order modeling of ion–surface interactions combined with molecular dynamics simulation,” *J. Phys. D: Appl. Phys.* **56**, 384005 (2023).
131. P. Vincent et al., “Extracting and composing robust features with denoising autoencoders,” in *Proc. 25th Int. Conf. Mach. Learn.*, ACM Press, Helsinki, pp. 1096–1103 (2008).
132. K. Sohn, H. Lee, and X. Yan, “Learning structured output representation using deep conditional generative models,” in *Adv. Neural Inf. Process. Syst.* **28** (2015).
133. T. Gergs, T. Mussenbrock, and J. Trieschmann, “Physics-separating artificial neural networks for predicting sputtering and thin film deposition of AlN in Ar/N₂ discharges on experimental timescales,” *J. Phys. D: Appl. Phys.* **56**(19), 194001 (2023).

134. K. M. Bal and E. C. Neyts, "On the time scale associated with Monte Carlo simulations," *J. Chem. Phys.* **141**, 204104 (2014).
135. M. J. Mees et al., "Uniform-acceptance force-bias Monte Carlo method with time scale to study solid-state diffusion," *Phys. Rev. B* **85**, 134301 (2012).
136. E. C. Neyts and A. Bogaerts, "Combining molecular dynamics with Monte Carlo simulations: implementations and applications," *Theor. Chem. Acc.* **132**(2), 1320 (2013).
137. T. Gergs, T. Mussenbrock, and J. Trieschmann, "Molecular dynamics study on the role of Ar ions in the sputter deposition of Al thin films," *J. Appl. Phys.* **132**(6), 063302 (2022).
138. T. Gergs, T. Mussenbrock, and J. Trieschmann, "Physics-separating artificial neural networks for predicting initial stages of Al sputtering and thin film deposition in Ar plasma discharges," *J. Phys. D: Appl. Phys.* **56**(8), 084003 (2023).
139. A. Jaramillo-Botero, S. Naserifar, and W. A. I. Goddard, "General multiobjective force field optimization framework, with application to reactive force fields for silicon carbide," *J. Chem. Theory Comput.* **10**, 1426–1439 (2014).
140. T. Gergs, T. Mussenbrock, and J. Trieschmann, "Charge-optimized many-body interaction potential for AlN revisited to explore plasma–surface interactions," *Sci. Rep.* **13**(1), 5287 (2023).
141. S. Berg et al., "Modeling of reactive sputtering of compound materials," *J. Vac. Sci. Technol., A* **5**, 202 (1987).
142. C. Woelfel, P. Awakowicz, and J. Lunze, "Model reduction and identification of nonlinear reactive sputter processes," *IFAC-PapersOnLine* **50**, 13728–13734 (2017).
143. C. Woelfel et al., "Model approximation and stabilization of reactive sputter processes," *J. Process Control* **83**, 121–128 (2019).
144. C. Woelfel et al., "Plasma state control of reactive sputter processes," in *Proc. 6th Int. Conf. Control, Dyn. Syst., and Rob. (CDSR'19)*, Ottawa, p. 107 (2019).
145. C. Woelfel et al., "Control-oriented plasma modeling and controller design for reactive sputtering," *IFAC J. Syst. Control* **16**, 100142 (2021).
146. D. Gidon, D. B. Graves, and A. Mesbah, "Predictive control of 2D spatial thermal dose delivery in atmospheric pressure plasma jets," *Plasma Sources Sci. Technol.* **28**(8), 085001 (2019).
147. D. Gidon et al., "Data-driven LPV model predictive control of a cold atmospheric plasma jet for biomaterials processing," *Control Eng. Pract.* **109**, 104725 (2021).
148. D. Rodrigues, K. J. Chan, and A. Mesbah, "Data-driven adaptive optimal control under model uncertainty: an application to cold atmospheric plasmas," *IEEE Trans. Control Syst. Technol.* **31**, 55–69 (2023).
