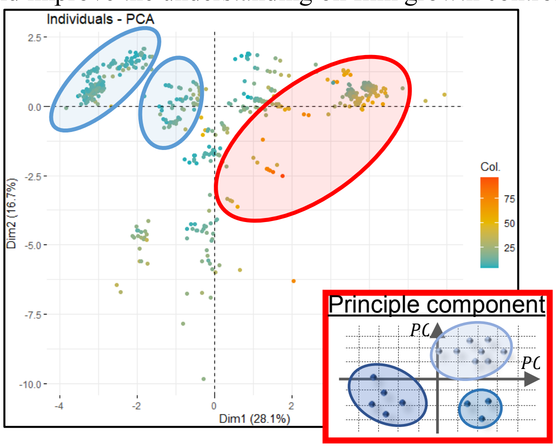
Machine learning-based analysis of the physio-chemical properties for the predictive thickness control of atomic layer deposition

Changsu Kim, Thai Ngan Do, and Jiyong Kim*

School of Chemical Engineering, Sungkyungkwan University, Suwon 16419, Republic of Korea
*(*E-mail address : jiyongkim@skku.edu).*

Abstract: Atomic layer deposition (ALD) is an outstanding thin film deposition technique based on the surface chemical reaction. As the conventional device fabrication method is reported to be ineffective under the 5nm process, ALD drew attention to its ability to control the film growth. To develop an efficient ALD process, controlling film thickness is one of the most important factors. The key requirement for the film thickness control is to understand their chemical reactivity upon adsorption on different substrate surfaces. However, the current research on viable ALD development have remained inefficient because of the expensive and time-consuming experiments. In this study, we aim to analyze and suggest a strategy for identifying the contribution of ALD process features on the device film thickness based on principal component analysis (PCA) method. First, the features of ALD experiment (Precursor properties, reactant type, substrate properties, operating conditions) are compiled to define the chemical dimension of ALD process. Then, the contributions of ALD experiment features on the thickness growth are analyzed by projecting ALD chemical dimension into the information space using PCA technique. As a result, we could identify highly sensitive features for ALD film growth and suggest the strategy for controlling film thickness. From providing the analysis for the film growth, we elucidate the importance of high dimension analysis of ALD process and improve the understanding on film growth control.



Keywords: Atomic layer processing; Machine learning; Artificial intelligence

1. INTRODUCTION

As the demand for greater integrated circuit functionality and performance increased, more precise and controllable deposition technique has been highly desired (G. Mazaleyrat et al., 2005). The deposition processes conducted in device fabrication conventionally are chemical vapor deposition (CVD) and physical vapor deposition (PVD). Even though the conventional deposition process has been helpful for device fabrication, there are many difficulties and limitations to produce the device with high circuit density at a lower cost. For example, decreasing the size of semiconductor devices is limited to 5nm with the conventional method (Rizwan Khan et al., 2018; Hyun Gu Kim et al., 2020). In this context, atomic layer deposition (ALD) is one of the most salient alternatives of deposition process for the thin film development (Chi Thang Nguyen et al., 2021). As ALD is able to control the film growth by modifying the surface properties, implementing ALD could achieve more uniform film with higher circuit density (Hyunhang Park et al., 2016).

However, even though ALD is the effective and powerful method for developing more precise devices, time-consuming and expensive experiments have been great challenges of ALD deployment (Hyun Gu Kim et al., 2020; Markku Leskelä Prof. et al., 2003). Especially, the film thickness is extremely difficult to be predicted according to the material types (e.g., precursor, reactant, substrate) and operating conditions (e.g., temperature, cycle, idle time) (Yangyao Ding et al., 2019). Hence, there is an urgent need for a deeper understanding of film growth to obtain the predictive film thickness control.

