Materials Informatics for Process and Material Co-Optimization

Fumiaki Tanaka, Hiroshi Sato, Naoki Yoshii, and Hidefumi Matsui

***Abstract*—In semiconductor manufacturing, fabrication processes and their materials should be properly co-optimized to achieve required processing results within reasonable development duration and acceptable cost. Unfortunately, it is a very time-consuming procedure, because the number of possible combinations of process/material candidates is very large. Here, we develop a methodology for co-optimization of processes and their materials. We successfully constructed a prediction model for dry-etching of high-k materials ($R^2 = 0.65$). Also, it was proven that considering both the materials and processes is needed for accurate prediction of etching rates. By trying only <0.00001% of all possible process/material candidates with this model and Bayesian optimization, we can find new combinations of gasses and their processes for more than 100 times higher etching rates than that with a traditional gas/process condition. Furthermore, we discussed that accurate prediction can be made by using a combination of the Bayesian optimization with LASSO and materials knowledge from related scientific papers. Future work will focus on validating the versatility of our methodology by applying it to other development items.**

***Index Terms*—Materials informatics, process optimization, semiconductor materials, artificial intelligence.**

I. Introduction

RECENTLY, a new research field called "Materials Informatics" has emerged, owing to tremendous advances in big data analysis, machine learning, materials database, and simulation techniques. In the field, novel materials that satisfy given requirements have been designed. For several years, some inorganic materials such as battery components have been successfully discovered [1]–[4]. Mo *et al.* used first principles calculation to search for a component of lithium ion battery [1]. Surprisingly, without any experiment in a lab, they succeeded in finding a new material for a solid-state electrolyte. Seko *et al.* also used first principles calculation to search for materials with low lattice thermal conductivities [2]. They found some promising materials for thermoelectric applications in silico. Hinuma *et al.* designed a novel crystal structure for a light-emitting material using first-principles calculations [3]. They actually synthesized the novel material and proved that it emitted red light as expected. Nishijima *et al.* used density functional theory calculations to design cathodes for lithium ion battery [4]. They proved that the designed materials showed long cycle-life time as a lithium ion battery. In these cases, the performances of materials were decided by only the compositions or structures of materials (this corresponds to the "Materials-Property/performance" plane in Fig. 1).

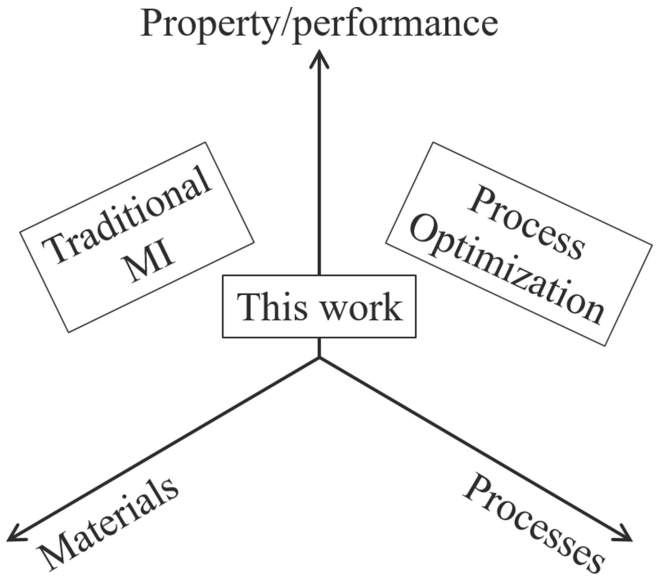

Fig. 1. Difference between traditional methodologies and ours. Our methodology co-optimizes both of materials and processes at the same time, while previous works optimized one of them.