149. G. Berkooz, P. Holmes, and J. L. Lumley, "The proper orthogonal decomposition in the analysis of turbulent flows," *Annu. Rev. Fluid Mech.* **25**(1), 539–575 (1993).
150. C. M. Greve, M. Majji, and K. Hara, "Real-time state estimation of low-frequency plasma oscillations in Hall effect thrusters," *Phys. Plasmas* **28**, 093509 (2021).
151. C. M. Greve and K. Hara, "Estimation of plasma properties using an extended Kalman filter with plasma global models," *J. Phys. D: Appl. Phys.* **55**, 255201 (2022).
152. K. J. Kanarik et al., "Human–machine collaboration for improving semiconductor process development," *Nature* **616**, 707–711 (2023).
153. H. Jaeger, "A tutorial on training recurrent neural networks, covering BPPT, RTRL, EKF and the "echo state network" approach," Technical Report, GMD-Forschungszentrum Informationstechnik Bonn, Bonn (2002).
154. S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.* **9**, 1735–1780 (1997).
155. F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: continual prediction with LSTM," *Neural Comput.* **12**, 2451–2471 (2000).
156. G. Borrelli et al., "Predicting the temporal dynamics of turbulent channels through deep learning," *Int. J. Heat Fluid Flow* **96**, 109010 (2022).
157. W. Zheng et al., "Disruption predictor based on neural network and anomaly detection on J-TEXT," *Plasma Phys. Control. Fusion* **62**, 045012 (2020).
158. B. H. Guo et al., "Disruption prediction on EAST tokamak using a deep learning algorithm," *Plasma Phys. Control. Fusion* **63**, 115007 (2021).
159. Y. Liu, J. N. Kutz, and S. L. Brunton, "Hierarchical deep learning of multiscale differential equation time-steppers," *Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci.* **380**, 20210200 (2022).
160. H. Jaeger and H. Haas, "Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication," *Science* **304**, 78–80 (2004).
161. N. A. K. Doan, W. Polifke, and L. Magri, "Physics-informed echo state networks," *J. Comput. Sci.* **47**, 101237 (2020).
162. S. Cuomo et al., "Scientific machine learning through physics–informed neural networks: where we are and what's next," *J. Sci. Comput.* **92**, 88 (2022).
163. Y. Wang and L. Zhong, "NAS-PINN: neural architecture search-guided physics-informed neural network for solving PDEs," *J. Comput. Phys.* **496**, 112603 (2024).

164. L. Lu et al., “Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators,” *Nat. Mach. Intell.* **3**, 218–229 (2021).
165. A. Anandkumar et al., “Neural operator: graph kernel network for partial differential equations,” in *Workshop on Integration of Deep Neural Models and Differential Equations, Int. Conf. Learn. Represent.* (2020).
166. Z. Li et al., “Fourier neural operator for parametric partial differential equations,” in *Int. Conf. Learn. Represent.* (2021).
167. Q. Cao, S. Goswami, and G. E. Karniadakis, “LNO: Laplace neural operator for solving differential equations,” arXiv:2303.10528 (2023).
168. N. Geneva and N. Zabarar, “Transformers for modeling physical systems,” *Neural Networks* **146**, 272–289 (2022).
169. S. L. Brunton, J. L. Proctor, and J. N. Kutz, “Discovering governing equations from data by sparse identification of nonlinear dynamical systems,” *Proc. Natl. Acad. Sci. U. S. A.* **113**, 3932–3937 (2016).
170. S. H. Rudy et al., “Data-driven discovery of partial differential equations,” *Sci. Adv.* **3**, e1602614 (2017).
171. E. P. Alves and F. Fiuza, “Data-driven discovery of reduced plasma physics models from fully kinetic simulations,” *Phys. Rev. Res.* **4**, 033192 (2022).
