Complexity visualization, dataset acquisition, and machine-learning perspectives for low-temperature plasma: a review

Osamu Sakai1,2* , Satoru Kawaguchi3,4 ©, and Tomoyuki Murakami3 ©

1Department of Electronic Systems Engineering, Faculty of Engineering, The University of Shiga Prefecture, 2500 Hassakacho, Hikone, Shiga 522-8533, Japan
2Regional ICT Research Center for Human, Industry and Future, The University of Shiga Prefecture, 2500 Hassakacho, Hikone, Shiga 522-8533, Japan
3Faculty of Science and Technology, Seikei University, 3-3-1 Kichijoji-Kitamachi, Musashino, Tokyo 180-8633, Japan
4Division of Information and Electronic Engineering, Graduate School of Engineering, Muraoran Institute of Technology, 21-1 Mizumoto, Muraoran, Hokkaido 050-8585, Japan

*E-mail: sakai.o@e.usp.ac.jp

Received April 2, 2022; revised May 19, 2022; accepted June 8, 2022; published online July 6, 2022

Low-temperature plasma plays various roles in industrial material processing as well as provides a number of scientific targets, both from theoretical and experimental points of view. Such rich features in variety are based on its complexities, arising from diverse parameters in constituent gas-phase species, working gas pressure, input energy density, and spatial boundaries. When we consider causalities in these complexities, direct application of machine-learning methods is not always possible since levels of complexities are so high in comparison with other scientific research targets. To overcome this difficulty, progresses in plasma diagnostics and data acquisition systems are inevitable, and the handling of a large number of data elements is one of the key issues for this purpose. In this topical review, we summarize previous and current achievements of visualization, acquisition, and analysis methods for complex plasma datasets which may open a scientific and technological category mixed with rapid machine-learning advancements and their relevant outcomes. Although these research trends are ongoing, many reports published so far have already convinced us of various expanding aspects of low-temperature plasma leading to the potential for scientific progress as well as developments of intellectual design in industrial plasma processes. © 2022 The Japan Society of Applied Physics

1. Introduction

Machine learning and its derivative schemes1–5) are currently popular in various industrial activities, which include the prediction of financial trends and analysis of consumer markets, and their applicability is absolutely wide with sufficiently-common algorithms and tools. This fact is also true up to now among many scientific categories in applied physics. For instance, material informatics based on machine learning5,6) works effectively for biomaterials, polymers, and solid-state materials. In cases where abundant datasets are available but no significant principle has been reported, machine learning techniques are powerful for finding possible optimized parameter sets or industrial solutions. We can find a good example in machine-learning application to bioinformatics,7,8) where genomics yields a huge number of datasets for genome sequences automatically. That is, one can obtain a clear and complete understanding of gene sequences, and of their resultant effects on life is possible to a certain extent, although biological processes of life still involve many veiled parts that remain unknown. In such a circumstance around bioinformatics, machine-learning methods are quite significant for confirming empirical laws that link genes and real life, and machine learning can play a key role in automatic predictions based on supervised classification and clustering. Here, we define machine learning as computation methods that are applicable in a semi-automatic manner (i.e. with very limited data handling and manual parameter tuning) to analyze a phenomenon with unknown and/or complicated underlying processes; it includes not only supervised learning like neural networks with one or multiple hidden layers,9,10) but also unsupervised learning without any training datasets.

Along with such trends around machine-learning-based informatics, plasmas may be possible targets in which typical or customized machine-learning methods are applicable. As shown later in this report, we found several studies on pioneering work for plasma phenomena analyzed in the scope of machine learning. One of the scientific motivations for performing machine learning for plasmas arises from various confusing outlooks and functions that plasmas exhibit. Such rich features in variety are based on its complexities: discharge species including atoms, molecules, radicals with ions and electrons, working gas pressures in more than 6-orders of magnitude,11,12) and various boundaries that may be solid-state material substrates,12,13) liquid surfaces12,14) and even soft matters in medical and agriculture treatments12,15,16) achieved by plasma. Accordingly, to perform optimization of external parameters without machine learning, the required knowledge is quite vast across many scientific categories like statistics, electrical engineering, inorganic and organic chemistry, and basic plasma physics. In contrast, when we use machine learning tools for plasmas (which have been already typical and suitable in some cases, as we will mention later), hypothetical information flow in a bundle system of mathematical functions is assumed; this system is given, for instance, as an artificial neural network (ANN)2,9,10) in simple supervised learning. Here, to fit output values in a given dataset, approximate functions are elaborately but automatically selected with a large number of tuned coefficients, which enables deliberate matching to complicated parameter dependences.

In principle, machine learning for plasma phenomena has close linkages to plasma diagnostics17,18) and numerical analysis19,20) machine learning uses datasets, which are obtained through plasma diagnostics and/or derived from numerical results. Several decades ago, by an available computational tool with insufficient calculation processors and small access memories, these deliberate tuning tasks in machine learning took so much time that machine-learning approaches were not practical. Then, although its framework like back-propagation algorithm was completed until 1990,2)
various efforts without machine learning have been performed on numerical calculations of differential equations in a model of plasma and experimental plasma diagnostics using elaborate designs, leading to scientific storage of rich understandings about internal plasma physics and/or chemistry processes. As described later, in conjunction with these historically-long efforts, machine learning with current calculation facilities can reinforce and provide some shortcuts for understanding underlying processes. We can say that replacement of such human consideration in diagnostics and numerical analysis by neural networks or other machine-learning tools makes the total procedure semi-automatic and useful in cases without sufficient professional human resources.

However, unfortunately, when we survey these previous attempts, i.e. diagnostics, numerical analysis, and machine-learning procedures to plasmas, we find some missing links. State-of-the-art diagnostics techniques are based on scientific rigorous evidence, and give us accurate plasma parameters; unfortunately, in most cases, the confirmed values are limited in a partial parameter domain (both in space and in a value range). Numerical computation works with high accuracy and abundant technical progress, although the validity range of the applied model on an actual experiment remains uncertain, resulting in obscure key processes analyzed in a complicated procedure. Machine learning can provide a suitable approximation, although performers hardly confirm the matching of approximated or extrapolated model(s) with actual internal events, which can be estimated as a black box. That is, the situations in plasma science and technology are not rigorously similar to the other categories in which machine learning has been quite successful relatively with ease; further efforts to achieve inside visualization of approximators are needed in cases of low-temperature plasma.

In this topical review, we survey the current status of machine-learning tools and their possible validities for plasma science and engineering and seek advanced methods to reinforce major drawbacks in machine learning for plasma science. To solve difficult and mysterious tasks in and around plasma science, we do not easily select machine learning which has been very popular both in scientific and industrial activities. In contrast, we keep our policy in which one applies machine learning tools if one certainly has a chance to obtain more rigorous analysis with small efforts to clarify underlying physical events. In the following sections, we look over the current status of machine learning in various aspects.

According to our survey performed so far, direct application of machine learning to specific low-temperature plasma processes is possible in limited cases, and scientific and/or technical additives in treatments work well to reinforce machine learning and to expand applicable cases. In most situations both in laboratory experiments and industrial fabrication practices, a particular additional superposition of exquisite care is usually required, and some examples of such tactics are summarized in this study. In a more general perspective, to overcome difficulties originating from plasma processes in which there exist too many degrees of freedom without adequate probing methods, complex network science may work well, whose insight is introduced to clarify underlying pathways in plasma reactors not only for reinforcement of machine learning but also for enhancement of visualization and understandings for complexities in plasma. We exclude a standing point in which one leaves internal processes in a black box and visualizes internal complexities using network topology. Interestingly, we intuitively find a common point between this method and the application of machine learning: the use of networks.

If we can successfully build up a hypothetical model based on machine-learning frameworks or visualize complex networks in underlying physical/chemical processes, we need
experimental evidence in the form of a dataset whose size is large with sufficient coverage of parameter ranges. Plasma diagnostics have historically played important roles to clarify and verify physical internal mechanisms in plasmas, and are triggered by such requirements in plasma machine-learning modelling, we are on the way to increasing measurement schemes for bigdata acquisition in a plasma. Here, whether corresponding techniques are state-of-the-art or not, we are aiming at (semi-)automatic collections of experimental data which cannot be detected using conventional electrical/optical probing methods inside plasma reactors. Recently, we successfully confirmed the operation of a tiny wireless electronic sensor, which is along the trend of IoT (Internet of Things) technology. Just by installing such multiple sensors near walls or in places where such devices do not disturb plasma activities in a reactor, data in time evolutions are automatically stored in a data server outside the vacuum chamber. Such a low-cost data acquisition system is quite suitable for our current purpose. In the following sections, we describe these approaches and mention other promising possibilities to build up suitable models for low-temperature plasma processes. One schematic view of our possible efforts can be displayed in Fig. 1. Here, on the side of the physical world, we set sensors and diagnostic tools around a targeted event with input-parameter setting and collection of outcomes. On the side of the virtual world or cyberspace, we develop a hypothetical approximator to reproduce the outcomes in the physical world.

In this report, we show a review of low-temperature plasma analysis from one point of view on information and communication technology (ICT), activated by general motivations in which machine-learning may be beneficial for further control in low-temperature plasma processes. To answer the question of whether machine-learning and other algorithm-based approaches are adjustable for analyses of low-temperature plasma or not, we overview features of a typical ANN, other supervised learning schemes, and unsupervised learning methods for candidates of machine learning. We also review the applicability of complex networks which are much more complicated than ANNs to cover the complexities of plasma in comparison with other media; this approach to complexity in plasma has just been proposed, and it may be available as a standard for plasma design when required in the near future. Then, a recent study on machine-learning-based analysis of cross-section analysis and other examples of supervised learning are reviewed, followed by an unsupervised-learning approach for plasma chemistry. Next, to increase substantially datasets for complex network visualization and supervised learning of low-temperature plasma, we show a sensor-installation approach based on the concept of IoT, and finally, we overview other promising reports that may support machine-learning and other algorithm-based analysis for plasma science in the near future.

2. Overview of plasma properties as targets of machine learning

2.1. Supervised learning based on ANNs

As the first case of machine learning in this report, we overview how a general ANN works in a regression or discriminant (classification) analysis, currently expanding in many topics in social, life, and physical science and engineering. A simple and small-size ANN is insufficient for modern technologies working for real scenes in societies, but the fundamentals of supervised-learning models based on ANNs, which include deep learning, are the same in almost all cases. We will mention later other methods of machine learning, like reinforcement learning and unsupervised learning, which are quite different from those based on neural networks. ANNs astonishingly work well in order to reproduce an outcome in a complicated system and to predict features of future events. When we can ask whether ANNs have the capability to work as a universal function hypothetically for various phenomena, there is a rigorously proven answer to this question, which is described below.