In contrast, in the field of semiconductor manufacturing, the fabrication processes as well as materials should be simultaneously optimized (in other words "co-optimized"). For decades, advances in semiconductor manufacturing processes have been increasing the number of materials used for device production [5]. Unfortunately, this results in an exponential increase of materials and their processes, and poses a difficult challenge to select the optimal combination from the enormous number of options. This leads to increase the cost and development period for device production in semiconductor industry. This hampers the growth of semiconductor manufacturing due to the reduction in business profits. To address this, methodologies for process optimization were well studied [6]–[8]. Chopra *et al.* developed a software tool

Manuscript received August 13, 2019; revised September 1, 2019 and September 14, 2019; accepted September 18, 2019. Date of publication September 24, 2019; date of current version October 29, 2019. *(Corresponding author: Fumiaki Tanaka.)*

F. Tanaka, H. Sato, and N. Yoshii are with the Development Strategic Department, Tokyo Electron Limited, Tokyo 107-6325, Japan (e-mail: fumiaki.tanaka@tel.com; hiroshi.sato@tel.com; naoki.yoshii@tel.com).

H. Matsui is with the Simulation Technology Department, Tokyo Electron Technology Solutions Limited, Yamanashi 407-0192, Japan (e-mail: hidefumi.matsui@tel.com).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSM.2019.2943162

for creating plasma etch recipes based on physical models and Bayesian inference [6] and applied it to prediction of the etching results [7]. They showed that the etching rate of SiO_2 with CF_4/Ar gasses could be predicted ($R^2 = 0.63$). Suzuki *et al.* adopted machine learning approaches to optimize the process results such as the uniformity of plasma enhanced atomic layer deposition [8]. They succeeded in achieving the target values and even outperforming the knowledgeable engineers. However, these works focused on process optimization rather than co-optimization of materials and their processes. In other words, they searched for optimum solutions in the "Processes-Property/performance" plane in Fig. 1. Although the co-optimization is needed in semiconductor manufacturing as mentioned above, such a methodology have not yet been established.

Herein, the purpose of this study is to develop a methodology that can reduce cost and time of the process development by co-optimizing the materials and their processes. By achieving this, we can efficiently find the novel optimal combinations of materials and processes that are difficult to find through traditional trial-and-error techniques by engineers. Such a methodology can decrease the number of experiments and eliminate the cost and time of the process development because they are approximately proportional to the number of experiments. Previously, in International Symposium on Semiconductor Manufacturing (ISSM), we develop a predictive method to co-optimize the processes and their materials by combining Bayesian optimization with data extracted from scientific papers, experimental data, and material databases [9]. Fig. 1 illustrates the novelty of our work compared to previous studies. To optimize the property or performance of materials and processes, we co-optimize both of them. In contrast to the materials informatics and process optimization mentioned above, we have focused on searching for optimal solutions in the entire space in Fig. 1, which consists of the "Materials", "Processes", and "Property/performance".

Here, we expanded our work to include latest results and considerations associated with discussion in the conference. Although we could find some candidates of gasses and processes for high-k etching, we were faced with two issues: (i) the etching rate was below the target value, and (ii) the candidates were not so non-trivial from an engineers' perspective, as pointed out by one of the conference's audience members. In this paper, we can find another promising candidate of gasses and processes for high-k etching. Interestingly, the candidate differed from what human engineers designed and resulted in 100 times higher etching rate than that with a traditional gas/process condition. In addition, we enriched "Discussion" section to clarify the effectiveness of our methodology and updated "Introduction" and "Materials and methods" to help the readers to understand.

II. Materials and Methods

A. Methodology for Co-Optimization

Our methodology is illustrated in Fig. 2 with the example of maximizing high-k etching rate. In short, the methodology can be divided into four parts; (1) producing a predictive model, (2) co-optimizing based on the model, (3) performing experiments according to the result of the co-optimization, and (4) improving the predictive model with the obtained experimental data. After part (4), the procedure returns to part (2). By repeating the procedures of parts (2-4), the predictive model can be improved continuously. Each part of our methodology is explained below.