Machine learning (ML) technique is artificial intelligence (AI) method for analyzing and predicting the complex system of natural phenomena (C. Pozo et al., 2012). Based on the observation data, ML could provide the practical data analysis method such as dimension reduction. Principal component analysis (PCA) technique is one of the well-established dimension reduction techniques (Hervé Abdi et al., 2010). From transforming the descriptor space to information space, the perplexing relationship of the descriptors could be more easily identified.

Therefore, in this research, we are aimed to analyze the dominant descriptors of ALD process for film growth and suggest the strategy for controlling the film thickness. For the analysis, the chemical dimension of ALD process is defined by collecting the critical ALD descriptors and experimental data. The ALD descriptors include the physio-chemical properties of materials (precursor, reactant, substrate) and operating conditions. Then, by using PCA technique, the contribution of descriptors for achieving thick (>30nm) and thin (≤30nm) thickness is compared and analyzed. Consequently, it was able to identify the material conditions and operating conditions for controlling film thickness.

2. METHOD

2.1 ALD experiment datasets

The precursor, reactant, and substrate types are the major components of ALD material conditions. Especially, the precursor design has a fatal role in ALD R&Ds. In ALD experiment datasets, there are popular 12 precursors (TTIP, Tris-dimethylamionisilane, TMA, $TiCl_4$, TEMAZ, TDMAH, $MeCpPtMe_3$, DMAI, DMADMS, $CoCp_2$, $CoCp(CO)_2$, Carish) as shown in Figure X. The material conditions and operating conditions are summarized in Table 1.

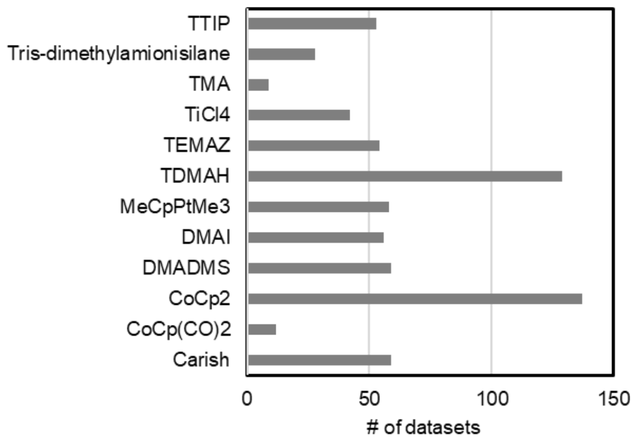


Figure 1. The precursor type counts

The precursors are then represented into the physio-chemical descriptors (ligand material, precursor molecular weight, Z value, ligand length, the number of oxides, precursor expose temperature, precursor expose time). The ligand materials means substances forming ligand, which are chlorine, carbon, cyclopentadiene, oxygen, and nitrogen. Total 533 ALD experimental datasets are compiled.

For the experiments, there are 4 reactants, which are H_2O, NH_3, O_2, and O_3. The reactant conditions incluye the number of H and O, molecular length, molecular weight, and expose time. These reactants are used to modify the substrate'surface properties, which later has the terminations of $–OH$, $-NH_2$, and -O.

In case of substrate, etched Si and SiO_2 are used in the experiment. Unlike other descriptors, substrate descriptor is not numerically represented. Because only 2 substrates are dealt thereby representing substrate as a categorical value (1 or 2). Note that the substrate temperature is sorted as the operating condition rather than substrate material conditions due to the process.

Operating conditions include 3 descriptors, which are idle time, substrate temperature, and cycle. Note that the cylce means the pulse-purge cycle of precursor and co-reactant. Lastly, the thickness of ALD process is also included in the datasets. The total counts of experimental datasets are 696 with 21 descriptors. Among 21 descriptors, 20 descriptors are dependant variables and 1 descriptor (thickness) is indenpendant variable.