Figure 2(a) depicts a very simple ANN in a style with accuracy to mathematical formulation. This is a classical multilayer feedforward perceptron that includes one hidden layer. ANNs have more complicated with a larger number of inner elements, but this simple structure sufficiently implies the functions they possess. The style in this figure is a network or a graph, in which constituents are nodes shown as circles and edges as connections in between, in the terminology of network science, with inlet and output ports. Edges in ANNs are represented by arrows since they are unidirectional in information flows. Every edge has a weight that multiplies flowing data through it, and inside every node, the summation of weighted flows in the previous layer is converted into one temporal output via a nonlinear function \( A(z) \) with a threshold value; one of the typical nonlinear mappings is the sigmoid or logistic function \( \sigma \), given as:

\[
A(z) \mapsto \sigma(z - \theta) = \frac{1}{1 + e^{-\alpha(z - \theta)}}, \quad (1)
\]

where \( \theta \) is a threshold for transition from 0 to 1, and \( \alpha \) represents the steepness of the transition. If \( z \ll \theta \), \( \sigma(z - \theta) = 0 \), while \( \sigma(z - \theta) = 1 \) if \( z \gg \theta \). The range around \( z \sim \theta \) is a transition zone, and between two regions \( \sigma \) varies continuously and smoothly (i.e. with the continuous change of its derivative), where \( \alpha \) determines varying gradients on these slopes.

By looking at Eq. (1) and Fig. 2(a), we should understand the following indications before any practices of supervised learning. First, we find plenty of weights and thresholds that should be fixed for the completion of an ANN model. For instance, in the case exemplified in Fig. 2(a), there are 13 edge weights that should converge throughout the training process of supervised learning, whose total flow is illustrated in Fig. 2(b), and each of them is linked to \( \sigma \) at a corresponding node. To determine all of the weight values, we need a sufficient number of datasets. Second, if we superpose logistic or sigmoid functions for \( \sigma \), an approximately arbitrary shape of mapping from inputs to outputs can be designed, as described in the following. Suppose that one threshold value for an activation function \( \theta_1 \) is slightly larger than the other one \( \theta_2 \). Then, the synthesized function \( f(z) = \sigma(z - \theta_1) - \sigma(z - \theta_2) \) yields a pulse-shaped curve where \( f(z) = 1 \) in a narrow range between \( \theta_1 \) and \( \theta_2 \), and \( f(z) = 0 \) when \( z \ll \theta \) or \( z \gg \theta \). This indicates that if we superpose \( f(z) \) by varying \( \theta_1 \) and \( \theta_2 \), we can create
arbitrary shapes along with the limitations coming from finite integration level of superposition. For our reference, we note that a superposition of linear functions creates only a linear function, and nonlinearity in a value-mapping process is essential for the achievements of universal approximators, which leads to the generalization of ANNs for untested datasets; this nonlinear function is achieved by the sigmoid function in Fig. 2(a), but other functions like binary-step and hyperbolic-tangent functions are successfully applied in practical performances.

Fig. 2. (Color online) (a) Simple form of an artificial neural network (ANN) model with mathematical functions. Between input and output layers, hidden layer(s) exist and functionalize this model as a universal approximator. A sigmoid function is shown as an activation function, which is not limited to this type. (b) Typical data processing in supervised learning. (c) Typical data processing in unsupervised learning.
The aforementioned two facts of ANNs indicate that, after successful supervised learning using datasets in the training process, we can derive a suitable function that reproduces a given experimental result. This success in the learning process depends on the balance between the complexity of the function and the available number of datasets; if the phenomenon is so complicated, with a number of abrupt changes and transitions, the appropriate approximator will be completed after training by available datasets with inputs (i.e., external parameters in an experimental setup) and outputs (resulting outcomes, like manufactured nano-meter profiles in dry etching). For instance, since 13 parameters have been found to be fixed as edge weights and thresholds in activation functions in the case of the ANN in Fig. 2(a), we have to collect more than 13 datasets for training ANNs at least in a mathematical point of views, and normally much more for automatic tuning in algorithmic iteration procedures if a given system is fairly complex. On the other hand, the required number of datasets is not necessarily determined only by unknown edge weights in an ANN model; in some cases, there have been cause-effect connections among edge weights, and a smaller number of datasets is possible to appropriate a suitable model; the influences of cause-effect interconnections in the system will be discussed in the later sections. Hereafter we call datasets for tuning ANN parameters as training datasets and those for unbiased evaluation of a final ANN model as test datasets. If optimization of ANN structure such as the number of hidden layers and nodes are required to suppress overfitting, which induces too complicated curve fitting, validation datasets are used for in-advance rough framing of the model in addition, as displayed in Fig. 2(b); overfitting is caused from various reasons, like by too many datasets and by too many nodes in the hidden layers.

There have been a number of reports which demonstrate ANN-based machine learning for analysis of experimental results in plasma experiments for material and device processing from the 1990s. A typical example of the ANN analysis is reported in an experimental study of atmospheric plasma spraying. Using experimental setup parameters as ANN inputs, where the parameters are arc current, gas flows, and injector size, the cross-sectional spatial parameters in the deposited spraying profiles are predicted as outputs of the ANN with two hidden layers in which seven and four nodes are set in the former and latter layer. The total datasets are collected from 22 cases with multiplication by a factor of seven. Some recent studies for these several years will be reviewed later. Dry etching was a target modelled by the ANNs in 1994, and 53 datasets are used for training and testing it. The inputs were composed of parameters in the external setup such as ignition power, electrode spacing, gas pressure, and flows, and the outputs are set to process output performances such as etch rate, etch uniformity, and selectivity with respect to oxide and photoresist. Plasma chemical vapor deposition is another target of the ANNs, and using 19 experimental datasets with fractional factorial design, which substantially increases the dataset volume. The inputs are the external setting parameters such as substrate temperature, ignition power, gas pressure and flows, whereas the outputs are set to be deposition rate, film-quality values like refractive index, permittivity, film stress, wet etch rate, and silanol/water concentration.

As we have reviewed above, overcoming dataset deficiency is a crucial task to attain a universal approximator that can create a proper prediction of outputs in cases with unknown experimental parameters. However, such deficient datasets may be useful for other purposes that can be handled automatically. One possibility is to use them for inverse problems in physics-informed neural networks. This analysis enables us not to obtain a universal approximator but to perform one kind of unsupervised learning in which unknown information in a process is approximately and automatically predicted with a limited number of (or only one) dataset; the general scheme of unsupervised learning is illustrated in Fig. 2(c) and described in some of the specific examples in the following sections.

The modeling methods based on physics-informed neural networks are briefly reviewed in the following. A given phenomenon in many cases is formulated by a partial differential equation (PDE); for events in low-temperature plasma, a diffusion equation with chemical rate reactions, the Helmholtz equation representing electromagnetic-wave propagations, and almost all other types of formulation are included in this phenomenological category in terms of mathematics. We need a mathematical or numerical technique in a certain level to solve a PDE \( F(u(x)) = 0 \), for instance, for field variables \( u(x) \) and position \( x \), but it is easy to obtain the calculation result on one side of a PDE \( F(u(x)) \) for a given \( x \). Comparing this result with zero in iterative calculations, we manipulate ANN to adjust them, where the input of the ANN is \( x \), getting \( u(x) \) as corresponding outputs. Here, the loss function is set to be the summation of differences between transient field variables in calculation iteration and the experimentally-observed constant values. The corresponding field profile, like the spatial one for electric fields in scattering microwaves, and the key parameter inside it, like neutral gas density that results in elastic collision frequency for the imaginary part of dielectric constant, are calculated in an ANN and are estimated using the loss function. Finally, obtaining a sufficiently small value of the loss function after iterations of ANN calculation and loss-value estimation, the field profiles, and the key parameter are deduced as an approximate value set in this inverse problem. The recent relevant study will be reviewed in the later section.

2.2. Data availability in plasma science and technology

As described in Sect. 2.1, even if we make use of a very simple ANN, a large number of datasets, one of which is composed of input and output values (for regression analysis) or labelled class types (for discriminant analysis), are required. This requirement is far from a conventional sense of mind in which three or four data points appear to be sufficient to identify a curve (normally a straight line) for linear regression analysis. While complexity is always present in plasma, we need a huge number of datasets for the prediction of future events, including possible disorder detections. As we noted above, in the case of Fig. 2(a), we set two input parameters and the required number of datasets (for two-input and one-output mapping) is...
13 at least; roughly speaking, we need the number of datasets by one order of magnitude larger than the input parameters or variables for a simple ANN, and more datasets are necessary for increasing nodes and layers (i.e. two or more, which is called the cases of “deep learning”) in the hidden part when the dependences of input parameters are complicated with strong nonlinear properties.

To fulfill these demands from both mathematical and practical points of view, an increasing number of datasets in low-temperature plasma experiments are inevitable, with revolutionary changes in the data collection step. Specifically, if one builds up scaling laws of one phenomenon in a given plasma reactor, the required number of datasets will depend on the count of input parameters and complexity in a hypothetical mapping. In another case, if a universal scaling law is expected for a given phenomenon over typical experimental reactors with a variety of their shapes and/or sizes, configuration parameters, such as the diameter of a disk electrode and the distance between electrodes, will form a part of input parameters, affecting the further increase of the required datasets. In the following, we briefly review two examples of collecting datasets as a bigdata in and around the scientific category of low-temperature plasma.

First, typical equipment in a fusion plasma reactor facilitates machine-learning analysis, although it has not been intended for it. In particular, a large fusion-oriented reactor provides a limited number of discharge shots, and researchers have been forced to collect data simultaneously as much as they can. Thus, it is quite common to equip a lot of diagnostic systems for a single fusion-oriented reactor,35–38) and they cover a large setup for diagnostics with data storage about electron density and temperature spatiotemporal profiles by infrared-light interferometer and Thomson scattering, ion-energy analyzer by a charge-exchange neutral-beam analyzer, various X-ray detectors, emission spectrometer for neutral particles, and so forth. Consequently and unintentionally, a huge number of datasets are abundant, so many studies related to machine learning methods have been reported so far.36–38) In the fusion-plasma society, we can also find several efforts for universal scaling of parameters, e.g. associated with particle/energy confinement. For instance, a simple regression analysis performed for scaling reactor size for confinement had a great impact on many researches that followed one report;30) such kinds of accomplishments conceptually link to a database which is of great importance for each foundation of machine-learning approaches in the plasma-fusion society.

Second, if a single research group feels difficulties in building up a complete database, collaborative data collection among several groups is possible by replicating a targeted process reactor. In the society of low-temperature plasma, the design and the data sharing on the Gaseous Electronics Conference RF Reference Cell (GEC cell, hereafter)40,41) is a well-known and popular project which is quite successful. Although a fusion reactor includes many complicated phenomena inside its plasma, a low-temperature plasma reactor also includes various factors that make its output vary, without sufficient reproducibility and reliability. From this point of view, to unveil its mysterious properties, several groups built up a similar vacuum chamber with the same configurations and equaled the external parameters with other members, they compared the results for the detectable quantities in common, and simultaneously shared the information which is detectable in one group but cannot be measured in other groups. This kind of collaborative researches is suitable for the foundation of a database, which is of significance for machine-learning approaches.

