Plasma Science and Technology Division
Room A106 - Session PS1+MS-ThM

AI/ML in Plasma Applications

Moderators: Robert Bruce, IBM Research, T. J. Watson Research Center, **Yu-Hao Tsai**, TEL Technology Center, America, LLC

8:00am **PS1+MS-ThM-1 Approaches to Accelerate Etch Process Optimization by Using Virtual Experiment**, ***Tetsuya Nishizuka**, R. Igosawa, T. Yokoyama, K. Sako, H. Moki, M. Honda,* Tokyo Electron Miyagi, Ltd., Japan **INVITED**

High Aspect Ratio Contact (HARC) hole etching is one of the processes which require a lot of efforts to optimize etch condition. As the aspect ratio increases, some issues such as "distortion" and "twisting" which are hole circularity degradation and deviation from vertical etch respectively, have been critical. Since they cause asymmetric profile along hole axis, not only vertical but also horizontal cross section is necessary to observe 3D profile image, it takes more time to conduct a series of experiments, and then it makes the optimization more difficult.

In this study, we created a model for a topography simulation which is based on Monte Carlo method, so that we can conduct "virtual experiment" on the simulator and expect to reduce the number of experiments by understanding etch mechanism. With respect to practicability for model building, we employed a procedure that representative ion and radical parameters which associate with etching behavior are carefully fitted to actual experimental results [1].

As the result, while this kind of asymmetric distortion profile is supposed to come from stochastic variation and charging in the hole [2], we found there is another systematic factor that is an interaction between re-deposition of sputtered etch material and initial mask profile by analyzing Amorphas Carbon Layer (ACL) etching precisely [3]. This model is consisted with the fact that the distortion deteriorates under low temperature condition. It was also applied to oxide-nitride (ON) layers etch and well reproduced twisting profile on the simulator.

Additionally, we attempted an automatic parameter fitting by using ML optimization for the purpose to minimize efforts in case of converting the model to the other applications than HARC.

[1] Ohmine et. al., Jpn. J. Appl. Phys. 50 (2011)

[2] Huang et. al., J. Vac. Sci. Technol. A 37 (2019)

[3] Igosawa et. al., Proceedings of international symposium on dry process 2022

8:40am **PS1+MS-ThM-3 Recipe Optimization for Plasma Etching with Machine Learning Model Trained by Initial Dataset Using D-Optimal Design**, ***Ryo Morisaki**, T. Ohmori,* Hitachi, Ltd., Japan

The development of semiconductor fabrication processes is becoming more difficult due to a growing need for the miniaturization of semiconductor devices to the nano-scale level. Furthermore, growing demands for cutting-edge semiconductor devices of superior performance necessitate the swift development of the fabrication processes.

Plasma etching is a pivotal technique for semiconductor processes. Machine learning (ML) methods have been applied to optimize the recipe for these processes, which is a control parameter set including items such as plasma generation power, wafer bias power, gas species for plasma generation, and the flow rate of gases[1][2]. Datasets for training the ML model consist of recipes and their corresponding etching profiles. Generally, the recipes are curated by expert process engineers to reduce the cost of the etching experiments. On the other hand, design of experiments (DoE) methods can be utilized to obtain the training datasets without expert knowledge. Therefore, DoE has the potential to increase of the number of engineers who can optimize recipes for difficult etching processes.

In this work, two distinct approaches for creating an initial dataset for the training are compared to evaluate the efficiency of recipe optimization using an ML model. In the first approach, the initial dataset is created on the basis of the plasma etching knowledge of expert engineers, as has been conventionally practiced, and in the second approach, it is created on the basis of elementary knowledge for etching tool operations and a DoE method with the D-optimal criterion is used[3]. In the latter approach, a preliminary range of values for the recipe parameters, in which plasma generation and etching can occur, is established on the basis of fundamental knowledge of plasma etching. Subsequently, D-optimization is conducted on the recipe parameters within the specified range to generate a high-quality and diverse initial dataset that can improve the ML model for optimizing the recipes and profiles. In contrast to the conventional DoE with orthogonal array, this DoE with the D-optimal criterion method has no limitation on the number of experiments, thus making it suitable for creating small initial datasets to reduce the cost of the etching experiments. We report detailed comparison results of the efficiency of the etching optimization using each approach.

[1]T. Ohmori et al., in Proc. Int. Symp. Dry Process, pp. 9–10 (2017).

[2]H. Nakada et al., in Proc. Int. Symp. Dry Process, pp. 53–54 (2019).

[3]J. Keifer, Journal of the Royal Statistical Society. Series B (Methodological), vol. 21, no. 2, pp. 272–319 (1959).

9:00am **PS1+MS-ThM-4 Digital Twin Model to Compensate for Variations in Plasma Etching Process**, *T. Nakayama, T. Ohmori,* Hitachi, Ltd., Japan; ***Naoto Takano**,* Hitachi High-tech America, Inc.

The miniaturization of semiconductor devices based on Moore's Law has necessitated increasingly demanding precision in the mass production of devices. To achieve a target etching profile with nano-scale accuracy during the manufacturing process, plasma etching systems must be equipped with technologies to minimize variations of the etching. A set of parameters for the control function of the etching system is called a recipe, which is used as input data for the system. While the same recipe is used for all systems in mass production, the etching profiles are varied due to the drift of etching chamber conditions or differences in the conditions of the inner chamber parts replaced at the time of chamber maintenance. Etching rate (ER) data are often utilized to check the variations of chamber conditions and construct a compensation model, but a large ER dataset is required for the model, which is time consuming.