Table 1. ALD experimental dataset

Feature name	Avg.	Std.	Min.	Max.
Cl (Chlorine)	0.24	0.95	0	4
C (Carbon)	0.82	1.12	0	3
Cp (Cyclopentadiene)	0.49	0.80	0	2
O (Oxygen)	0.72	1.45	0	4
N (Nitrogen)	1.34	1.76	0	4
Precursor M.W. (g/mol)	253	97.6	92.5	411.3
Z value	38.4	23.2	13	78
Ligand length (nm)	7.01	2.37	3.69	12.44
Number of oxide	1.24	0.94	0	2
Temperature (℃)	57.7	23.6	0	100
Expose time (hr)	3.42	2.15	0.5	15
# of H	1.35	1.23	0	3
# of O	1.58	1.18	0	3
Molecular length (μm)	1.72	0.21	1.23	1.89
Reactant M.W. (g/mol)	29.7	14.3	17	48
Expose time (hr)	3.88	4.73	0.3	60
Functional group	-	-	1	2
Idle time (hr)	0.97	3.95	0	60
Wafer temperature (℃)	262	77	100	428
Cycle	245	185	30	1500
Thickness (nm)	18.8	13.1	3.75	96.4

2.2 Principal component analysis (PCA)

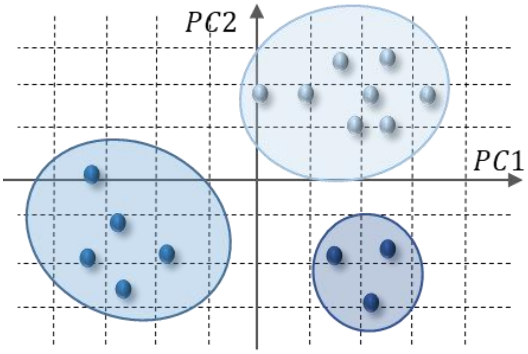


Figure 2. Principal component analysis

PCA is a powerful technique to visualize and interpret the high-dimensional data by extracting the most valuable chemical descriptors and transforming the chemical space into information space. Principal components (PC) are the dimension of information space and decided by Eigen-decomposition procedure. From finding the most critical orthogonal vectors maximizing the covariance of datasets, the ALD data could be projected in lower dimensions.

The mathematical concept of PCA includes the calculation of the covariance and covariance matrix. Firstly, the Euclidean space that has N number of ALD descriptors is constructed. For the given series of datasets x and y, the variances of each dataset are calculated as follows:

$$\mathrm{cov}(x, y) = E\big[(x - m_x)(y - m_y)\big] = E[xy] - m_x m_y \qquad (1)$$

m_x and m_y indicate average value of x and y.

The covariance matrix C is as followed:

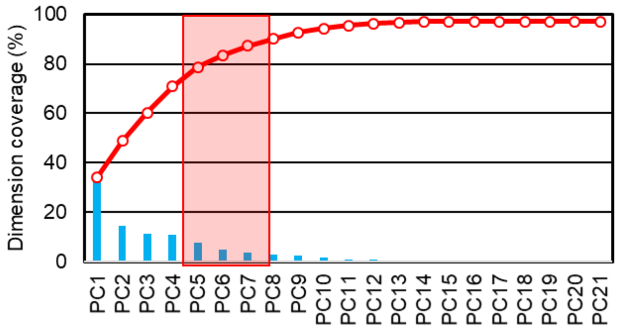
$$C = \begin{pmatrix} cov(x,x) & cov(y,x) \\ cov(x,y) & cov(y,y) \end{pmatrix} = \begin{pmatrix} \frac{1}{n}\sum(x_i - m_x)^2 & \frac{1}{n}\sum(x_i - m_x)(y_i - m_y) \\ \frac{1}{n}\sum(x_i - m_x)(y_i - m_y) & \frac{1}{n}\sum(y_i - m_y)^2 \end{pmatrix} \qquad (2)$$

Then, Eigen-decomposition using orthogonal matrix (P) and diagonal matrix (P^{T}) produce the Eigenvector (e) and Eigenvalue (λ) of information space.

$$C = P\Sigma P^T = (e_1 \quad \cdots \quad e_n)\begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_2 \end{pmatrix}\begin{pmatrix} e_1^T \\ \vdots \\ e_n^T \end{pmatrix} \qquad (3)$$

The Eigenvector and Eigenvalue are score and loading, respectively. Score is the coordinates in information space and loading is the contribution of the descriptor on Eigenvector. Among N number of Eigenvectors (PCs) produced, PCs that explain the 80-90% of original datasets are selected. Also, in order to identify the importance of descriptors, the loadings of descriptors are compared and analyzed.