After discussing the above examples, one may notice that data openness or its availability is a key issue, schematically summarized in Fig. 3. Open-data approaches, which enhance and confirm accessibility to various data in public, have been pursued as worldwide activities, one of which is guided by Creative Commons,42) an American nonprofit organization, that takes the balance of openness in contrast to personal copyrights. For instance, in a case similar to the GEC cell, experimental studies published in scientific journals may contribute to forming an open dataset found in the machine-

![Fig. 3. A map of various data categories with axes of data openness and mechanism clarification.](image-url)
learning database. However, in other cases, the creation of an open dataset will suffer from many difficulties, partly because (unintentionally) missing information exists like details of wall conditions affected by other previous setting parameters, and partly because of a kind of business principle that inclines to keep them in secret. The level of openness is closely related to a white- and black-box status, in which we can or cannot understand underlying mechanisms, respectively; if we open some data to the public, many volunteers may share it with the data supplier to overcome the difficulties, similar to main developments of open software tools in computer science.43)

From another point of view, a machine-learning analysis may leave this black box veiled since we feel heavy difficulties in comprehension of parameters inside ANNs,44) while users can get some hypothetical input/output mapping after one procedure of machine-learning analysis. Temporarily this situation would be acceptable for assuring convinced fabrication operation, and chances to clarify underlying mechanisms in plasma are still in our hands, as possible keys to open this black box. To bridge this gap of partially understandable elements and mysterious residues which might be massive, we should remind ourselves of the importance of the fact that there have been huge efforts of plasma diagnostics17,18) and modelling19,20) in the long history of low-temperature plasma science and engineering. Related to this issue, we show an alternative method for a more instant grasp of underlying mechanisms at every level, which is based on complex network science.26) Plus, with recent research progresses, we revisit this topic in the final section of this review. Before that, we have to piece out our knowledge from another point that we cannot miss in this review: causality in networked elements in plasma.

2.3. Causality in plasma datasets
More or less, causality or cause-effect connection, which governs a system that contains interrelations among physical/chemical factors or elements, exists in plasma. Yet, typically it is frequently simplified or even ignored since it is very difficult to clarify it completely, except for some attempts in numerical modeling with a huge formulation of particle and momentum balance equations of test plasma particles or in a fluid model, although such numerical modeling does not directly reveal causal effects indirectly reproduce phenomena in mathematical equations. In contrast, when we perform machine-learning analysis, whether it is classification or regression analysis, we have to set input(s) and output(s), which are equivalently and analogically cause(s) and effect(s). Here we review the theory of causality applicable to machine learning and obtain valuable knowledge before performing it for analysis of low-temperature plasma.

Causal analysis was a scientific category for legal and moral reasoning,45–49) and has contributed to statistical science for these two or three decades as a logical (and algorithmic) tool that handles association in data distributions, and is being explored as a scientific way of thinking that creates human-level machine intelligence (for instance, see discussion of counterfactual machine learning49)). It is worthwhile to watch out for this way of thinking; in this article, we aim at clarifying a complex network in low-temperature plasma as much as we can, but our attempt might not be perfect, in some cases failing in completion of ongoing element interactions. Let us list a few examples of these, picking up a partial part in Figs. 4 and 5; because plasma chemistry is too complicated to be handled rigorously, we might have to shorten its detail analysis in an optimization procedure at a given factory reactor, leading to unsuccessful
parameter tuning. Or, there may be some unknown but seriously influential reaction that will be an obstacle to completing a total chemical system, and such material processing will not be realized in industry. In cases similar to the aforementioned examples, machine-learning techniques will help the situations recover, and cause-effect reasoning is an essential issue that we should not ignore.

Suppose that one has an experimental data plot in which the deposition rate of amorphous Si is proportional to the intensity of SiH* at 414 nm observed by optical emission spectroscopy (OES) during the deposition process. This is not a cause-effect relationship in causal reasoning but is an associative relation simply deduced from statistical data because the main cause should be the energy level of the input power supply \( C \) which increases both the deposition rate \( B \) and the intensity of SiH* \( A \) through the generation of electrons; both the emission of SiH* and the Si thin-film deposition are caused by SiH4 decomposition, and the main precursor of this deposition is not SiH but SiH3 radicals \(^{50}\) [Fig. 4(b)]. In a practical ANN application, one might choose SiH* as an input parameter when we set the deposition rate at an output node, but taking a good look at the complex chemistry, SiH* would be better to be recognized as another output, or at least, as nodes or edge weight(s) in the hidden layers. In the terminology of causal reasoning, the role of the power supply is located in a confounding position, and the input power from the electric supply directly affects both the SiH* intensity and the deposition rate, with a very weak cause-effect linkage between SiH* and the deposition phenomenon. To pick up a confounding element in a cause-effect relationship, one can reveal a solution, finding a back door in a causal network composed of many elemental factors, as solved through a simple algorithm using graphical representations.\(^{45}\)

In another case, suppose that one has a good experimental scaling law in which the energy level of an external electric power supply (say \( A \)) monotonously increases the deposition rate of amorphous Si (\( B \)) when we use SiH4 as a discharge gas. A typical and specific question that can be treated using causality analysis would be: What will happen if we mix a small amount of H2 in a discharge gas? In this case, insertion of H2 is estimated as a role of interventions \( D \), which results in complicated outcomes in this complex SiH4 plasma chemistry. Without such knowledge of the SiH4 plasma chemistry with additive H2 or in a case where sufficient information from the literature is not available, conditional distributions \( P(B \mid A, D) \) for cases with and without H2 plus the distribution of \( P(D) \) will verify intervention effects. Thus, we can draw a suitable causal inference with links among \( A, B, \) and \( D \).

The aforementioned examples are so simple that we can recognize their meanings smoothly, but in a practical case, a cause-effect relation is typically too complicated to identify and to infer intervention and confounding elements instantly. This background knowledge of possible difficulties in clarification arising from plasma complexities implies that an overly simplistic view of the machine-learning approach to plasma science may not work well to understand and comprehend plasma processing towards optimization of control methods. Most of these difficulties currently remain unsolved, and many of you may hesitate to make attempts to solve them. However, we can stress that there may be suggestions and signs proposed and discovered so far. Before taking into account a specific machine-learning approach to phenomena in plasma, we visualize this complexity using general complex networks which is too large to understand but are nearly complete and rigorous, leading to assistance for intuitive grasp and its further understanding.

Figure 5 handles a rough but total view of active and underlying mechanisms that determine internal plasma parameters and process outputs. This is one typical model, in a variety of pathway exchanges. External parameters like gas pressure and electric power input result in gas-phase parameters like collision frequency and confinement times and determine fundamental low-temperature plasma parameters such as electron density and temperature. These basic parameters play the roles of inputs in the system, and observable phenomena and industrial-technological demands are set as system outputs.
parameters are the origins of three large networking elements in complex chemical reactions, electron excited-state transitions, and physical processes like ion bombardments in a sheath region. Finally, process outputs include dry etching and thin-film deposition phenomena. As we fix a reactor configuration, there are only several external parameters, although detached gases from surrounding walls can be neither completely controlled nor precisely detected; this number of external variables is feasible to set input parameters in an ANN model. Next, as the knowledge of a required dataset number discussed above, we face the challenge of assuming the level of complexities in this very complicated nonlinear system. A layered network structure is found near the inputs with recurrent edges, but following these layers, complex networks are located before outputs. To solve this problem, the total data acquisition is needed in principle, with the reduction of the cost as much as possible, and additionally, we have to pay attention to cause-effect relations; what we need is a dataset that sufficiently includes independent data. For instance, simple data acquisition in OES spectra might be collections of data points for dependent elements in Fig. 4; one may use OES data points for inputs of ANNs, but they should be on the output side as shown in Fig. 5 according to this causality. If one made use of OES spectra as inputs, they would rather play a role of parameters in spatial profiling in plasma, representing a corresponding reactor configuration. Instead, signals detected from plasma, like OES spectra, are likely to be suitable for analyses of unsupervised learning, in which labeling of exiting data points are automatically possible, not for input data in supervised learning like ANNs (Fig. 2).

As schematically shown in Fig. 5, internal underlying processes in plasma are so complicated that, if no other method than machine learning is available, we cannot confirm the way how we select suitable methods of machine learning. To overcome this difficulty, we propose a method to visualize and analyze the complexity in the next section. We show examples of a chemical reaction system in low-temperature plasma, but a similar visualization is possible for other complexities such as a transition system among electron excited states stimulated by electrons in plasma.  

3. Visualization and comprehension of plasma complexities

General networks, which are composed of much more nodes/edges than ANNs, with different styles in topology and consequently more degrees of freedom, are useful for surveying and analyzing general complicated phenomena. When the number of nodes is so large that networks include some statistical properties, they are widely studied in complex network science, which is an interdisciplinary category that extends over topics in mathematics, statistics, social science, biology, and physics. After Sakai et al. proposed that concepts and methods developed in complex network science are applicable to plasma chemistry, several reports demonstrated the potential of this graphical modeling for visualizing and analyzing complex chemistry in low-temperature plasma, and here we take a closer look at these plasma complexities in this graphical view of topology. In addition, we review what are new aspects that the concept of complex networks can introduce to plasma science.

3.1. Graphical representations of plasma chemistry

Both complex networks and ANNs have common components of graphs, i.e. nodes and edges linking two nodes, which are either directed or non-directed. First, we compare these two representations to see if we find common and different points, which are expected to promote our awareness and comprehension. An example we review here is a graphical representation of chemical reactions in plasma, where we define edges as information flow in an irreversible reaction and they start from reactants to products in one chemical reaction equation. We note that reversible reactions are also represented in this rule by using two directed edges. This is a simple and sufficient rule for plasma chemistry, suitable for seeking cause-effect connections and for assuming a hypothetical model for plasma chemistry. In

![Complex-network outlook](image-url)
previous reports, this style or its derivative method works to deduce reactant-product (cause-effect) relations, to rescale complex plasma chemistry, and to reveal reaction-in-network analysis based on connectivity matrix.54)

Figure 6 illustrates two drawings of the same chemical reaction network. One is in a style free from arrangements, similar to that in complex networks, whose network topology is of significance, nodes with a large number of edges being centered while those with a few edges are in the periphery. Another is a layered network, similar to ANNs that are feed-forward networks, in which input nodes are on the left side and outputs are located on the right side. Assuming that the purpose of this process is the formation of carbon nanoparticles and carbon nanotubes, we set CH₄ and catalysis species (M) with CH (generated by electron impacts on CH₄) as input species and select molecules and arrange molecules or radical species with higher-order C atoms as output species. From the latter network in a layered style, we notice that there exist many recurrent edges that direct the same or upstream layers. This fact implies the following two points: (1) recurrent elements may yield drastic nonlinear and/or memory (hysteresis) effects, and (2) a model based on a simple ANN, with monotonous information flow from the left to the right side, is insufficient to represent a process input–output function in CH₄ chemistry. Namely, we are now uncovering this complex plasma chemical system from a mixed point of view developed in complex network science and machine learning.

From the above item (1), roles of recurrent edges in general complex networks are being explored in ongoing research, while, from a classical point of view, we can learn the effects of feedback loops, represented by the knowledge in analog control theory.55) When a recurrent edge exists in parallel to a transfer element of information flow, this edge may induce an unstable response in time evolutions. Second, such a recurrent edge will form a recurrent neural network or a long short-time memory, which is usually adapted to time-series datasets and involves (time) evolutions; this is not the case for the reconstruction of the stationary state. At least, we consider such recurrent effects in solutions.

From item (2), shall we abandon possibilities for the application of simple ANNs to plasma processing? This is not always the case. As we apply network rescaling to smaller ones with a reduced number of species, which will be addressed later, such recurrent edges may be neglected in some cases, while we should pay attention to causality in primary reactions. This possibility will be associated with the robustness of chemical networks, which frequently exist from the previous studies and are also addressed later. In short, we should carefully examine the framework of given chemical reactions, and by neglecting recurrent edges with reasonable treatments or by grouping nodes in which recurrent interchanged edges are included to make a cluster or a community, the resultant framework may be evaluated to be in an ANN configuration.