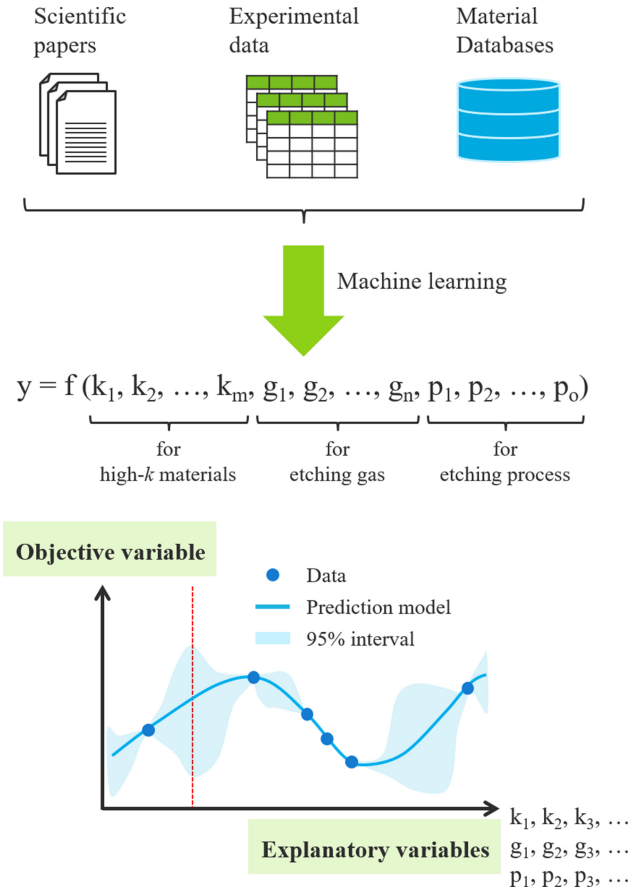

Fig. 2. Schematic illustration of the methodology for process and material co-optimization. Predictive model for etching rate (denoted by y) based on machine learning, including three types of explanatory variables; those for high-k materials ($k_1, k_2, \ldots, k_m$), etching gas ($g_1, g_2, \ldots, g_n$), and etching process ($p_1, p_2, \ldots, p_o$). The function f is implicitly modeled in the case of Bayesian optimization.

In order to validate our methodology, we must choose an appropriate development item for proof of concept such that (1) the results depend on both materials and their processes and (2) we can obtain enough training data to make a predictive model. With these criteria in mind, we chose a development item that searches for gas materials and their processes for dry cleaning of high-k thin films (ZrO_2, HfO_2, and Al_2O_3). In this case, the model predicts the etching rate from three types of explanatory variables; those for high-k materials, etching gas, and the etching process. To co-optimize processes and materials, a predictive model for the etching rate was formulated using machine learning such as Gaussian process regression (GPR). GPR is a nonparametric method to provide a regression model [10]. Although GPR is based on a linear model, it can be expanded to a nonlinear regression model

using appropriate kernels. We used a well-known kernel, called "radial basis function (RBF)" to obtain a nonlinear predictive model. GPR can consider even the interaction effect between explanatory variables by using kernels, such as RBF. In fact, GPR is known to be mathematically equivalent to some type of deep learning algorithms [11]. Compared to other regression models, GPR shows some benefits, such as working well on small datasets and can provide confidence intervals as well as predicted values. These advantages among others are well fitted to the problem addressed in this study. In particular, confidence interval values have proven to possess valuable information since it describes the uncertainty of predictions. Using predicted values and their confidence intervals, the most promising parameters (i.e., combinations of materials and their process parameters) can be estimated statistically as shown in the red dashed line in Fig. 2. Using the suggested parameters, the next experiment is decided and then the result is used to update the predictive model. Through a repetition of this procedure, the predictive model is continuously improved after each experiment, resulting in improving the suggested parameters.