3. RESULTS.



PC1	PC2	PC3	PC4	PC5	PC6	PC7
34.2%	14.6%	11.4%	10.8%	7.8%	4.8%	3.8%

Figure 3. Dimension coverage (blue) and the cumulative dimension coverage (red) of PCs

Figure 3 shows the dimension coverage and the cumulative variance of PCs. To keep the original features of the data, it is essential to choose the number of PCs. As shown in Figure 3, the PCs that explain 80-90% of datasets are at least PC1 to PC5. It means that we have to consider at least PC1 to PC5 to extract the essential information of ALD process datasets.

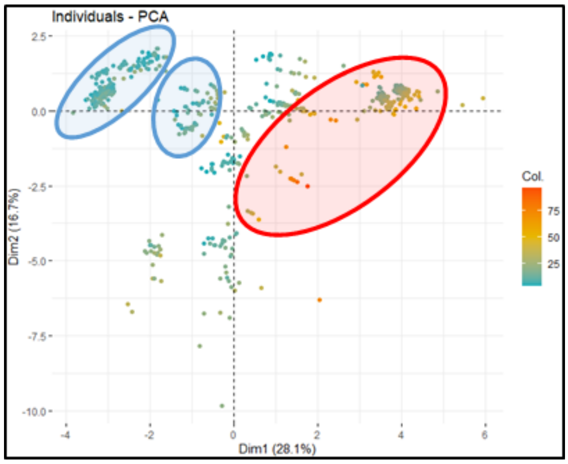
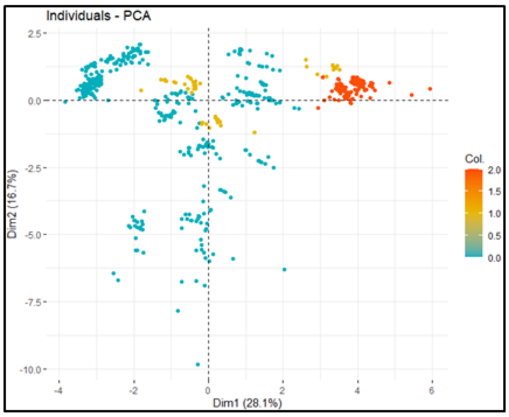


Figure 4. Thick (red) and Thin (blue) thickness of ALD process

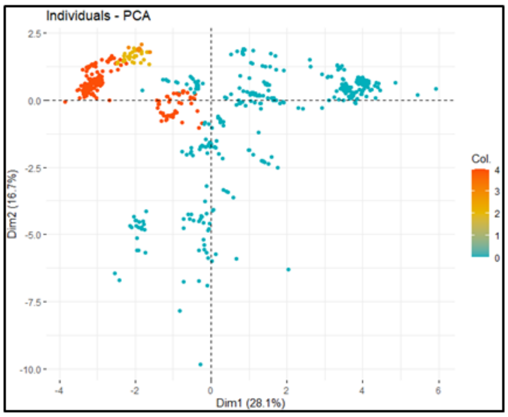
In order to visualize the relationship of ALD process feature and film thickness, the PCs with the biggest dimension coverage are extracted. The ALD process data is projected in dim1 (PC1) and dim2 (PC2) coordinates. PC1 and PC2 are covering 48.8% of original ALD data. The thickness and ALD process features are displayed in Figure 4 and Figure 5, respectively.