3.2. Microscopic view of complex network in plasma chemistry: identification of species roles

Before discussing the possible simplification of a complex reaction network, we take a good look at complex chemical networks microscopically, namely, to identify a role of an individual species. Microscopic examination will lead to a complete understanding of cause-effect relations in a chemical reaction, which can be replaced by an agent-product (or reactant-product) relationship. The standing point of this insight is similar to several previous reports, in which methods of pathway analysis or other graphical representations have been already demonstrated.52–64)

In the category of network science, there have been several centrality indices defined and applied to many analyses.52) Since most networks have no clear upstream or downstream directionality in their network topology, these measures focus on classifications of central or peripheral locations. In general cases in chemistry, some of which have been investigated in complex network science so far,53,56) reactions are frequently bidirectional, such as protein reactions in biological fields. In plasma chemistry, however, many reactions are directional and reaction flow from original sources to the final target species is apparent. Thus, based on the centrality indices proposed in network science, we make use of some of them, or newly modify others in their definition and indication.

Reactant-product (or reactant-subproduct) connections in a reaction system are basic in material processing in plasma chemistry, which clarifies the causal explanation. However, as recurrent edges often point to original sources in molecular plasma reactions, the reactant-product connection seems to be ambiguous. Let us apply the rule for generating a reaction network as follows: a directional edge starts at a reactant (i.e. on the left-hand side of a chemical reaction equation) and ends at a product (on its right-hand side). In most cases in plasma chemistry, chemical reactions are irreversible, and directional edges are suitable for configuring networks. After completion of the entire network, for instance, as shown in Fig. 6, the reaction system of plasma chemistry in CH₄ gas has a recurrent edge, where CH₄ is the very origin of gas species supplied from the gas inlet of carbon-nanoparticle reactors. Thus, its density in the gas phase does not always a pure function of the single variable represented by the inlet CH₄ flow. In their first report, Sakai et al. made use of the PageRank index that was proposed by researchers in a well-known website platformer to measure importance with accumulated effects of directed edges in a network, and this centrality index works to point out the level of products in complex chemical reactions.51) In the following, we list the centrality indices which have been applied to plasma chemistry so far.

3.2.1. Degree (indegree and outdegree).52,53) The number of edges connected to a given node, degree, is a fundamental numeral representing assignments and contributions in an entire network. The more edges the node is linked to, the more influential it is for other nodes and to the network topology. Since edges (or reactions) are directional in many cases of plasma chemistry, we can count incoming and outgoing edges; these counts are indegree and outdegree, respectively. If the in-degree is superior to the outdegree, such a species more frequently plays a role of a product in a reaction system.

3.2.2. PageRank in normal and reverse directions of edges.51,55,56) This index indicates substantial importance in a network with directional edges. The number of degrees is just a numeral of summation, but PageRank is a kind of eigenvector centrality index, reflecting a role throughout the network. That is, even if the indegree is 1, the PageRank
value of the corresponding node will be high when the edge that pointed to the node originates from another node that has a number of important incoming edges. Thus, in plasma chemistry, a PageRank index represents the level of importance as a product. Then, applying this to the network with all edges with the reverse direction to the initial and normal one, we will obtain importance as a reactant by PageRank index; derivation of PageRank in the normal direction according to irreversible reactions and in the hypothetically reversed direction allows us to obtain levels as products and reactants, respectively. If both PageRank indices in both directions are high, we can conclude that such a node is important as a product in some cases and also as a reactant in other cases; this is, indeed, important as an intermediate.

3.2.3. Betweenness. \[52,53\] This index is for evaluating the quality of intermediate species; betweenness centrality is defined as a number of passes along shortcuts between every pair of nodes throughout the network. Even if both values of PageRank are balanced, the species is not central in the network topology when it is located along one of the pathways between reactants and products. Important intermediates are likely to be located along a limited number of such pathways, and almost all information flows have to pass them.

3.2.4. Closeness. \[52,53\] While the betweenness centrality index provides essentials of intermediates in the network topology, the closeness index is defined as an inverse of summation among all path lengths to all nodes. That is, this quantity means a level of a topological central location, as a node located visibly in the center. This quantity also reflects upstream positions of information flows when the edges are directional; if the positions of two given nodes are symmetric for the exact center point, this index shows higher values when the species are at nearby locations of reactants in plasma chemistry.

We note that these indices are quantitative, not limited to discriminant or classification analysis; levels of reactant, product or intermediate, or those in the core or around the outer region, can be evaluated by their quantities. After exemplifying their values and comparing them with the previous qualitative facts assured in the literature, we will confirm their validities in each case of reactions. For instance, Sakai et al. investigated the case of CH₄ chemistry using PageRank, and summarize the effects of induced reactions by electron collisions.\[51\] In a network without electrons, C₂H₂ gains high values of PageRank in the normal directions of edges, while H and H₂ increase their PageRank indices drastically with electrons, indicating that electrons enhance decompositions of molecules or molecular radicals. CH₃ is evaluated as an important intermediate species with balanced two PageRank scores. In another case, in SiH₄ plasma chemistry, \[52,55,56\] which is a typical chemical system in semiconductor fabrications, role evolutions of typical species are analyzed as reactions proceed to successive reactions. Mizui et al. revealed that \[56\] in an initial state of reactions starting from electrons and SiH₄ as a mother gas, SiH₂ is classified as a product since its PageRank score in the normal direction is high, whereas that in the reverse direction becomes large and its role as a reactant in the complex chemical system is fixed afterward. These successive reactions, which are likely to observe in many cases when we use simple molecules as a source gas species, will be discussed again in the next section.

4. We can further stress that this microscopic network analysis is useful for substantial clarification of cause-effect connections in complicated networking. Even though a few recurrent edges are found in a given network as shown in Fig. 6, we can obtain substantial measures for causal analysis (see Fig. 4) as described below. Reactants in a complex chemical system are influential as elements with causes, which are labelled as species with high-level PageRank in the reverse direction or closeness centrality indices. On the contrary, products are estimated as elements affected by others, which are found as species with high-level PageRank in the normal direction of low-level closeness centrality indices.

Another important point that these defined quantities at nodes make is to examine their histograms or spectra of belonging species at each quantitative level and to use them in a two-dimensional diagram with axes of two different indices. The spectra in the case of degree are very well-known as degree distributions, and diagrams were recently investigated by Murakami et al. \[53\] and Sakai et al. \[52\]. These statistical profiles build up macroscopic properties of complex networks in plasma chemical reactions, as described in the next section.

3.3. Macroscopic view of complex network in plasma chemistry: properties of scale-free networks and resultant dimensionality reduction

When we examine distributions of nodes as a function of one centrality index, we might have the illusion that the type of a distribution strongly depends on a specific chemical system. However, surprisingly in many sample systems of plasma chemistry, \[53,56\] it is not the case as well as many other systems investigated so far in complex network science. \[65,66,68,69\] They frequently belong to a universal principle in which, for instance, the degree distribution is in a scale-free manner. \[24–26\] This fact itself is really significant, and additionally as discussed below, it leads to the context of dimensionality reduction. \[53,70–74\] which is one of the unsupervised learning methods [Fig. 2(c)]. This framework may lessen the plasma complexity examined in this study and simplify the system which is a target in various analyses of low-temperature plasma science. This informative and useful conclusion will be carefully deduced in this section.

For instance, in the SiH₄ plasma chemistry, tens of species are in the reaction set, and hundreds of chemical reactions are listed in the literature. \[50\] A large number of species are involved in a few reactions, and a few species are found in hundreds of reactions. Thus, the degree ranges from a few to hundreds, and the number of species decreases as the degree increases. In the report by Mizui et al. \[56\] the degree distribution or existing probability of nodes as a function of degree \( k \) is proportional to be \( \exp(-\gamma/k) \) with \( \gamma = 1.05 - 1.48 \). This is a power-law degree distribution, and the graph with this property is a scale-free network. \[24–26\]

Scale-free networks is generally robust against an external intentional (or unintentional) attack; they have a capability of error tolerance where one node at most degree levels is removed. This principle indicates that the power-law distribution is maintained only with slight changes of degrees in the connected nodes, except in the cases where very essential
nodes near the highest degree are removed and large damage takes place.

Before we discuss the merits of plasma chemistry with the property of a scale-free network, we speculate the reason why reaction networks in plasma chemistry are likely to be scale-free. The Barabási–Albert model is well known to reproduce power-law degree distribution in general complex networks with a huge number of nodes, and it is partially applicable to plasma chemistry. This model assumes two ingredients in network formation: growth, which indicates that a new node joins the network one by one, and preferential attachment, which means that the larger the degree the node has, the more likely it is to be linked to the new one it is. In plasma chemistry, when pure molecules (say, SiH₄, in Fig. 7) and high-energy electrons coalesce as reactants, atoms (H) and radicals (SiH, SiH₂, etc.) are generated as products via decomposition, and these products are newly linked to the products by edges. As a next step, all of these species play the roles of reactants, and other species like SiH and Si₂H₃ create new nodes, which are newly attached to this growing network. Interestingly, these radicals and atoms generated at early steps are more fundamental and active species in reactions, consequently more likely to react with other species, which is a preferential property.

As we noted above, this kind of scale-free network is robust against accidental removals of nodes, which is a favorable macroscopic property and also leads to possible simplification of the complexities. For instance, protein reaction networks, which are scale-free networks, always suffer from mutations and protein misfolding, but they are sound in their functions. Due to various climate changes and current human invasions of nature, species are deleted from food webs, which are scale-free networks, but this web has managed to be active, although this is still an issue of major concern for ecology and our environments. The Internet and the World Wide Web also possess scale-free networks, suffering from random node removals, by attackers or hackers; however, by random node removals up to 60%, their connectivity is still at a good level estimated by the distance between an arbitrary node pair. As far as the extension to unsupervised learning is concerned, this drastic reduction of nodes suggests that keeping its macroscopic property, the robustness in a complex network assures a reduced-number model for simplicity (see Fig. 8), which is associated with dimensionality reduction in large-data handling, which is described in the following.

As an example case in plasma chemistry, a reaction system in atmospheric-pressure plasma is visualized using complex networks, and its rescaling method into a smaller size is addressed with verification of its validity using numerical results. Following a series of studies of numerical analysis that deals with more than one thousand reactions in the zero-dimensional model, Murakami et al. rescale a complex network by considering a scale-free property mixed with other statistical perspectives, whose schematic outline is shown in Fig. 8. In atmospheric-pressure He plasma, although He is the inlet gas, O₂, N₂, H₂O, and other minority gases in the ambient air are involved in a reaction system. This reaction set consists of 65 species and 1360 reactions in total, which yields a huge data process when we numerically analyze the corresponding rate equations in the computational calculation. O₂ has the highest degree of 2546, which means that the same number of edges are linked to O₂, and there are a number of multi-links that bridge O₂ and the other partner species simultaneously. From another point of view, when we make an ANN model of this chemical system, this number of elements with interchange reactions is too complicated to obtain a suitable layered feed-forward network.