The detailed explanatory variables employed in this case are listed in Table I. The variables of high-k materials and gasses (e.g., bond strength) were extracted from two material databases [12], [13]. However, there was no guarantee that all explanatory variables in Table I were effective in the accurate prediction of the etching rate. If invalid variables were included in the predictive model, it could bring an overfitting, which may reduce the prediction accuracy. Thus, before modeling, effective variables for etching rate prediction were selected using one well-known sparse modeling method called "Least Absolute Shrinkage and Selection Operator (LASSO)". This modeling method provides a linear regression model while trying to produce zero coefficients when the corresponding variable respond negatively to the prediction accuracy by adding a penalty to the coefficient values [14]. Therefore, the model can be used to exclude invalid variables that insignificantly contribute to the prediction. Note that we used LASSO only to exclude invalid explanatory variables. Our predictive model was constructed using GPR rather than LASSO, due to the inability of LASSO to construct a nonlinear regression model. After variable selection by LASSO, we adopted 19 variables by excluding four invalid variables (as denoted '*' in Table I). For the data used for the predictive model, we used 36 pieces of our own experimental data and 335 pieces of experimental data extracted from 24 scientific papers. Before training the predictive model, the output data (i.e., etching rates) were normalized by the average (μ) and standard deviation (σ) using the equation of (etching rate $- \mu)/\sigma$.

With the model mentioned above, co-optimization of the etching process and materials was done by Bayesian optimization with Gaussian processes. Bayesian optimization as described, is an optimization method used to search for maximum (or minimum) values of black-box functions (i.e., functions whose shapes are unknown) [15]. This optimization method typically uses the predictive values and their confidence intervals obtained from GPR to suggest the most promising candidate parameters that are expected to obtain the maximum (or minimum) value of the black-box function. How the candidate parameters are promising is evaluated by an acquisition function. In this study, we used a typical acquisition function, known as "Upper Confidence Bound", with which parameters are considered to be the most promising such that the summation of the predictive value and its confidence interval is maximum or minimum. Compared to other machine learning approaches, the advantage of Bayesian optimization is the use of confidence intervals to evaluate the candidate based on an acquisition function. This advantage is critical to search for extrapolated candidates that deviate from a range of well-known parameters. As such, Bayesian optimization suggests the most promising gas combination and the process based on the predicted value and the confidence interval (Fig. 2). The model was programmed in Python with scikit-learn library.

TABLE I
EXPLANATORY VARIABLES FOR PREDICTION OF ETCHING RATE

Process parameters	Flow rate of etching gas 1 (sccm)
	Flow rate of etching gas 2 (sccm)
	Flow rate of additive gas (sccm)
	Flow rate of carrier gas (sccm)
	Temperature (deg C)
	Pressure (Torr)
	Plasma power (W)
	Substrate bias (V) / Substrate bias (W)
High-k material properties [a]	Formation energy (eV) *
	Band gap (eV) *
	Volume ($Å^3$)
	Number of sites in the unit cell *
	Density (g/cm^3)
	Bond strength (eV) *
Etching gas 1 [b] properties	Bond strength (eV)
	Polarizability
Etching gas 2 [c] properties	Bond strength (eV)
	Polarizability
Additive gas [d]	Bond strength (eV)
	Polarizability
Carrier gas [e]	Bond strength (eV)
	Polarizability

* these parameters were removed as a result of variable selection by LASSO
[a] ZrO_2, HfO_2, or Al_2O_3, [b] BCl_3, CH_4, CF_4, C_2F_6, ClF_3, HBr, NF_3, SF_6, or $SiCl_4$, [c] Cl_2 or C_4F_8, [d] CO, O_2, or H_2, [e] Ar or N_2

TABLE II
PROCESS CONDITIONS

Substrate temperature	400 - 700 (deg C)
Total pressure	0.01 - 5 (Torr)
Total gas flow	10 – 1,500 (sccm)
Process time	5 - 30 (min)

As shown in Table I, we considered nine process parameters, combined with two types of etching gasses, additive gas, and carrier gas (Table I). All the combination of the candidate gasses is 359 ($= 10 \times 3 \times 4 \times 3 - 1$, see the footnote of Table I). Thus, the number of all possible combinations of process/material candidates was 701,171,875 ($= 5^9 \times 359$), when each process parameter has five levels. Compared to a traditional design of experiments based on orthogonal arrays, Bayesian approach has some advantages in designing experiments. First, Bayesian optimization can be used even when the size of parameters and their levels are huge. In the case of this study, it was difficult to utilize the traditional design of experiments, since the number of parameters and their levels was beyond the size of any orthogonal array. In addition, another advantage of Bayesian optimization is its ability to use the latest experimental result for designing the next experiment. By updating the regression model after each experiment, the next experiment is well designed through the complete utilization of the experimental results. In contrast, with the traditional approach, all experimental conditions are fixed through the experiments until every experiments are done based on the orthogonal array.