Among 20 dependent variables, Figure 5 highlights the descriptor with a noticeable trend on film growth. The noticeable descriptors are Cp, precursor molecular weight, ligand length, N, precursor Z value, the number of oxide in precursor, the number of H in the reactant, the number of O in the reactant, reactant molecular weight, and the number of experiment cycle.

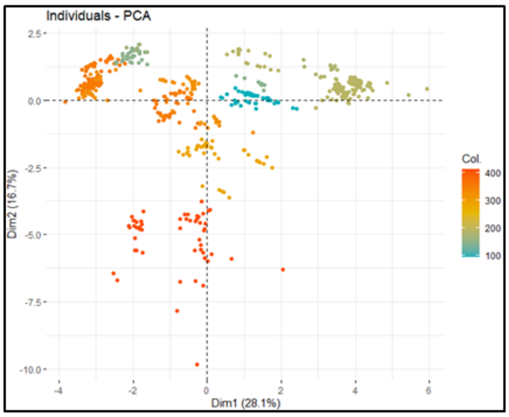
a) C_p (Ligand component)



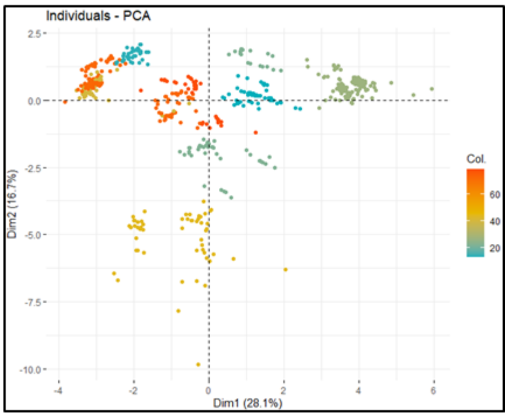
b) N (Ligand component)