The method proposed for data reduction is described as follows. By deriving the degree distribution (i.e., a diagram of node number and degree), we find that the distribution is proportional to be \( \exp(-\gamma/k) \) with \( \gamma = 1.1 \), indicating a scale-free property. We can simply perform random removals of nodes, as exemplified in the case of the Internet, but here three constraints are applied for the node reduction rule. That is, three different macroscopic (statistical) data plots are taken into account: a diagram of outdegree and indegree, that of betweenness centrality and closeness centrality, and that of clustering coefficient and degree, where clustering coefficient is the interconnection rate between two nodes linked to a given node and indicate a kind of network binding force. Considering all four diagrams, a selection of removed nodes is performed to maintain data scattering trends in all diagrams. Although other selection possibilities might exist, the remaining species are: electron, He (ground state), O₂ (ground state), O (ground state), He(2P), O(3P), O(1D), O₂(a’D), He⁺, He²⁺, O²⁺, and O²⁺. To confirm the validity of this selection, a numerical simulation based on this selected set was performed, and reasonable agreement with other selected sets was observed.

If all 65 species existed independently in a system, we should search for a solution for a dataset in the 65-dimensional space. It is sufficient in the above case to take account of 12 species, which is a significant reduction of dimensionality. Dimensionality reduction is one of the topics that are being explored as a machine-learning method, categorized as unsupervised learning. Other methods like signal sampling in a graph are also currently investigated, while this network rescaling method, arising from visualization and microscopic careful observation, is one way to reduce the network complexities existing in plasma chemistry; the significant point pursued in this method is that selections of nodes are from every level of statistical properties, not only from the top level, and that the robustness against removals is basically validated from its scale-free property. Also, this network-based method shares essential points with research in other chemical societies.

After this simplification, we understand more smoothly what are cause-effect connections among species, and it is more reliable to judge whether ANN models are applicable or not. Furthermore, the visualization as a complex network is a kind of white-box approach and allows us to select representatives reasonably, as we described above and shown in Fig. 1. Moreover, as we will discuss in the next section, after this visualization of complex networks in low-temperature plasma, we can find more choices to perform other machine-learning methods.

The fact of the removability of species in the scale-free-network model leads to another example of algorithm-based methods for supporting experimental approaches. This example in the following, which is practical for
experimentalists, is relevant to nano-particle formation in chemical processes in plasma. The nano-particle formation is one of the active scientific topics in plasma chemistry. In fact, we point out a number of studies reported for Si and carbon nano-particle formation using low-temperature plasma;\(^75^–77\) high-throughput nano-crystal particle formation is possible without any tools for forced-shaping frames. Figure 9 shows the case of polymer growth in silane plasma, where polymers

Fig. 7. (Color online) Growing chemical system from one molecule species (SiH\(_4\)) in plasma.\(^56\) (a) At the very initial phase. (b) Growth and preferential attachment in SiH\(_4\) chemistry with a specific gas residence time. Reproduced with permission from Mizui et al., R. M. Benito et al. (Eds.): COMPLEX NETWORKS 2020, SCI 944, p. 398–409, 2021. Copyright 2021 Springer Nature Switzerland AG.
are expressed as $\text{Si}_x\text{H}_y$ with non-zero $x$ and $y$ integers and they are seeds of growing particles. When we examine locations of species of polymers, in the degree distribution as their size increases, we find changes from high- to low-degree areas in the spectrum. After the lower-degree shifts, such species are in the group of removable-class species. Actually, with nano-particles being removed from the reactive space, this chemical system continues to be stable. This description certainly allows us to confirm the reasoning why nano-particle formation in complex reaction systems in plasma chemistry is so natural that it keeps on creating particles. Although those microscopic/macroscopic network analyses based on complex network science are not in the category of machine learning, we can assess automatically this reasoning, with one information data processing and without any profound knowledge of plasma science.

4. Current progress in machine learning for analyzing plasma

Machine learning is one of the highlighted topics in a wide range of science and technology studies. Its basis like the backpropagation algorithm was established several decades ago, and recent progress in machine-hardware capabilities enhances activities in machine-learning research and development, which makes it popular. In the meanwhile, in plasma science and technology, although ANNs have applied to data analysis around 2000 as we reviewed above, we notice most of the prominent achievements are based on other approaches. The long-range historical efforts have been established in mechanical, electrical, and optical diagnostics developed by sensors and effective tool designs. Purely analytical approaches based on equation formulation have been also active, and computational approaches based on numerical finite-difference methods for multiple sets of differential equations, have promoted scientific and technological basis. Then when we seek effective and reasonable hybridization of these historical heritages and newly-born modern machine learning, what kind of analysis is favorable and what can we expect from this fusion? After we visualize one part of the complexities that plasma possesses in the previous sections, a clearer image or ongoing path can be found for this purpose. One way is simply to increase datasets available for machine learning drastically, which will be discussed in the next section. Here, we show other promising cases, using a few examples of a new application of machine learning for plasma science, despite the shortage of datasets, most of which are well matched with the historical heritages in plasma science and engineering.

4.1. Community detection or clustering: an example of unsupervised learning

Before complex network science emerged around 2000, relatively small-size networks have been investigated in graph theory for several decades. In such a study, topology or geometrical structure of node-edge combination is the main targeted topic, and mixing with mathematics, various application-oriented researches such as traveling salesman problem and vehicle routing problem have been addressed. As the size of an aimed network increases, the structure and function of a part of the network, or in other words, subgraph, cluster, and so forth, have attracted attention. After the targeted network size becomes huge as the machine hardware with high-speed data memory is enhanced, analyses of subgraphs and cluster properties were examined by various network-based techniques; hereafter, we focus on community detection, which is a cluster classification by automatic (or semi-automatic) algorithms. Community detection is, by definition, the division into clusters that is completed by removing the minimal number of edges also inferred as clustering analysis. If we separate a network into two parts by breaking the minimal number of edges, the cutting location across the edges is identical. However, when we search for more than three clusters in one network, the cutting pattern differs in locations, depending on applied rules, and several methods have been investigated in
network science. Applicable data for community detection is not only for networking data but also for any data with numerical indicators in multi-dimension; even if the linkage between data points is unclear, one can select a suitable rule, such as distance defined by Euclidean metric, and such a hypothetical network is a target of community detection. In other words, community detection shares the common concept and methods with clustering analysis, which is one of the general unsupervised learning methods. Here we show a community detection method applied to classify species in plasma chemistry complexity, but similar techniques can be expanded to non-networking data; an easy-to-detect and somewhat trivial dataset like optical emission spectra, mass spectra, signals detected by Fourier-transformed infrared rays, and surface morphology of manufactured micro-devices are similarly applied as problems of community detection or clustering analysis, where these data are very easy to be accumulated but seem to be meaningless in some cases. Even when causality is unclear, this unsupervised learning may be valuable, in the classification of regular/irregular operation of a plasma reactor in fabrication processes.

Again based on a complex SiH4 chemical system, we perform community detection to classify belonging groups in which species are for the same purpose. Figure 10 shows the case in which edge betweenness is used to decompose this intertwined network. Edge betweenness is defined for edges, similarly to that for nodes described above; if this centrality measure for a given edge is high, the corresponding edge is located along a large number of short paths between arbitrary node pairs. According to the divisive algorithm based on the edge-betweenness, the edge with the highest centrality is removed first, and after recalculating the centrality values for the network without the first removed edge, the same iteration of edge removal is repeated step by step. Finally, all the nodes are isolated after all-edge removals. This node-isolation process is visible using a dendrogram or hierarchical tree, as shown in Fig. 10. In the following, we discuss classification using the hatched clusters in Fig. 10, but other cluster separations are possible with more rough or more fine subgraphs using such dendrograms.

From this classification, we find two large groups. One of them is composed of electrons, SiH4, SiH3, H2, H, several kinds of ions, and molecules like SiH8 without dangling bonds. Almost all the species listed here contribute to film deposition like amorphous or hydrogenated nano-crystalline Si thin films. The other large group consists of SiH and higher-order radicals like SiH2x+2−z, where x and z are non-zero integers. These species work for polymerization and gas-phase nanoparticle formation, and the variation of the degree levels has been already shown in Fig. 9. Using the centrality index, i.e., edge betweenness, which never includes direct information of chemical properties in the computation, community detection based on network topology automatically and successfully separates one group from the other, with some other residue species. These two facts are deduced from Figs. 9 and 10 clarify that ICT-based methods notify underlying mechanisms that cannot be pointed out by other conventional methods.

Unsupervised learning applicable to plasma science and engineering is not limited to such clustering classification. Analysis based on principal components is a well-known procedure, listed as one of the unsupervised learning and historically comprehended in multivariate statistical analysis. A dimensionality reduction technique that has already been described in the previous section is another, and many of them work to classify existing data into different categories, regardless of dataset size. If the volume of datasets is limited, the choice of unsupervised learning will be more practical and reliable than supervised learning methods that generally require a dataset with a huge size.

4.2. Supervised learning working with other theoretical frameworks

As described above, around the research of low-temperature plasmas, there has been plenty of historical knowledge in variety, and the hybridization of such rich information with
machine-learning methods will have substantial significance to promote our smooth understanding. Here we review a good example: solution for Boltzmann equation analysis. Electron swarm parameters, as well as electron energy/velocity distribution functions, form the basis of low-temperature plasma processing, and the Boltzmann equation, governs them. However, the complete set of its suitable parameters cannot be solved analytically, and various numerical methods have been proposed so far. This classical but still difficult problem will become a good benchmark for the hybridization of plasma science and machine learning.

Tezcan et al. apply an ANN to obtain electron energy distribution functions (EEDFs) from electron swarm parameters. Mapping from the swarm parameters to the energy distribution function is a targeted task, and ANNs are trained to realize a substantial mapping scheme; this is not in a discrete form of a mathematical function but through a kind of ANN window, which is an approximator of mathematical function(s). Based on cross-sectional data of gases, a finite difference method in the steady-state Townsend condition is used to solve the Boltzmann equation. In detail, an effective ionization coefficient and the corresponding EEDF are iteratively obtained while exchanging them as input and output values until they settle down. After fixing them, other swarm parameters are derived from EEDF. That is, in this work, the Boltzmann equation is completely solved in this numerical procedure. Next, the ANN is arranged as follows. The input layer, in which four nodes exist, is composed of these four swarm parameters at a given electric field normalized by gas number density. After the hidden layer with 50 nodes, the ANN has the output layer consisting of 100 nodes or 100 data points in the EEDF. The training (82 in quantity) and test (20) datasets, which are digitized, are selected from the fixed solution which consists of continuous curved plots in the Boltzmann equation analysis. As a result, the trained ANN provides a smooth fitting, which means that the ANN model is a good generalized approximator even for untrained data points. We note that, although the training dataset for this ANN configuration, which has 50 nodes in the hidden layer, is deficient in size or quantity to rigorous mathematical

Fig. 10. (Color online) Community detection in SiH₄ chemical system and group identification using edge betweenness index. On the upper side: dendrogram based on closeness in network topology. On the lower side: classification of communities in color. Two large groups are for nanoparticle formation and for film deposition.
requirements, it is still sufficient for the physical requirements to reproduce an EEDF curve as a function of electron energy, which is smooth without any fine ripples. This is an example of typical supervised learning cooperation with historical agreements in the Boltzmann equation analysis.