172. A. A. Kaptanoglu et al., “Permanent-magnet optimization for stellarators as sparse regression,” *Phys. Rev. Appl.* **18**, 044006 (2022).
173. A. A. Kaptanoglu et al., “Sparse regression for plasma physics,” *Phys. Plasmas* **30**, 033906 (2023).
174. E. Kaiser, J. N. Kutz, and S. L. Brunton, “Sparse identification of nonlinear dynamics for model predictive control in the low-data limit,” *Proc. R. Soc. A: Math. Phys. Eng. Sci.* **474**, 20180335 (2018).
175. J. D. Lore et al., “Time-dependent SOLPS-ITER simulations of the Tokamak plasma boundary for model predictive control using SINDy*,” *Nucl. Fusion* **63**, 046015 (2023).
176. B. Settles, “Active learning literature survey,” technical Report, University of Wisconsin-Madison Department of Computer Sciences (2009).
177. A. Diaw et al., “Multiscale simulation of plasma flows using active learning,” *Phys. Rev. E* **102**(2), 023310 (2020).
178. S. Bozinovski, “Reminder of the first paper on transfer learning in neural networks, 1976,” *Informatica* **44**, 291–302 (2020).
179. K. D. Humbird and J. L. Peterson, “Transfer learning driven design optimization for inertial confinement fusion,” *Phys. Plasmas* **29**, 102701 (2022).
180. L. H. Gilpin et al., “Explaining explanations: an overview of interpretability of machine learning,” in *IEEE 5th Int. Conf. Data Sci. and Adv. Anal. (DSAA)*, pp. 80–89 (2018).
181. M. T. Ribeiro, S. Singh, and C. Guestrin, ““Why should I trust you?” Explaining the predictions of any classifier,” in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. and Data Mining, KDD '16*, Association for Computing Machinery, New York, pp. 1135–1144 (2016).
182. C.-Y. Wang, T.-S. Ko, and C.-C. Hsu, “Machine learning with explainable artificial intelligence vision for characterization of solution conductivity using optical emission spectroscopy of plasma in aqueous solution,” *Plasma Processes Polym.* **18**(12), 2100096 (2021).
183. S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in *Proc. 31st Int. Conf. Neural Inf. Process. Syst., NIPS'17*, Curran Associates Inc., Red Hook, New York, pp. 4768–4777 (2017).
184. Y.-J. Wang and C.-T. Lin, “Runge-Kutta neural network for identification of dynamical systems in high accuracy,” *IEEE Trans. Neural Networks* **9**, 294–307 (1998).
185. M. Zhu, B. Chang, and C. Fu, “Convolutional neural networks combined with Runge-Kutta methods,” *Neural Comput. Appl.* **35**, 1629–1643 (2023).
186. S. De Pascuale, D. L. Green, and J. D. Lore, “Data-driven linear time advance operators for the acceleration of plasma physics simulation,” *Phys. Plasmas* **29**, 113903 (2022).
187. J. C. S. Kadupitiya, G. C. Fox, and V. Jadhao, “Solving Newton’s equations of motion with large timesteps using recurrent neural networks based operators,” *Mach. Learn.: Sci. Technol.* **3**, 025002 (2022).
188. A. Sanchez-Gonzalez et al., “Graph networks as learnable physics engines for inference and control,” in *Proc. 35th Int. Conf. Mach. Learn., PMLR*, pp. 4470–4479 (2018).
189. A. Sanchez-Gonzalez et al., “Hamiltonian graph networks with ODE integrators,” in *2nd Workshop on Mach. Learn. Physical Sciences, Neural Inf. Process. Syst.*, Vancouver, Canada, (2019).
190. Z. Li et al., “Graph neural networks accelerated molecular dynamics,” *J. Chem. Phys.* **156**(14), 144103 (2022).
191. S. L. Brunton, B. R. Noack, and P. Koumoutsakos, “Machine learning for fluid mechanics,” *Annu. Rev. Fluid Mech.* **52**(1), 477–508 (2020).