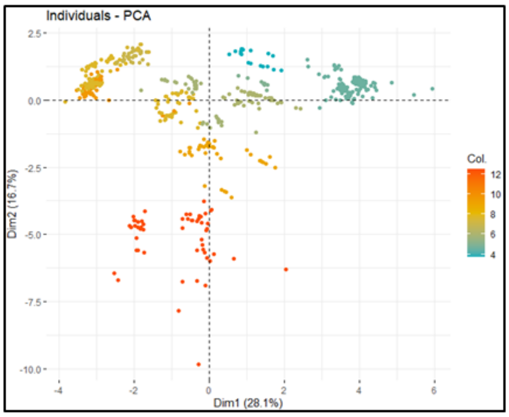
c) Precursor molecular weight



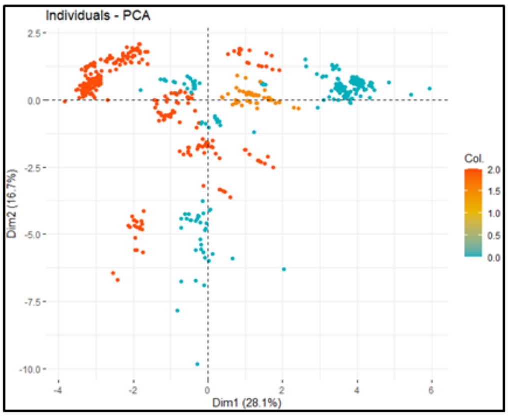
d) Precursor Z value



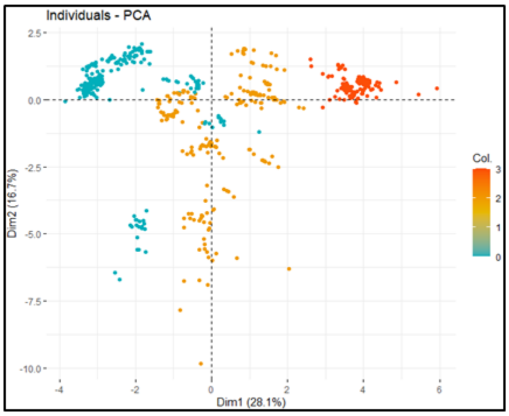
e) Ligand length



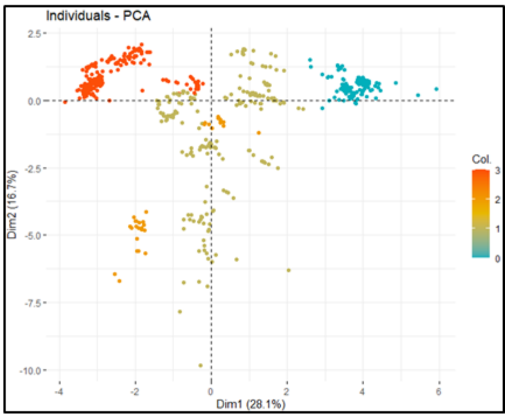
f) The number of oxide in the precursor



g) The number of H in the reactant



h) The number of O in the reactant



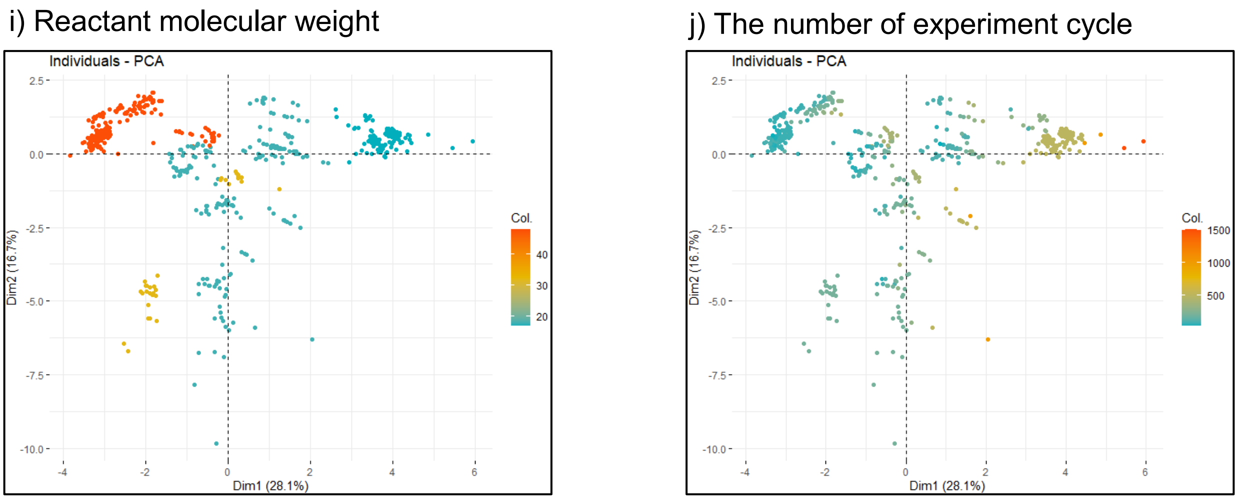


Figure 5. The distribution of features (a. Cp, b. N, c. Precursor molecular weight, d. Precursor Z value, e. Ligand length, f. The number of oxide in precursor, g. The number of H in the reactant, h. The number of O in the reactant, i. The reactants molecular weight, k. The number of cycle) in information space

As shown in Figure 4, thin film and thick film data are located in the upper left area and upper right area of the plot. In case of both film growth, the distinguishable trend are found in precursor conditions, reactant conditions, and operating conditions.

For the thin film growth, the precursor tends to have relatively shorter ligand length and more N than other precursors. Also, according to Figure 5 c and d, the presence of heavy precursor does not have a clear contribution on the film thickness. In the reactant, high number of O and high molecular weight are clearly favored while the presence of H is not favored.