Therefore, we have been investigating a compensation method that requires only a small amount of data. In our method, first, a reference digital twin (DT) model utilizing neural networks is trained by sufficient data. A large amount of the training data consisting of recipes and ER can be prepared in advance by experiments using a reference chamber. Further, in addition to the experimental ER, simulation data of ion and radical fluxes correlating with ER (calculated by a plasma simulator) are used for the training data [1]. The simulation data can also be prepared in advance. Next, a small amount of ER data is obtained from a target chamber that has a different ER from that in the reference chamber. A target DT model is obtained by additionally training the reference DT using the small amount of ER data. Finally, a recipe that compensates for the ER difference is predicted using the target DT.

We prepared several hundred experimental ER data and calculated fluxes data using the reference chamber and the plasma simulator, respectively. Several tens of ER data were obtained as a small amount of data using a target chamber. Predicted recipes by the trained target DT model for ER compensation were experimentally verified in the target chamber. As a result, the ER difference used to check the variation of chamber conditions was decreased by our compensation method using DT models.

[1] T. Nakayama et al., Proc. Gaseous Electronics Conf., Sendai (2022).

9:20am **PS1+MS-ThM-5 Deep Learning-Enabled Plasma Equipment Design Optimization in Semiconductor Manufacturing**, *S. Ahn, **Jinkyu Bae**, S. Yoo, S. Nam,* Samsung Electronics, Republic of Korea

In plasma reactors, a focus ring which surrounds the wafer plays an important role in improving uniform fluxes and energy across the wafer. To ensure the uniformity and consistent processes, sophisticated design for the focus ring is necessary. However, focus ring design is very challenging due to the complexity of the design space and the high-dimensional geometry of the focus ring. Furthermore, the multi-scale and multi-physics nature of low-temperature plasmas (LTPs) makes it difficult to develop accurate simulation models that can capture the dynamics of plasma discharges. Simulating LTPs is the need to consider multiple physical and chemical processes that occur simultaneously, such as ionization, recombination, excitation, and attachment. These processes can be highly nonlinear and require a large range of integrating time scales from picosecond to millisecond. In this study, we present a Deep Neural Network framework employing the DeepONet, which is pre-trained deep neural operators between each physical quantities on behalf of physical governing equations. The training data is generated using HPEM (The Hybrid Plasma Equipment Model) plasma simulation solver. The framework involves two network types. The first network reduces the dimensions of the focus ring geometries to a latent representation. We used a geometric attention mechanism in Variational-Auto-Encoder (VAE) allowing us to discover the latent geometric features of focus ring parts. The high-dimensional design space was effectively reduced by the neural network model. We proposed the concept of using this latent representation in combination with the pre-trained neural networks. We pre-trained deep neural operators that can predict independently physical quantity fields, given general inputs. It is an efficient way of incorporating the plasma physics without embedding the partial differential equations into the loss function of the neural network.The proposed framework is shown to be efficient and effective in optimizing the focus ring design for different objectives, and the effects of variations in the design are thoroughly investigated based on very few measurements using pre-trained deep neural operators. This paper aims to develop framework for predicting plasma dynamics and carrying out focus ring design in reactors.

9:40am **PS1+MS-ThM-6 Wafer Arcing Detect Algorithm Using LSTM Autoencoder in Hardmask Strip Equipment with CCP Source***, **Heewoong Shin**,* PSK, Republic of Korea

“Wafer arc” is one of the phenomenon that occur in semiconductor manufacturing equipment that utilizes plasma, rather than being limited to arc discharge in plasma science. In this study, we discuss the possibility of successfully classifying normal manufacturing process of semiconductor equipment using CCP and the abnormal data by using deep learning methodology. In general, since wafer arcs data have an obvious characteristics that engineers can easily notice, it is thought that they can also easily be detected by general SPC methodology. But in some cases, there are pattern anomalies that cannot be detected by SPC, or there are problems such as requiring the subjectivity of engineers for determining the data is normal or not because it is hard to find common in each abnormal data. Because wafer arcs might cause serious malfunctions in most equipment and are a major cause of yield reduction, so a function that can diagnose abnormalities in advance is required. However, wafer arcs are generally known to be difficult to reproduce to make abnormal data, and there are problems such as requiring high-resolution optical equipment to detect the arc phenomenon which also costs additional charge. In this study, we used specific parameters in lotlog of CCP equipment as an input of DL model. By training with LSTM-autoencoder, it shows the possibility of classifying normal and abnormal data successfully through simple learning. When the result of this research can be applied in mass-production, it is highly expected that it will effectively detect and predict wafer arcs and other anomaly related to electrical/plasma parameters, and will greatly benefit the yield of semiconductor equipment.

Author Index

Bold page numbers indicate presenter

— A —
Ahn, S.: PS1+MS-ThM-5, 2
— B —
Bae, J.: PS1+MS-ThM-5, **2**
— H —
Honda, M.: PS1+MS-ThM-1, 1
— I —
Igosawa, R.: PS1+MS-ThM-1, 1
— M —
Moki, H.: PS1+MS-ThM-1, 1
Morisaki, R.: PS1+MS-ThM-3, **1**
— N —
Nakayama, T.: PS1+MS-ThM-4, 1
Nam, S.: PS1+MS-ThM-5, 2
Nishizuka, T.: PS1+MS-ThM-1, **1**
— O —
Ohmori, T.: PS1+MS-ThM-3, 1; PS1+MS-ThM-4, 1
— S —
Sako, K.: PS1+MS-ThM-1, 1
Shin, H.: PS1+MS-ThM-6, **2**
— T —
Takano, N.: PS1+MS-ThM-4, **1**
— Y —
Yokoyama, T.: PS1+MS-ThM-1, 1
Yoo, S.: PS1+MS-ThM-5, 2