B. Validation of our Methodology

The predictive model obtained by GPR was evaluated statistically using the following typical indicators accompanied with a scatter plot of observed vs. predicted values:

- *Coefficient of Determination (R^2):* $\left(\frac{s_{pe}}{s_p s_e}\right)^2$,
- *Root Mean Square Error (RMSE):* $\sqrt{\frac{\sum_{i=1}^{N}(pred_i - expel_i)^2}{N}}$,
- *Mean Absolute Error (MAE):* $\frac{\sum_{i=1}^{N}|pred_i - expel_i|}{N}$,

where s_{pe} denotes the covariance between predictions and experiments; s_p and s_e denote the standard deviation of predictions and experiments, respectively; and $pred_i$ and $expel_i$ denote the i-th prediction and experimental values, respectively. The predictive values were estimated by leave-one-out cross validation, where one of the data is used as a test data, while the others are used to train the model.

The suggested gasses and processes were experimentally tested whether high-k films can be etched with the conditions. Sputtered high-k films (∼1 μm) on a sapphire substrate were processed with a remote plasma source in an electric furnace under the suggested condition. The process conditions are listed in Table II. We tried 55 conditions (i.e., gasses and processes) to maximize the etching rates of high-k films based on the co-optimization procedure mentioned above. The target value of the etching rate was more than 500 nm/min. Unfortunately, 335 pieces of data obtained from the 24 scientific papers could not achieve this target. Thus, to achieve the target, the extrapolated data should be explored in a region deviated from the known parameters. In summary, the problem we focused on was to achieve higher etching rate than that of any training data within a limited number of experiments, which was only 14.8% ($= 55/(335+36)$) of the total amount of training data.

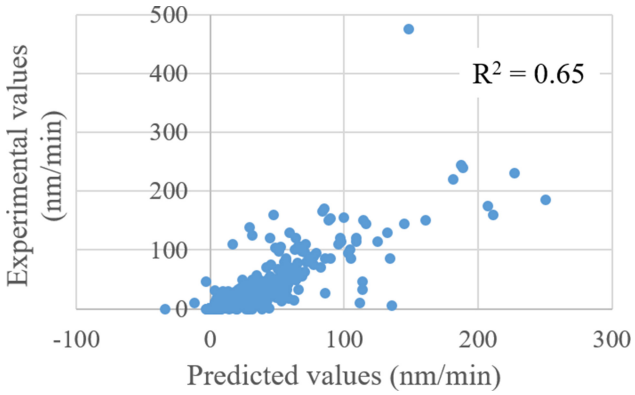

Fig. 3. Comparison between predictions by the model and experimental data from our laboratory and published research.

To evaluate the found materials and processes through our co-optimization procedure, we compared with the results obtained by a traditional gas combination and its process provided by Nakamura *et al.* [16]. They etched a thin film of HfO_2 at 75 deg C using gasses of BCl_3, Cl_2, and O_2 (37.5:62.5:5) with a total flow rate of 40 sccm at 5 mTorr and achieved a relatively high etching rate of ∼150 nm/min. To fairly compare with this result, we used the same gas combination and their processes as much as possible.