Oppositely, for the thick film growth, the precursor tends to have a longer ligand length and 5re N 5re other precursors. While precursor molecular weight and Z value has no clear 5ren on the thin film growth, in thick film growth, Figure 5 c and d show that the modest precursor molecular weight and Z value are favored. In the reactant, a low number of O and low molecular weight are clearly favored.

In the material aspect, the favored conditions for the thin and thick film growth tend to have the opposite trend. The operating condition for the thin and thick film growth was opposite as well. As there are more cycle, the film tends to have thicker film.

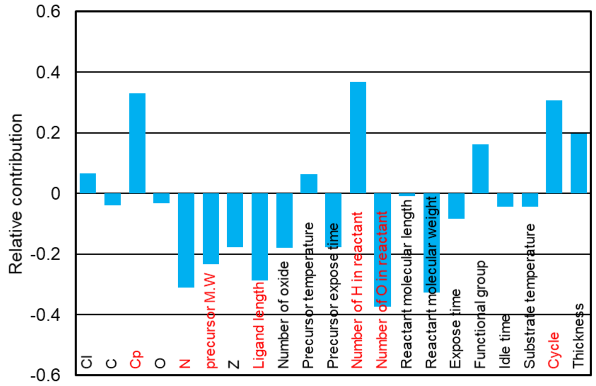


Figure 6. The relative contribution of ALD descriptors on PC1

As well as in Figure 5, the contributions of 21 descriptors on PC1 construction are analyzed in Figure 6. Note that the absolute value of the relative contribution is taken into account for the interpretation. The noticeable descriptors are highlighted in red. For PC1, there are 7 descripors (Cp, N, Precursor moleculear weight, ligand length, number of H in reactant, number of O in reactant, cycle) that have relatively more important than others. Especially, the importance of reactant conditions are the most important among them.

Meanwhile, there are less important descriptors: Cl, C, O, precursor temperature, reactant molecular length, idle time, and substrate temperature. Interestingly, the contribution of Cp and N are more visible than that of other ligand materials. Also, all kind of temperatures (the precursor temperature and substrate temperature) shows the relatively less contribution than other descriptors. It means that film thickness of ALD process does not heavily rely on the temperature.

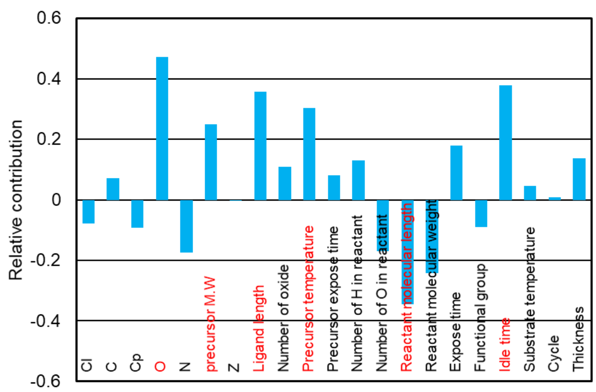


Figure 7. The relative contribution of ALD descriptors on PC2

The contributions of 21 descriptors on PC2 construction are analyzed in Figure 7. The most distinguishable difference from PC1 contribution in Figure 6 is that some of the high contributors (O, precursor temperature, reactant molecular length, idle time) are the less important descriptors in PC1. Interestingly, the reactant molecular length and idle time that were not important in the PC1 are heavily contributed in PC2.

Meanwhile, in PC2, some of descriptor (Cp, number of H in reacatnat, number of O in reactant, cycle) that are important in PC1 are less important in PC2.

Even though PC1 and PC2 agree that the precursor weight and ligand length are important descriptors, the result from Figure 5 shows that the precursor weight shows no clear contribution to the thickness. Hence, it is suggested that designing appropriate ligand size is important to control the film thickness of ALD process. Specifically, ligand size around 8 nm is suitable for thin film growth, while ligand sizes less than 8nm are for thick film growth.

