PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Machine learning for predictive electrical performance using OCD

Sayantan Das, Joey Hung, Sandip Halder, Guillaume Schelcher, Roy Koret, et al.

Sayantan Das, Joey Hung, Sandip Halder, Guillaume Schelcher, Roy Koret, Igor Turovets, Mohamed Saib, Anne-Laure Charley, Matthew Sandelbach, Avron Ger, Philippe Leray, "Machine learning for predictive electrical performance using OCD," Proc. SPIE 10959, Metrology, Inspection, and Process Control for Microlithography XXXIII, 109590F (26 March 2019); doi: 10.1117/12.2515806

SPIE.

Event: SPIE Advanced Lithography, 2019, San Jose, California, United States

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 28 Mar 2019 Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Machine learning for predictive electrical performance using OCD

Sayantan Das[a], Joey Hung[b], Sandip Halder[a], Guillaume Schelcher[a], Roy Koret[b], Igor Turovets[b], Mohamed Saib[a],
Anne-Laure Charley[a], Matthew Sendelbach[c], Avron Ger[b], Philippe Leray[a]

[a]Imec, Kapeldreef 75, B-3001 Leuven, Belgium
[b] Nova Measuring Instruments, LTD, P.O. Box 266, Weizmann Science Park, Rehovot 76100, Israel
[c] Nova Measuring Instruments, Inc., 3090 Oakmead Village Drive, Santa Clara, CA 95051, USA

Abstract

With growing process complexity and the increasing number of process steps, early prediction of device performance has become an important task in semiconductor manufacturing process control. Machine learning (ML) techniques allow us to link in-line measurements to End-Of-Line (EOL) electrical tests. In our paper, we use reflectance spectra obtained from the scatterometry tool to predict both metal-line resistance and capacitance. We used IMEC N-14 process flow with LELE EUV double patterning at the M1 stage. Special designs-of-experiments (DOE) for multiple parameters allowed us to create a metrology solution for the entire process window and test its accuracy for all POI. Induced variations of both line CDs and space CDs, together with specially designed measurement sites, created wide variations both in metal-line resistance and capacitance. Reflectance spectra were collected in-line at two process steps defining metal lines: HM etch and Cu CMP at multiple targets, including E-test measurement sites, together with reference metrology for overlay (OV) and CD (by using Diffraction-Based-Overlay (DBO) and CD SEM). EOL MT1 electrical test results were used for the ML training procedure for early prediction of patterning effects (both CD and OL) on electrical performance enabling early decisions and cost reduction by discarding out-of-spec wafers before they reached the electrical test. It was shown that ML OCD predictive techniques are complimentary to the OCD model-based solutions for geometrical parameters widely used for in-line APC.

1. Introduction

The continuous shrinkage of device features has resulted in process costs at different steps (litho, etch, deposition, CMP) becoming more and more prohibitive unless extremely high yields are achieved. The advanced technology nodes require tight process control and accurate CD measurements. Metrology techniques like optical critical dimension scatterometry (OCD) and CDSEM are typically used for in-line CD measurements. However, both techniques have limitations and advantages [1]. For instance, discrepancy in material properties (n&k) and long model-optimization times restrict OCD scatterometry techniques, while resist shrinkage and its charging effect impact the measurement performance of the CDSEM tool. In this context, having an early insight into the electrical performance and variability can be a significant game changer for high-volume manufacturing (HVM), such that proper action can be taken in a timely manner either to scrap or rework the wafer, and improve or monitor the process.

Metrology, Inspection, and Process Control for Microlithography XXXIII, edited by
Vladimir A. Ukraintsev, Ofer Adan, Proc. of SPIE Vol. 10959, 109590F © 2019
SPIE · CCC code: 0277-786X/19/$18 · doi: 10.1117/12.2515806

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 28 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

As we progress to more advanced nodes, device structure and layout architecture get very complicated, sometimes requiring a greater number of process and metrology steps requiring greater process monitoring and control. This increase in the number of OCD steps requires a longer time to solution, especially in R & D. Moreover, tighter process control budgets are compromised by model errors. Modelling of non-periodic structures can make OCD modeling problematic. For these reasons, machine learning solutions will become an attractive tool for future process control and monitoring purposes.

Scatterometry involves light diffraction from periodic structures, and the intensity of the diffracted light is measured. The measured intensities are then compared with modeled data to extract parameters of interest. The modelling toolbox consists of a full geometrical model with Rigorous Coupled Wave Analysis (RCWA) [2]. Predicting data from OCD using machine learning is also now available. The advantage of scatterometry is that it is sensitive to the change in geometry and material properties. It is also a fast and non-destructive technique. However, it is an indirect measurement, and the parameters need to be obtained by computer modelling.

The methodology involves cross combining the OCD spectra with reference data. Then, a mathematical estimator is generated using a set of machine learning algorithms. The data used to create the mathematical estimator is the training set. Once the training is complete and shows a good correlation to the reference data, we can use the OCD spectra to predict different outcomes, such as CD, overlay and electrical performance for other wafers. This technique is less dependent on structural complexity. Model optimization is done by DOE. The technique works complimentary to OCD modelling.