Kawaguchi et al. also solved the Boltzmann equation to obtain electron velocity distribution function (EVDF) $f$ using an ANN, whose hidden layers are deeper, partly using the concept of physics-informed neural networks as shown in Fig. 1. All the terms in the Boltzmann equation are set on the left-hand side of the equation, which is minimized by the output of the ANN, so that the elements of this differential equation form the main part of the cost function $J$ in the computation procedure, with other parts composed of the boundary conditions. The inputs of the ANN are two components of electron velocity $v_x$ and $v_y$ along the two axes, and the output is one representing variable of EVDF, while there are three hidden layers one of which has 450 nodes. Then, on behalf of the training data in input-output pairs, this ANN is trained so as to minimize the cost function $J$ down to zero (or less than the neglectable error $\sigma$) at the points in the two-dimensional velocity space where EVDF should be derived; the number of points is 4000 inside the parameter area $S$ with 400 on the area boundaries $\partial S$. In the meanwhile, the test datasets that are conventionally used to estimate modeling validity is replaced by the completely independent datasets, which are the results in the Monte Carlo simulation.

Here we show one calculation result by this method in Figs. 11(b) and 11(c). EVDFs are generally anisotropic, where $v_x$ is in the vertical direction of externally-applied DC fields and $v_y$ is parallel to them, and such anisotropy is clearly observed in the calculated EVDFs. Using the corresponding EEDF as a subproduct of the calculation, we can confirm the validity of EVDFs since the calculated EEDF is consistent with that solved by Monte Carlo simulations. That is, while Tezcan et al. used training datasets that are calculated in the Boltzmann equation analysis, Kawaguchi et al. applied the Boltzmann equation itself within the concept of physics-informed neural networks, by calculating it hypothetically by means of ANN functions as universal approximators. In both cases, the historical heritage of comprehending and handling the Boltzmann equation is well mixed with the rapidly growing machine-learning technology. Hence, these studies are somewhat different from the other black-box tools based on ANNs; they revisit our previously established knowledge and overcome difficulties that have not been solved well so far and remain a concern.

We note that physics-informed neural networks are classified into unsupervised learning in some cases, although they still belong to supervised learning when we define them as a method with universal approximators. That is, the boundary between supervised and unsupervised learning is unclear in some cases. However, one key factor that is crucial for both methods is sufficient dataset accumulation, whose volume definitely improves the accuracy of predictions and outputs. We continue to discuss applicable and suitably supervised learning for low-temperature plasma in the next section, regarding dataset collections that are available using current and recently-developed experimental and factorial tools.

5. Progress and potentials for bigdata acquisition from plasma

Next, we discuss details and possible frameworks of experimental dataset acquisition for machine learning and other ICT schemes. So as to perform this task for low-temperature plasma processing, in comparison with computational approaches based on numerical machine calculations, we need more time and effort for adjustments of experimental equipment since large-size datasets, which are typically difficult to be acquired manually, are forced to be collected automatically or extracted from existing data. However, as we review in this section, we can find suitable and available methods that are ready for application for our purpose or may become powerful tools for large-dataset acquisition in the near future.

5.1. Methods suggested in Internet-of-Things technology

In the current progress in sensor technology, many devices for monitoring trivial data in our daily life and in infrastructural and industrial scenes are connected to networks of wireless communication, frequently up to the Internet, with local battery or energy-harvesting power supply, leading to self-standing but connected sensor networks. Such individual tiny terminal devices belong to the so-called Internet-of-Things (IoT) technology. However, we did not find many low-temperature plasma sensors in the category of IoT, partly because IoT sensors are designed not in vacuum but in the ambient air at atmospheric pressure. The IoT sensors have been verified to work for accumulating various datasets that are applicable for machine learning due to their low-cost property and their possible installation in many locations. In this point of view, Sakai et al. designed in-vacuum active sensors with a function of wireless data transfer and verified validity in small-chamber experiments.

Figure 12 illustrates the concept of in-vacuum active sensors [Fig. 12(a)] in comparison with conventional plasma diagnostic schemes. So far we set sensors outside a vacuum chamber [Fig. 12(b)], and only a probe component was inserted into it if the insertion is necessary, except for some sensors like pressure gauges that are in vacuum; nondestructive testing of plasma is always of prior importance. In the meanwhile, packaging with sufficient seal levels and environmentally-friendly components has been widely accepted and promoted to use in electronic devices, and Li-ion batteries are now required to keep their capabilities in low-pressure conditions for aircraft. Then, in cases of inside-space vacancy without harsh plasma fluxes, discrete/integrated electronic devices and batteries, which are almost all elements of an IoT sensor, can be installed in vacuum. Sufficient ground level in its circuitry part should be required in all cases, and removals of inappropriate electronic parts have to be performed in advance; in such conditions, also thanks to various successful operations of sensors in spacecraft. IoT sensors are now ready for use in vacuum [Fig. 12(a)]. Clear advantages of these schemes are removal of difficulties in handling wiring across vacuum and air and reducing stray inductance and capacitance along cables, which are always the concerns shared by chamber designers in industry and experimentalists in laboratories. Using this sensor, signal-to-noise ratios are enhanced partly because noise invasion through cables is drastically reduced and partly because
detected signals are instantly converted into digital numerals in bits. Simultaneously, by setting multiple sensors, we can easily increase sensing locations in vacant spaces in a vacuum chamber.

After careful in-advance testing of all electronic parts, Sakai et al. observed successful installation and the functions of an in-vacuum active color sensor.\(^{23}\) Constant output signals are transferred from the sensor inside a glass chamber to a data-storage personal computer outside it while the gas pressure around the sensor changes from the atmospheric pressure to 30 Pa. After plasma ignition, it successfully detects plasma emissions, and the signal trends are the same as those detected through a quartz vacuum window outside the chamber. Recently, we experimentally confirmed that this sensor is still active in Ar plasma at 100 Pa, as shown in Fig. 12(c), which means that the electronic circuitry can be inserted into plasma if sufficient metallic housing encloses it. Of course, as a diagnostic tool, such disturbance against plasma is not accepted, and its location should be outside the discharge region, near the wall, for instance.

![Diagram](image.png)

**Fig. 11.** (Color online) Method of solution for deriving electron velocity and energy distributions by a deep neural network, following Ref. 85. (a) Schematic view of solution scheme. (b) Electron velocity distribution function in Ar at \(E/N = 500\) Td, where \(E\) is the electric field with gas number density \(N\). (c) Electron energy distributions in Ar in case of (b). Numerical results calculated by Monte Carlo simulation are also plotted.
This sensor is composed of a sensing-head chip with an A/D converter (S11059-02DT, Hamamatsu Photonics K.K.), a microcontroller (AVR Microcontroller MEGA88V, Atmel Corp.), and a Bluetooth module (Bluetooth module RN42, Microchip Technology), and inclusion of the microcontroller is quite significant since its software module governing the input–output control of digital sensor data can be freely rewritable. Then, this sensor enables to perform not only digital-data wireless transfer to the data storage device outside the chamber but also small-size data analysis inside the chamber. This is exactly equivalent to an edge-computing device; both data cleansing and immediate signal feedback to plasma can be achieved using similar circuitry.

5.2. Other methods feasible currently or in the near future

Due to its flexibility and universality for performing appropriate predictions, the topics we discussed above, i.e. complex science, machine learning, and IoT, are being explored to be valid in various scientific and technological fields. In this topical review, we review various aspects of low-temperature plasma science and technology and methods arising from information science and technology. This hybridization to create interdisciplinary science and technology has been pursued so far, with a number of successful accomplishments described in the previous sections. Compared with other science categories, however, this scientific hybridization seems to be somewhat slow, partly because datasets for analysis are insufficient. On the other hand, fusion-oriented plasma research has been well mixed with machine-learning technology; for low-temperature plasma, similar and/or alternative efforts can be possible with current situations and may be feasible in the near future. In the final section of this topical review, we search for four possibilities that will grow up soon.

First, data collection and/or data reinforcement by numerical simulation is promising to supplement datasets for analysis in complex networks and machine learning. For instance, when experimental data collection is difficult with deficiency of diagnostic equipment, first, a small number of experimental data are compared with some results of the numerical simulations to confirm their consistency, and second, datasets are deduced from numerical computation in calculation computers with many input parameter sets; this is a synthesis of physical and virtual experiments for dataset collection. In other words, a given phenomenon is well understood down to the basic and fundamental levels but its mathematical formulation is not simple since the linkages between equations are too complicated, one can set equations carefully by differentiation and integration, with the addition of some statistical parameters like random numbers, resulting in good-quality training data for supervised learning.

Actually, Kruger et al. collected datasets from a Monte Carlo simulation code for plasma sputtering phenomena, and perform supervised learning based on an ANN. In detail, one set of inputs consists of incident ion energy distributions (from 0 to 1800 eV in every 8 eV), and the corresponding set of outputs are energetic and angular distributions of target material atoms (with energy in 30 levels and directional cosine in 20 levels for three species). To train three hidden layers one of which has 1000 nodes, 439 samples in the datasets are used for training, validation, and testing; this number is selected to suppress overfitting in the ANN training process. In this study, Kruger et al. stressed the importance of causal relationships between inputs and outputs data for the ANN, and we share the importance of causality in the previous section.

Second, another scheme that has been already established for possible dataset collection for analysis is data logging of events in fabrication processes in industrial activities. In a plasma fabrication process in mass production, a huge number of wafers or substrates are handled with a specific recipe. If the testing process is in a step of the post process, parameters of recipe and quantities, classes, or images are evaluated as one corresponding quality level. During this process, one can set equations carefully by differentiation and integration, with the addition of some statistical parameters like random numbers, resulting in good-quality training data for supervised learning.

Actually, Kruger et al. collected datasets from a Monte Carlo simulation code for plasma sputtering phenomena, and perform supervised learning based on an ANN. In detail, one set of inputs consists of incident ion energy distributions (from 0 to 1800 eV in every 8 eV), and the corresponding set of outputs are energetic and angular distributions of target material atoms (with energy in 30 levels and directional cosine in 20 levels for three species). To train three hidden layers one of which has 1000 nodes, 439 samples in the datasets are used for training, validation, and testing; this number is selected to suppress overfitting in the ANN training process. In this study, Kruger et al. stressed the importance of causal relationships between inputs and outputs data for the ANN, and we share the importance of causality in the previous section.

Second, another scheme that has been already established for possible dataset collection for analysis is data logging of events in fabrication processes in industrial activities. In a plasma fabrication process in mass production, a huge number of wafers or substrates are handled with a specific recipe. If the testing process is in a step of the post process, parameters of recipe and quantities, classes, or images are evaluated as one corresponding quality level. During this
testing step, instead of ignoring the detected raw data, they can be transferred into the outputs of the training process in machine learning analyses, whereas the corresponding inputs are in the recipe. In this case, the database will be dependent on a given individual machine used in the process and valid for the specific reactor. Jalali et al. investigated time to failure (TTF), which is for planning scheduled maintenance for dry etching reactors. TTF was analyzed using both supervised and unsupervised learning methods, and the training datasets were recorded in an industrial processing reactor for a six-month period. They used the logging data such as gas flows, discharge voltage signals, the time needed until etching stops, and other applied recipe parameters, which are pure input parameters, together with several resultant signals detected during operation. Both regression and classification analyses are performed using unsupervised learning methods with algorithms of linear regression and support vector machine and supervised learning methods with algorithms of ANNs and random forest. Consequently, for example, the factor of a machine health state is successfully predicted quantitatively, which is directly linked to TTF.