III. RESULTS

With parameters listed in Table I, the predictive model can estimate the etching rates of high-k films. Fig. 3 shows the comparison between experimental values extracted from our laboratory and published research and predicted ones estimated by the model. The result indicated a high coefficient of determination ($R^2 = 0.65$), which was calculated with 371 pairs of predicted and experimental values. In general, the coefficient of determination is considered as high according to the interpretation of the R value, denoting that $R \geq 0.7$ (i.e., $R^2 \geq 0.49$) can be regarded as "a strong linear relationship" [17]. In addition, the value was equivalent to that of the prediction by Chopra *et al.* [7]. They achieved a similar coefficient of determination ($R^2 = 0.63$) to predict the etching rate of SiO_2 with CF_4/Ar gasses. In this case, the materials were fixed, whereas process parameters were variable. In contrast, in our case, both materials and process parameters are variable. Nevertheless, the coefficient of determination achieved by our model was equivalent to that of the model by Chopra *et al.*, indicating that the coefficient of determination by our model was sufficiently high. Furthermore, when we apply a simple

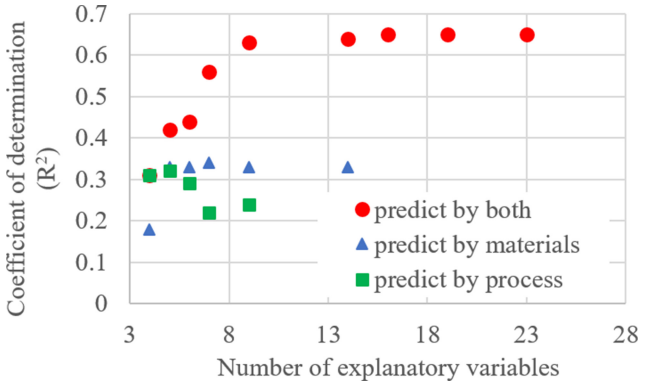

Fig. 4. Comparing prediction accuracy when using (1) both etching process and materials, (2) only materials, and (3) only the etching process.

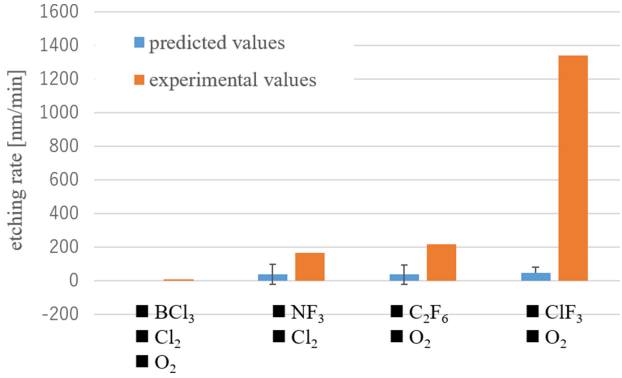

Fig. 5. Discovered gas combination and resultant etching rates of HfO_2 using discovered processes at 500 deg C. The condition on the left (BCl_3 + Cl_2 + O_2) is a traditional one obtained from literature [16], although the process is a little bit different, because high vacuum was not available.

outlier detection method, the extraordinary data (experimental value: 475 nm/min in Fig. 3) can be excluded because the value exceeds the threshold (= average + 3 standard deviation). When the data was excluded, the coefficient of determination gets higher ($R^2 = 0.70$). These results supported our hypothesis that the etching rates can be estimated by material properties and process parameters. Furthermore, RMSE and MAE were 29.4 and 15.0, respectively. The ratio of RMSE to MAE was 1.96, which exceeded the ideal value of 1.25 when the model suffered from only the Gaussian noise. This implied that more explanatory variables were needed to predict the etching rates accurately (see Discussion).

In order to validate the effectiveness of considering both etching process and materials (high-*k* films and gas), we investigated the prediction accuracies when using either etching process or materials (Fig. 4). When considering both of them together, the coefficient of determination was sufficiently high between 9 to 19 explanatory variables ($R^2 > 0.64$). In contrast, when considering only one of them, the coefficient was quite low ($R^2 < 0.33$). Note that the former prediction accuracy was better even when the number of explanatory variables was less than that of the latter. This result demonstrates that considering both the etching process and materials is very important in accurately predicting the etching rate. This also indicates that co-optimization of the etching process and materials is necessary in the field of semiconductor manufacturing.