192. R. Vinuesa and S. L. Brunton, “Enhancing computational fluid dynamics with machine learning,” *Nat. Comput. Sci.* **2**, 358–366 (2022).

193. S. Park et al., “Predictive control of the plasma processes in the OLED display mass production referring to the discontinuity qualifying PI-VM,” *Phys. Plasmas* **27**, 083507 (2020).
194. J. Degraeve et al., “Magnetic control of tokamak plasmas through deep reinforcement learning,” *Nature* **602**, 414–419 (2022).
195. T. Grassi et al., “Reducing the complexity of chemical networks via interpretable autoencoders,” *Astron. Astrophys.* **668**, A139 (2022).
196. K. S. Tang and M. Turk, “Reduced order model for chemical kinetics: a case study with primordial chemical network,” in *Workshop on Mach. Learn. Astrophys., Int. Conf. Mach. Learn.* Baltimore (2022).
197. E. Chiavazzo et al., “Reduced models in chemical kinetics via nonlinear data-mining,” *Processes* **2**, 112–140 (2014).
198. S. Goswami et al., “Learning stiff chemical kinetics using extended deep neural operators,” *Comput. Meth. Appl. Mech. Eng.* **419**, 116674 (2024).
199. L. Campoli, E. Kustova, and P. Maltseva, “Assessment of machine learning methods for state-to-state approach in nonequilibrium flow simulations,” *Mathematics* **10**, 928 (2022).
200. M. Savarese et al., “Machine learning clustering algorithms for the automatic generation of chemical reactor networks from CFD simulations,” *Fuel* **343**, 127945 (2023).
201. R. Mao et al., “DeepFlame: a deep learning empowered open-source platform for reacting flow simulations,” *Comput. Phys. Commun.* **291**, 108842 (2023).
202. The OpenFOAM Foundation Ltd., “OpenFOAM: open source field operation and manipulation,” 2023, www.openfoam.org
203. R. Collobert, K. Kavukcuoglu, and C. Farabet, “Torch7: a Matlab-like environment for machine learning,” in *NIPS 2011 Workshop BigLearn* (2011).
204. D. G. Goodwin, “Cantera C++ user’s guide,” tech. rep., California Institute of Technology (2002).
205. T. Mou et al., “Bridging the complexity gap in computational heterogeneous catalysis with machine learning,” *Nat. Catal.* **6**, 122–136 (2023).
206. D. Chen, C. Shang, and Z.-P. Liu, “Machine-learning atomic simulation for heterogeneous catalysis,” *NPJ Comput. Mater.* **9**, 2 (2023).
207. A. V. Phelps and Z. L. Petrovic, “Cold-cathode discharges and breakdown in argon: surface and gas phase production of secondary electrons,” *Plasma Sources Sci. Technol.* **8**(3), R21 (1999).
208. M. Daksha et al., “A computationally assisted spectroscopic technique to measure secondary electron emission coefficients in radio frequency plasmas,” *J. Phys. D: Appl. Phys.* **49**(23), 234001 (2016).
209. C. Schulze, Z. Donkó, and J. Benedikt, “A computationally assisted technique to measure material-specific surface coefficients in capacitively coupled plasmas based on characteristics of the ion flux-energy distribution function,” *Plasma Sources Sci. Technol.* **31**, 105017 (2022).
210. H.-Y. Chang, A. Alvarado, and J. Marian, “Calculation of secondary electron emission yields from low-energy electron deposition in tungsten surfaces,” *Appl. Surf. Sci.* **450**, 190–199 (2018).
211. J. T. Su and W. A. Goddard, “Mechanisms of Auger-induced chemistry derived from wave packet dynamics,” *Proc. Natl. Acad. Sci. U. S. A.* **106**, 1001–1005 (2009).