Third, we conceptually propose the use of image datasets, which is possible to get in an experimental laboratory with one plasma vacuum chamber, without any particular facilities. Images have been examined by supervised learning to distinguish specific items displayed in them, and such pattern recognitions are one of the most successful cases in machine learning. A charge-coupled device (CCD) in a mobile phone is now a very popular sensor around us, and in particular, one photo image taken by this device possesses a very large amount of digital information data, like around 1 MB. That is, if we take a photo image of plasma emissions by a conventional CCD, the potential information accumulated in images is tremendous, some of which may include significant elements in underlying processes in plasma. For instance, a spatial profile of emissions definitely contains information on plasma spatial distribution, leading to the classification of homogeneity levels after some data processing, hopefully by machine learning. Information of spatial distribution is not usually pure input parameters to a specific process, and such data will be very close to the cause level in causality. If we put adequate wavelength filters to the inlet of the CCD, detected intensity can be a representative factor for chemical reaction enhancement, with its spatial tendency; although detailed information on chemistry is not fully included, with some other simple input parameters, the mixed dataset will get close to the sufficient set for inputs of machine learning. We again note that similar to OES spectra, plasma emissions are rather in an effective position in causality, and they will work for some classification analysis.

Fourth, and last but not least, the combinatorial plasma process, which is an experimental system with a facility dedicated to this purpose, has a high capability to supply large databases; it was initially proposed for database acquisition for parameter maps, but it can be ready for that for machine learning. Moon et al. and Setsuhara et al. proposed the concept of a combinatorial plasma etching process and verified its validity using experimental results. The initial concept of combinatorial approaches to chemistry was proposed and verified in general cases in which continuous parameter variation in multi-dimensions is possible to obtain experimental samples as corresponding outputs. The machine designed for this purpose in plasma chemistry exhibited performances to create spatial-gradient reactive plasma space on one wafer, and plasma diagnostic tools for detecting plasma and radical densities can be swept to obtain a database in a single experiment. Simultaneously, a processed wafer with thin films, which are affected by plasma with spatial-gradient parameters at each location, is analyzed by various ex-situ measurements. These plasma diagnostics performed on a single substrate can provide inputs, and the ex-situ measurements can be for obtaining outputs for supervised learning; these input–output datasets are completed as cause-effect connections, and large databases are ready for use in machine learning.

As a tabular summary, focusing on ANNs, we show in Table I historical achievements and currently-developing methods which have been described in this topical review. Future and possible advances are not restricted to the above-mentioned promising research seeds; these examples are within the best of our knowledge, and as the technology of machine learning or other plasma diagnostic methods rapidly increase and are being developed, feasibilities for the hybridization between ICT-based methods and plasma science and technology will expand day by day.

### 6. Concluding remarks

This topical review surveys the complexities in low-temperature plasma and the application of machine-learning methods to unveil these complexities. Since various phenomena in plasma are so complicated with cause-effect interconnections, conventional formats like a small set of mathematical equations are frequently insufficient to describe them. To perform these tasks, we show the way how the complexities are exhibited in a style of complex networks and how we can analyze them. Also, due to these complexities, advances in low-temperature plasma science and engineering sometimes cannot catch up with the rapid expansion of industrial demands, and machine learning can be a promising candidate to reinforce them. To select an adequate machine-learning method for configuring a model, we have to pay attention to cause-effect connections in a given plasma system, and sufficient experimental datasets are required, whose amounts are far beyond the level in the conventional measurement.

| Years       | Reported scientific/technological progresses                                      | Examples                       |
|-------------|-----------------------------------------------------------------------------------|--------------------------------|
| ~1960       | Proposals of perceptrons (basic ANNs)                                             | Ref. 1                         |
| 1989–1993   | Proofs of ANN capability for universal approximator                                | Refs. 9, 10                    |
| 1990–2000   | Establishment of fundamental bases for ANNs                                       | Refs. 2–4                      |
| 1994–2004   | ANN model for plasma chemical vapor deposition                                      | Ref. 22                        |
| 2000–2015   | ANN model for dry etching                                                          | Ref. 23                        |
| 2010–2015   | ANN model for atmospheric plasma spray                                              | Ref. 21                        |
| 2015–2015   | Collaborative model with numerical simulation                                     | Refs. 84, 85                   |
| 2015–2015   | Machine-learning model for fusion plasma                                           | Ref. 90                        |
| 2015–2015   | ANN model based on data logging in factories                                       | Refs. 36–38                    |

Table I. Summary of ANN-related studies described in this study, in general points of view and around plasma science and technology.
procedures; thus, progress in data acquisitions are also needed in the base of plasma diagnostics, and we list capable tools and schemes that are ready for use. Then, both unsupervised-learning and supervised-learning methods can be possible to clarify a new fact that would not be found out otherwise. Not only ANNs, which are typical supervised-learning tools in free software and work as universal approximators but also other unsupervised-learning methods for clustering classification and dimensionality reduction are useful with practical outputs for plasma processing. Matching such automatic procedures in the virtual space with the historical plasma in science is also of major importance, and we show a machine-learning scheme mixed with the fundamental theory in plasma elementary collisions as an example. Stimulated by the rapid progress of ICT-based approaches, an increasing number of new schemes will be proposed from now on, and they will contribute to the enrichment of plasma science and technology along with ongoing research activities based on other schemes.

Acknowledgments

One of the authors (O.S.) thanks Mr. Y. Mizui for his data cleansing for Figs. 7(a) and 7(b), and Mr. K. Sakurai for reactor setting in the experimental performance shown in Fig. 12(c). One of the authors (O.S.) also thanks the members of the Regional ICT Research Center for Human, Industry and Future, The University of Shiga Prefecture, for useful discussion on ICT-based methods. This work was partly supported by a Grant-in-Aid for Scientific Research from the Japan Ministry of Education, Culture, Sports, Science and Technology (MEXT/JSPS KAKENHI) with Grant Nos. JP18H03690, JP18K18756, JP19K03813, and JP20KK0089.

ORCID iDs

Osamu Sakai https://orcid.org/0000-0003-0959-5146
Satoru Kagawuchi https://orcid.org/0000-0003-3115-8719
Tomoyuki Murakami https://orcid.org/0000-0003-0618-2055