Then, using Bayesian optimization, we could find some new kinds of gas combinations to achieve high etching rate (>1,200 nm/min). Fig. 5 shows the discovered gas combination and the resultant etching rates of HfO_2. As shown in the figure, the etching rates achieved by discovered gas combinations were 100 times higher than that of the traditional gas combination obtained from the literature [16]. We emphasize that only 55 experiments (equivalent to <0.00001% of all possible process/material candidates) were needed to discover the new conditions. Furthermore, the achieved etching rate (>1,200 nm/min) exceeded the target value (500 nm/min) by performing only 14.8% of the times of experiments in the total amount of the training data, although our previous work did not achieve this target value [9].

To quantitatively compare with the conventional methodology, we estimated the necessary number of experiments, which is proportional to the cost and development time. For this comparison, we hypothesized that an experimenter would try the same gas combinations we used and optimize the process parameters by using a traditional design of experiment. In our case, we tried 15 gas combinations within 55 experiments. For the process optimization in each gas combination, the experimenter can utilize the L_{50} orthogonal array, which includes 11 five-level parameters. If the values of nine process parameters in Table I are limited to five-level, the L_{50} orthogonal array can cover them. In such a case, 750 ($= 15 \times 50$) experiments should be carried out in total. Therefore, under the hypothesis mentioned above, we needed only 7.3% of experiments compared to the conventional methodology. In other words, we could eliminate 92.7% of cost and development time by the methodology for efficiently co-optimizing materials and their process parameters.

IV. Discussion

Experimental values in Fig. 5 were much greater than the predicted ones (e.g., the experimental value was 1,343.0 nm/min, while the predicted one was 46.6 nm/min in the case of ClF_3 and O_2). This is because the found gasses were obtained as a result of the exploration of extrapolated parameters. Because the target value (>500 nm/min) was higher than any training data (i.e., mainly published research) as shown in Fig. 3, we should search the extrapolated parameters far from well-known parameters extracted from the training data. In general, it is very difficult to accurately predict the results of such extrapolated parameters with any machine learning technique. Nevertheless, Bayesian optimization worked well to find gasses and processes to increase the etching rate of HfO_2. This was because the predictive model could exclude unpromising parameters whose etching rates were probably low and suggest the relatively promising ones even if the prediction accuracy was not good due to their extrapolation. As exemplified in Fig. 5, Bayesian optimization is one of the appropriate tools for promising but unknown processes and materials as a result of their co-optimization.

There are two important reasons for our methodology to find appropriate conditions. One is a good selection of gasses. We chose gas candidates according to the relevant papers.

Therefore, promising candidates (halogens in this case) can be selected in advance. Another reason is usage of LASSO to decrease the number of explanatory variables in the predictive model. Because it limited the search space consisting of processes and materials, it can increase the search efficiency to find good solutions within a limited time. In today's semiconductor manufacturing world, more than 61 elements from the periodic table are used, which have resulted in an increased cost of research and development [5]. The increased number of elements lead to an enormous number of materials and associated processes, which can increase the research and development duration. Thus, by efficiently optimizing the combination of materials and their processes with smaller number of experiments as demonstrated in this paper, the research and development duration can be shortened, resulting in a decrease in the costs.

To further improve the predictive model and the search efficiency for appropriate gasses and processes, larger training data sets are needed. In addition, inclusion of more model parameters, such as increasing gas parameters, by-products, and hardware properties are another promising approach to improve the results.

Our methodology can be applied to totally different development items (e.g., high-quality film deposition using CVD/ALD) by changing the objective and explanatory variables if only we had sufficient experimental data. If this can be done successfully, the research and development duration can be decreased by the efficient optimization of materials and their processes. In the future, we plan on validating the versatility of our methodology by applying it to other development items.

REFERENCES

[1] Y. Mo, S. P. Ong, and G. Ceder, "First principles study of the $Li_{10}GeP_2S_{12}$ lithium super ionic conductor material," *Chem. Mater.*, vol. 24, no. 1, pp. 15–17, 2012. doi: 10.1021/cm203303y.

[2] A. Seko, A. Togo, H. Hayashi, K. Tsuda, L. Chaput, and I. Tanaka, "Prediction of low-thermal-conductivity compounds with first-principles anharmonic lattice-dynamics calculations and Bayesian optimization," *Phys. Rev. Lett.*, vol. 115, Nov. 2015, Art. no. 205901. doi: 10.1103/PhysRevLett.115.205901.