212. M. Pamperin, F. X. Bronold, and H. Fehske, “Ion-induced secondary electron emission from metal surfaces,” *Plasma Sources Sci. Technol.* **27**(8), 084003 (2018).
213. F. X. Bronold and H. Fehske, “Invariant embedding approach to secondary electron emission from metals,” *J. Appl. Phys.* **131**, 113302 (2022).
214. J. F. Ziegler and J. P. Biersack, “The stopping and range of ions in matter,” in *Treatise on Heavy-Ion Science*, D. A. Bromley, Ed., pp. 93–129, Springer, New York (1985).
215. H. Wang et al., “Deep learning inter-atomic potential model for accurate irradiation damage simulations,” *Appl. Phys. Lett.* **114**, 244101 (2019).
216. H. Niu et al., “A machine-learning interatomic potential to understand primary radiation damage of silicon,” *Comput. Mater. Sci.* **218**, 111970 (2023).
217. M. A. Caro et al., “Growth mechanism and origin of high sp³ content in tetrahedral amorphous carbon,” *Phys. Rev. Lett.* **120**, 166101 (2018).
218. M. A. Caro et al., “Machine learning driven simulated deposition of carbon films: from low-density to diamondlike amorphous carbon,” *Phys. Rev. B* **102**, 174201 (2020).
219. A. P. Thompson et al., “LAMMPS: a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales,” *Comput. Phys. Commun.* **271**, 108171 (2022).
220. J. S. Lim et al., “Evolution of metastable structures at bimetallic surfaces from microscopy and machine-learning molecular dynamics,” *J. Am. Chem. Soc.* **142**, 15907–15916 (2020).
221. G. Henkelman and H. Jónsson, “A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives,” *J. Chem. Phys.* **111**, 7010–7022 (1999).
222. G. Henkelman, B. P. Uberuaga, and H. Jónsson, “A climbing image nudged elastic band method for finding saddle points and minimum energy paths,” *J. Chem. Phys.* **113**, 9901–9904 (2000).

223. G. Henkelman and H. Jónsson, "Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points," *J. Chem. Phys.* **113**, 9978–9985 (2000).
224. S. Franke et al., "Plasma-MDS, a metadata schema for plasma science with examples from plasma technology," *Sci. Data* **7**, 439 (2020).
225. E. Carbone et al., "Data needs for modeling low-temperature non-equilibrium plasmas: the LXCat project, history, perspectives and a tutorial," *Atoms* **9**, 16 (2021).
226. I. C. Siffa, J. Schäfer, and M. M. Becker, "Adamant: a JSON schema-based metadata editor for research data management workflows," *F1000Research* **11**, 475 (2022).
227. M. D. Wilkinson et al., "The FAIR guiding principles for scientific data management and stewardship," *Sci. Data* **3**, 160018 (2016).

Jan Trieschmann is a junior professor at Kiel University, Germany. He received his diploma and PhD degrees in electrical engineering and information technology from Ruhr University Bochum, in Germany, in 2012 and 2017, respectively. His current research interests include low-temperature plasma modeling and simulation, multiscale modeling using data-integrated schemes and machine learning, as well as the processing of memristive devices and systems and their application in neuromorphic engineering.

Luca Vialetto is a postdoctoral fellow at Stanford University, USA. He received his BSc and MSc degrees in physics from the University of Padua, in Italy, in 2015 and 2017, respectively. He received his PhD in applied physics from the Eindhoven University of Technology, The Netherlands, in 2021. His research interests include plasma physics and chemistry, swarm physics, and multi-scale modeling of plasmas for application in energy research.

Tobias Gergs is a postdoctoral researcher at Kiel University, Germany. He received his BSc, MSc, and PhD degrees in electrical engineering and information technology from Ruhr University Bochum, in Germany, in 2013, 2016, and 2023, respectively. His current research interests include plasma–surface interactions, neuromorphic engineering, molecular dynamics simulations, and machine learning.