1) F. Rosenblatt, *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms* (Spartan, Washington, D.C., 1962).
2) J. Hertz, A. Krogh, and R. G. Palmer, *Introduction to the Theory of Neural Computation* (Westview Press, Boulder, CO, 1991).
3) M. Bishop, *Pattern Recognition and Machine Learning* (Springer, Singapore, 2006).
4) M. I. Jordan and T. M. Mitchell, “Machine learning: trends, perspectives, and prospects,” *Science* **349**, 255 (2015).
5) R. Ramprasad, R. Batra, G. Pilaman, A. Mammoli-Kanakkithodi, and C. Kim, “Machine learning in materials informatics: recent applications and prospects,” *Comput. Mater. Sci.* **53**, 4 (2017).
6) J. Schmidt, M. R. G. Marques, S. Botti, and M. A. L. Marques, “Recent advances and applications of machine learning in solid state materials science,” *Comput. Mater. Sci.* **53**, 83 (2015).
7) P. Baldi and S. Brunak, *Bioinformatics: The Machine Learning Approach* (MIT Press, Cambridge, MA, 2001).
8) P. Larrañaga et al., “Machine learning in bioinformatics,” *Briefing in Bioinf.* **7**, 86 (2006).
9) G. Cybenko, “Approximation by superpositions of a sigmoidal function,” *Math. Control Signal Syst.* **2**, 303 (1989).
10) K. Hornik, “Some new results on neural network approximation,” *Neural Netw.* **6**, 1059 (1993).
11) Y. P. Raizer, *Gas Discharge Physics* (Springer, Berlin, 1987).
12) I. Adamovich et al., “The 2017 plasma roadmap: low temperature plasma science and technology,” *J. Phys. D: Appl. Phys.* **50**, 232001 (2017).
13) S. M. Sze, *VLSI Technology* (McGraw-Hill, New York, 1983).
14) P. J. Bruggeman et al., “Plasma–liquid interactions: a review and roadmap,” *Plasma Sources Sci. Technol.* **25**, 053002 (2016).
15) M. G. Kong, G. Kroezen, G. Morell, T. Nosenko, T. Shimizu, J. van Dijk, and J. L. Zimmermann, “Plasma medicine: an introductory review,” *New J. Phys.* **11**, 115012 (2009).
16) M. Itô, T. Ohta, and M. Hori, “Plasma agriculture,” *J. Korean Phys. Soc.* **60**, 937 (2012).
17) O. Auciello and D. L. Flamm, *Plasma Diagnostics* (Academic, New York, 1989).
18) I. H. Hutchinson, *Principles of Plasma Diagnostics* (Cambridge University Press, Cambridge, 2002).
19) H. C. Kim, F. Iza, S. S. Yang, M. Radmilović-Radjenović, and J. K. Lee, “Particle and fluid simulations of low-temperature plasma discharges: benchmarks and kinetic effects,” *J. Phys. D: Appl. Phys.* **38**, R283 (2005).
20) M. J. Kushner, “Hybrid modelling of low temperature plasmas for fundamental investigations and equipment design,” *J. Phys. D: Appl. Phys.* **42**, 194013 (2009).
21) G. Guessasma, F. I. Trifa, G. Montavon, and C. Coddet, “Al2O3-13% weight TeO2; deposit profiles as a function of the atmospheric plasma spraying parameters,” *Mater. Des.* **25**, 307 (2004).
22) B. Kim and G. S. May, “An optimal neural network process model for plasma etching,” *IEEE Trans. Semicond. Manuf.* **7**, 12 (1994).
23) S. S. Han, M. Ceiler, S. A. Bidstrup, P. Kohl, and G. May, “Modeling the proper tyes of PECVD silicon dioxide films using optimized back propagation neural networks,” *IEEE Trans. Compon. Packag. Manuf. Technol. A* **17**, 174 (1994).
24) R. Albert and A.-L. Barabási, *Statistical mechanics of complex networks,* Rev. Modern Phys. **74**, 47 (2002).
25) M. E. J. Newman, “The structure and function of complex networks,” *SIAM Rev.* **45**, 167 (2003).
26) A.-L. Barabási, *Network Science* (Cambridge University Press, Cambridge, 2016).
27) O. Sakai, T. Kitagawa, K. Sakurai, G. Itami, S. Miyagi, K. Noborio, and K. Taguchi, “In-vacuum active colour sensor and wireless communication across a vacuum-air interface,” *Sci. Rep.* **11**, 1364 (2021).
28) G. Gershenfeld, R. Krishnan, and D. Cohen, “The Internet of things,” *Sci. Am.* **291**, 76 (2004).
29) M. Kranz, P. Holleis, and A. Schmidt, “Embedded interaction: interacting with the internet of things,” *IEEE Internet Comput.* **14**, 46 (2010).
30) L. Atzori, A. Iera, and G. Morabito, “The internet of things: a survey,” *Comput. Netw.* **54**, 2787 (2010).
31) Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature* **521**, 436 (2015).
32) E. D. Kolaczyk, *Statistical Analysis of Network Data: Methods and Models* (Springer, Berlin, 2009).
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**, 866 (2019).
34) Y. Chen, L. Lu, G. E. Karniadakis, and L. D. Negro, “Physics-informed neural networks for inverse problems in nano-optics and metamaterials,” Opt. Express **28**, 11618 (2020).
35) S. R. Haskey, B. A. Grierson, C. Chrystal, A. Ashourvan, K. H. Burrell, R. J. Groebner, E. A. Bell, L. Stagner, D. J. Battaglia, and T. Stoltzfus-Dueck, “Main ion and impurity edge profile evolution across the L- to H-mode transition on DIII-D,” *Plasma Phys. Contrib. Fusion* **60**, 105001 (2018).
36) E. A. Baltz, E. Trask, M. Binderbauer, M. Dikovskiy, H. Gota, R. Mendoza, J. C. Platt, and P. F. Riley, “Achievement of sustained net plasma heating in a fusion experiment with the optimetrist algorithm,” *Sci. Rep.* **7**, 6425 (2017).
37) C. Rea and R. S. Granetz, “Exploratory machine learning studies for disruption prediction using large databases on DIII-D,” *Fusion Sci. Technol.* **74**, 89 (2018).
38) J. Kates-Harbeck, A. Svyatkovskiy, and W. Tang, “Predicting disruptive instabilities in controlled fusion plasmas through deep learning,” *Nature* **568**, 526 (2019).
39) T. F. Connor and J. B. Taylor, “Scaling laws for plasma confinement,” *Nucl. Fusion* **17**, 1047 (1977).
40) M. L. Brake, J. Pender, and J. Fournier, “The gaseous electronic conference (GEC) reference cell as a benchmark for understanding microelectronics processing plasmas,” *Phys. Plasmas* **6**, 2307 (1999).
41) J. K. Olthoff and K. E. Greenberg, “The gaseous electronic conference rf reference cell—an introduction,” *Res. Natl. Inst. Stand. Technol.* **100**, 327 (1995).
42) Creative Commons (accessed April 2, 2022) (https://creativecommons.org/)
43) A. Bonaccorsi and C. Rossi, “Why open source software can succeed,” *Res. Policy* **32**, 1243 (2003).
O. Sakai, K. Nobuto, S. Miyagi, and K. Tachibana, M. Prosperi, Y. Guo, M. Sperrin, J. S. Koopman, J. S. Min, X. He, S. Rich, M. J. Kushner, Y. Mizui, T. Kojima, S. Miyagi, and O. Sakai, Jpn. J. Appl. Phys. Y. Mizui, S. Miyagi, and O. Sakai, S. T. Reweis and L. K. Saul, S. Brin and L. Page, O. N. Temkin, A. V. Zeigarnik, and D. Bonchev, D. Montgomery, C. Jennings, and M. Kulahci, I. Fishtik, C. A. Callaghan, and R. Datta, Y. Mizui, S. Miyagi, and O. Sakai, Complex Networks VIII (Springer, Cham, 2017), p. 135. Y. Mizui, S. Miyagi, and O. Sakai, “Statistics of growing chemical network originating from one molecule species and activated by low-temperature plasma,” Complex Networks & Their Applications IX (Springer, Cham, 2020), p. 231. E. B. Lee, Foundations of Optimal Control Theory (Wiley, Hoboken, NJ, 1967). D. Montgomery, C. Jennings, and M. Kulach, Introduction to Time Series Analysis and Forecasting (Wiley, Hoboken, NJ, 2008). O. N. Temkin, A. V. Zeigarnik, and D. Bonchev, Chemical Reaction Networks (CRC Press, Boca Raton, FL, 1996). I. Fishik, C. A. Callaghan, and R. Datta, “Reaction route graphs: III. Nonlinear minimal kinetic mechanisms,” J. Phys. Chem. B 110, 2710 (2005). P. D. Leenheer, D. Angel, and E. D. Sontag, “Monotone chemical reaction networks,” J. Math. Chem. 41, 295 (2007). C. D. Bie, J. Dijk, and A. Bogaerts, “The dominant pathways for the conversion of methane into oxygenates and syngas in an atmospheric pressure dielectric barrier discharge,” J. Phys. Chem. C 119, 22313 (2015). A. N. Gorban and G. S. Yablonsky, “Extended detailed balance for systems with irreversible reactions,” Chem. Eng. Sci. 66, 5388 (2011). T. Murakami, K. Niemi, T. Gans, D. O’Connell, and W. G. Graham, “Chemical kinetics and reactive species in atmospheric pressure helium-oxygen plasmas with humid-air impurities,” Plasma Sources Sci. Technol. 22, 015003 (2013). R. Albert, H. Jeong, and A.-L. Barabási, “Error and attack tolerance of complex networks,” Nature 406, 378 (2000). H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási, “The large-scale organization of metabolic networks,” Nature 407, 651 (2000). S. Brin and L. Page, “The anatomy of a large-scale hypertextual web search engine,” Comput. Netw. ISDN Syst. 30, 107 (1998). E. Agliari, A. Barra, S. Bartolucci, A. Galluzzo, F. Guerra, and F. Mauroa, “Parallel processing in immune networks,” Phys. Rev. E 87, 42701 (2013). E. Agliari, A. Barra, L. D. Schiavi, and A. Moro, “Complete integrability of information processing by biochemical reactions,” Sci. Rep. 6, 36314 (2016). A. Ortega, P. Frossard, J. Kova’cevic, J. M. F. Moura, and P. Vandergheynst, “Graph signal processing: overview, challenges and applications,” Proc. IEEE 106, 808 (2018). A. R. Koss et al., “Dimensionality-reduction techniques for complex mass spectrometric datasets: application to laboratory atmospheric organic oxidation experiments,” Atmos. Chem. Phys. 20, 1021 (2020). S. T. Rewetis and L. K. Saul, “Nonlinear dimensionality reduction by locally linear embedding,” Science 290, 2323 (2000). A. L. Ferguson, A. Z. Panagiotopoulos, P. G. Debenetti, and I. G. Kevrekidis, “Systematic determination of order parameters for chain dynamics using diffusion maps,” Proc. Natl. Acad. Sci. 107, 13597 (2010). M. Dworinik, S. Mukherjee, C. Jayaprakash, and I. Das, “Dramatic reduction of dimensionality in large biochemical networks owing to strong pair correlations,” J. R. Soc. Interface 9, 1824 (2012). K. D. Bleecker, A. Bogaerts, W. Goedheer, and R. Gijbels, “Investigation of growth mechanisms of clusters in a silane discharge with the use of a fluid model,” IEEE Trans. Plasma Sci. 32, 691 (2004). A. Ivey, H. Lowen, G. Fortioli, and C. P. Royall, Complex Plasmas and Collodial Dispersions: Particle-resolved Studies of Classical Liquids and Solids (World Scientific, Singapore, 2012). S. H. Hwang, T. Okumura, K. Kamataki, N. Itagaki, K. Koga, and M. Shiratani, “Size and flux of carbon nanoparticles synthesized by Ar + CH4 multi-hollow plasma chemical vapor deposition,” Diam. Relat. Mat. 109, 108090 (2020). J. Fox (ed.) Proceedings of the Symposium on Generalized Networks (Polytechnic Press, New York, 1966). K. Menger, “Das botenproblem,” Ergebnisse Eines Mathematischen Kolloquiums 2, 11 (1932). N. Christofides, A. Mingozzi, and P. Toth, “Exact algorithms for the vehicle routing problem, based on spanning tree and shortest path relaxations,” Math. Program. 20, 255 (1981). S. Fortunato, “Community detection in graphs,” Phys. Rep. 486, 75 (2010). T. W. Anderson, An Introduction to Multivariate Statistical Analysis (Wiley, New York, 1984). G. J. M. Hagelaar and L. C. Pitchford, “Solving the Boltzmann equation to obtain electron transport coefficients and rate fluid models,” Plasma Sources Sci. Technol. 14, 722 (2005). S. S. Tezcan, M. A. Akcayol, O. C. Ozerdem, and M. S. Dincer, “Calculation of electron energy distribution functions from electron swarm parameters using artificial neural network in SF6 and argon,” IEEE Trans. Plasma Sci. 38, 2332 (2010). S. Kawaguchi, K. Takahashi, H. Okhama, and K. Satoh, “Deep learning for solving the Boltzmann equation of electrons in weakly ionized plasma,” Plasma Sources Sci. Technol. 29, 025021 (2020). C. Kuenzer, M. Ottinger, M. Wegmann, H. Guo, C. Wang, J. Zhang, S. Decha, and M. Wikelski, “Earth observation satellite sensors for biodiversity monitoring: potentials and bottlenecks,” Int. J. Remote Sens. 35, 6599 (2014). G. Chandra, B. L. Markham, and D. L. Helder, “Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors,” Remote Sens. Environ. 113, 893 (2009). J. P. Lebreton et al., “An overview of the descent and landing of the Huygens probe on Titan.” Nature 438, 758 (2005). M. Abe et al., “Near-infrared spectral results of asteroid Itokawa from the Hayabusa spacecraft,” Science 312, 1334 (2006). F. Krüger, T. Gers, and J. Trieschmann, “Machine learning plasma-surface interface for coupling sputtering and gas-phase transport simulations,” Plasma Sources Sci. Technol. 28, 035002 (2019). A. Jalal, C. Heisstracher, A. Schindler, B. Haslhofer, T. Nemeth, R. Glawar, W. Sihn, and P. De Boer, “Predicting time-to-failure of plasma etching equipment using machine learning, 2019,” Proc. IEEE Int. Conf. on Prognostics and Health Management (San Francisco). C. S. Moon, K. Takeda, M. Sekine, Y. Setsuhara, M. Shiratani, and M. Horii, “Combinational plasma etching process,” Appl. Phys. Express 2, 096001 (2009). Y. Setsuhara, K. Cho, K. Takenaka, M. Shiratani, M. Sekine, and M. Horii, “Combinational analyses of plasma polymer interactions,” Surf. Coat. Technol. 205, 5484 (2011). H. Koizuma and I. Takeuchi, “Combinational solid-state chemistry of inorganic materials,” Nat. Mater. 3, 429 (2004). Osamu Sakai received B.S., M.S., and Ph.D. degrees from Kyoto University, Kyoto, Japan, in 1990, 1992, and 1996, respectively. He was a Research Engineer (1995-2002) with the Sharp Corporation, Nara, Japan, and also he was an Assistant Professor and an Associate Professor with the Department of Electronic Science and Engineering, Kyoto University (2003-2014). He is currently a Professor with the Department of Electronic Systems Engineering and a Chair with the Regional ICT Research Center for Human, Industry and Future at the University of Shiga Prefecture. His current research interests include network science, monitoring technology, plasma science and engineering, and metamaterial science.
Tomoyuki Murakami received a B.E. degree from the Tokyo Science University, Chiba, Japan in 1993 and M.E. and Ph.D. degrees from the Tokyo Institute of Technology, Yokohama, Japan, in 1995 and 1998, respectively. He was an Assistant Professor with the Tokyo Institute of Technology and has been a Visiting Research Professor in the School of Mathematics and Physics, Queen’s University Belfast, NI, UK. He is currently a Professor with the Department of Science and Technology, Seikei University, Tokyo, Japan. His current research interests include numerical modeling on plasma science and technology.

Satoru Kawaguchi was born in Hokkaido, Japan, in 1990. He received a Dr. Eng. degree from Muroran Institute of Technology, Hokkaido, Japan, in 2018. He was a Research Fellow of Japan Society for the Promotion of Science from 2017 to 2019 and an Assistant Professor with the Department of Systems Design Engineering, Seikei University from 2019 to 2022. He is currently an Assistant Professor with the Division of Information and Electronic Engineering, Graduate School of Engineering, Muroran Institute of Technology. His current research interests include plasma science and physics-informed machine learning.