[3] Y. Hinuma *et al.*, "Discovery of earth-abundant nitride semiconductors by computational screening and high-pressure synthesis," *Nat. Commun.*, vol. 7, Jun. 2016, Art. no. 11962. doi: 10.1038/ncomms11962.

[4] M. Nishijima *et al.*, "Accelerated discovery of cathode materials with prolonged cycle life for lithium-ion battery," *Nat. Commun.*, vol. 5, p. 4553, Aug. 2014. doi: 10.1038/ncomms5553.

[5] M. L. Green *et al.*, "Fulfilling the promise of the materials genome initiative with high-throughput experimental methodologies," *Appl. Phys. Rev.*, vol. 4, no. 1, 2017, Art. no. 011105. doi: 10.1063/1.4977487.

[6] M. J. Chopra, R. Verma, Z. A. Lane, C. G. Willson, and R. T. Bonnecaze, "A method to accelerate creation of plasma etch recipes using physics and Bayesian statistics," in *Proc. SPIE*, vol. 10149, 2017, Art. no. 101490X. doi: 10.1117/12.2263507.

[7] M. J. Chopra, S. Helpert, R. Verma, Z. Zhang, X. Zhu, and R. T. Bonnecaze, "A model-based, Bayesian approach to the CF4/Ar etch of SiO_2," in *Proc. SPIE*, vol. 10588, 2018, Art. no. 105880G. doi: 10.1117/12.2297482.

[8] Y. Suzuki, S. Iwashita, T. Sato, H. Yonemichi, H. Moki, and T. Moriya, "Machine learning approaches for process optimization," in *Proc. Int. Symp. Semicond. Manuf.*, 2018, Art. no. PC035. doi: 10.1109/ISSM.2018.8651142.

[9] F. Tanaka, H. Sato, N. Yoshii, and H. Matsui, "Materials informatics for process and material co-optimization," in *Proc. Int. Symp. Semicond. Manuf.*, 2018, Art. no. PO020. doi: 10.1109/ISSM.2018.8651132.

[10] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006. [Online]. Available: http://www.gaussianprocess.org/gpml/

[11] C. K. I. Williams, "Computation with infinite neural networks," *Neural Comput.*, vol. 10, no. 5, pp. 1203–1216, 1998. doi: 10.1162/089976698300017412.

[12] A. Jain *et al.*, "Commentary: The materials project: A materials genome approach to accelerating materials innovation," *APL Mater.*, vol. 1, no. 1, 2013, Art. no. 011002. doi: 10.1063/1.4812323.

[13] R. D. Johnson, III, *NIST Computational Chemistry Comparison and Benchmark Database, Database Number 101, Release 19*, NIST, Gaithersburg, MD, USA, 2018. [Online]. Available: http://cccbdb.nist.gov/

[14] R. Tibshirani, "Regression shrinkage selection via the LASSO," *J. Roy. Stat. Soc. B (Stat. Methodol.)*, vol. 58, no. 1, pp. 267–288, 1996. doi: 10.1111/j.2517-6161.1996.tb02080.x.

[15] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. D. Freitas, "Taking the human out of the loop: A review of Bayesian optimization," *Proc. IEEE*, vol. 104, no. 1, pp. 148–175, Jan. 2016. doi: 10.1109/JPROC.2015.2494218.

[16] K. Nakamura, D. Hamada, Y. Ueda, K. Eriguchi, and K. Ono, "Selective etching of high-k dielectric HfO_2 films over Si in BCl_3-containing plasmas without rf biasing," *Appl. Phys. Exp.*, vol. 2, no. 1, 2009, Art. no. 016503. doi: 10.1143/APEX.2.016503.

[17] B. Ratner, "The correlation coefficient: Its values range between $+1/-1$, or do they?" *J. Targeting Meas. Anal. Market.*, vol. 17, no. 2, pp. 139–142, 2009. doi: 10.1057/jt.2009.5.