4. CONCLUSIONS

The critical descriptors and trend of film growth of ALD process are analyzed by using PCA to understand the relationship of ALD system. In this research, the ALD process descriptors, including physio-chemical properties of materials and operating conditions, are collected to develop the preliminary chemical dimension of ALD process. As a result, high-dimensional analysis using PCA reveals the most important descriptor for the film thickness control and suggests the strategy for thick (>30nm) and thin (≤30nm) film growth. Based on this study, we could provide fundamental knowledge for film thickness control. As future work, we aimed to analyze the ALD process with a larger volume of data and develop the strategy for precisely predicting the thickness of ALD process.

REFERENCES

C. Pozo, R. Ruíz-Femenia, J. Caballero, G. Guillén-Gosálbeza, L. Jiméneza (2012), "On the use of Principal Component Analysis for reducing the number of environmental objectives in multi-objective optimization: Application to the design of chemical supply chains", *Chemcial Engineering Science*, 146-158, vol. 69, issue 1

Chi Thang Nguyen, Bonwook Gu, Taehoon Cheon, Jeongwoo Park, Mohammad Rizwan Khan, Soo-Hyun Kim, Bonggeun Shong, and Han-Bo-Ram Lee* (2021), "Atomic Layer Modulation of Multicomponent Thin Films through Combination of Experimental and Theoretical Approaches", *Chemistry of Materials*, 4435-4444, vol. 33, issue 12

G. Mazaleyrat, A. Estève, L. Jeloaica, M. Djafari-Rouhani (2005), "A methodology for the kinetic Monte Carlo simulation of alumina atomic layer deposition onto silicon", *Computational Material Science*, 74-82, vol. 33, issue 1-3

Hervé Abdi, Lynne J. Williams (2010), "Principal component analysis", WIREs computational statistics, 433-459, vol. 2

Hyun Gu Kim, Miso Kim, Bonwook Gu, Mohammad Rizwan Khan, Byeong Guk Ko, Sumaira Yasmeen, Chang Su Kim, Se-Hun Kwon, Jiyong Kim, Junhyuck Kwon, Kwangseon Jin, J. -S. Chun, Byungchul Cho, Bonggeun Shong, and Han-Bo-Ram Lee (2020), "Effects of Al Precursors on Deposition Selectivity of Atomic Layer Deposition of Al2O3 Using Ethanethiol Inhibitor", *Chemistry of Materials*, 8921-8929, vol. 32, issue 20

Hyunhang Park, Bogyeong Kim, Sung Hoon Lee, and Hyunbin Kim (2016), "Study of a Vanadium Precursor for VO2 Thin-Film Growth in the Atomic Layer Deposition Process by Multiscale Simulations", *The Jouranl of Physical Chemistry C*, 28193-28203, vol. 120, issue 49

Markku Leskelä Prof., Mikko Ritala Prof. (2003), "Atomic Layer Deposition Chemistry: Recent Developments and Future Challenges", *Angewandte Chemie international edition*, 5548-5554, vol. 42, issue 45

Rizwan Khan, Bonggeun Shong*, Byeong Guk Ko, Jae Kwang Lee, Hyunsoo Lee, Jeong Young Park, Il-Kwon Oh, Shimeles Shumi Raya, Hyun Min Hong, Kwun-Bum Chung , Erik J. Luber, Yoon-Seok Kim, Chul-Ho Lee, Woo-Hee Kim*, and Han-Bo-Ram Lee* (2018), "Area-Selective Atomic Layer Deposition Using Si Precursors as Inhibitors", Chemistry of Materials, vol. 30, issue 21

Yangyao Ding, Yichi Zhang, Yi Ming Ren, Gerassimos Orkoulas, Panagiotis D. Christofidesa (2019), "Machine learning-based modeling and operation for ALD of SiO 2 thin-films using data from a multiscale CFD simulation", *Chemical Engineering Research and Design*, 131-145, vol. 151