2. Mask Layout and description of electrical structures

The minimum design CD and pitch is 24nm and 96nm, respectively. We apply a litho-etch-litho-etch approach whereby the first M1A layer is exposed. Afterwards, the pattern is transferred to an oxide HM, followed by exposing the second metal layer, to achieve 48nm pitch. There are 25 sub-dies within a field. Both are LF field masks. We then use the NTD process to print trenches.

3. Test vehicle description

3.1 Process flow

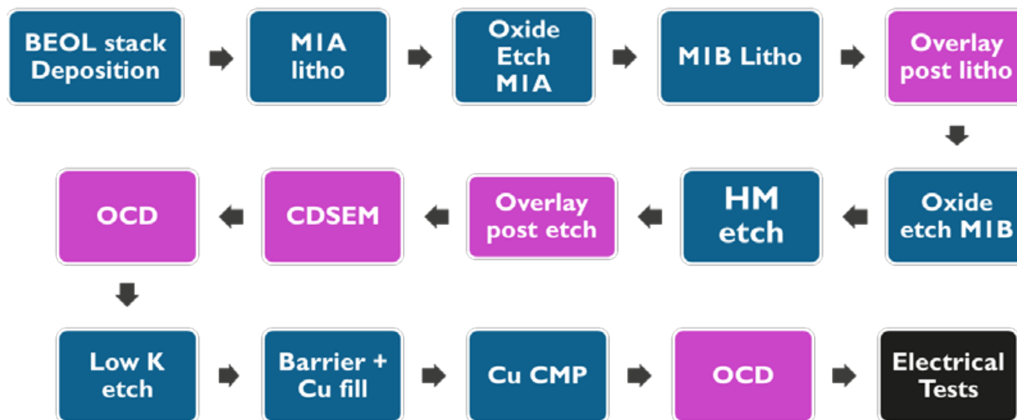


Fig. 1. Process flow showing LELE approach to achieve 48nm pitch

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 28 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Fig.1 displays the basic process steps involved in achieving 48nm metal pitch. The first step in the process includes BEOL stack deposition followed by M1A litho exposure under an ASML immersion scanner. The target CD post litho is approximately 40nm L/S. A negative tone development process was chosen to print trenches. A 30nm SOG on top of 100nm SOC was used to obtain the desired CD post etch. The oxide etch step involves opening the SOG/SOC step and transferring the pattern onto an Oxide layer. A second litho exposure involving the M1B layer is then performed, followed by pattern transfer onto the oxide layer. Both patterns are then transferred onto the TiN HM (hard mask) layer. At this stage, the CDSEM data, post-etch overlay data, as well as scatterometry/OCD data, are collected. Finally, the HM is used to pattern the low k trenches, which is filled with Cu, followed by a CMP step. A 5nm thin layer of SiCN was then deposited on top of the wafer to prevent oxidation of the Cu metal lines. The wafers are then sent for further scatterometry measurements on all resistance and capacitance targets before undergoing electrical tests.

3.2 Test structure description

Our study involved measuring CDSEM, collecting OCD spectra and measuring line resistance at 4 different locations. The site names ending with WR and NR signify that these targets either have designed roughness (WR) or are without designed roughness (NR). The length of the line is approximately 600 microns and CD =24nm. Each sub-die also has 12 (AB1, AB2, ..., AB6 and BA1, BA2......, BA6) vertically placed Fork-Fork structures to correlate with the Y overlay and 12 horizontally placed structures to correlate with the X overlay. The CD is 24nm and the period is 192 nm. The different designs have different spacing between them.

3.3 CD and Overlay fingerprint

To obtain a good training set for machine learning, we created a sub-recipe for the scanner job such that there is a translational offset of 0 to 7.5nm in the X and Y directions for the four columns, as shown in Fig. 2. These are referred to as programmed overlay wafers for the remainder of this paper.

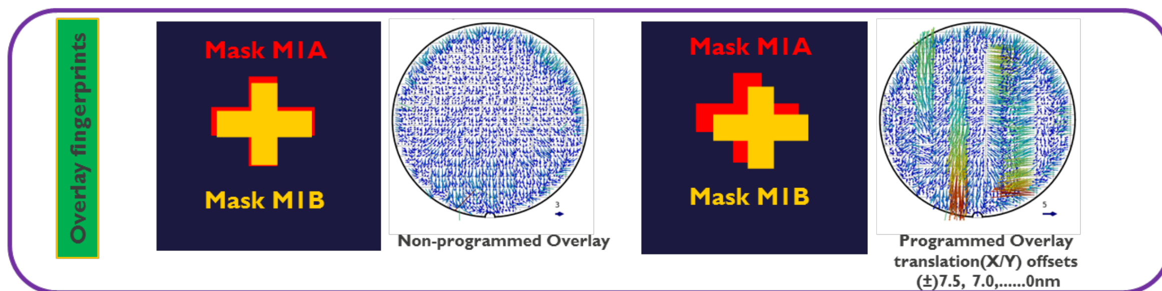


Fig. 2 Programmed and non-programmed (POR) overlay fingerprints post litho.

In order to induce CD change and obtain a good machine learning training set, etch DOE was performed by manipulating the oxide etch recipe at the M1B oxide etch step, keeping the etch process at M1A step in POR condition. Consequently, the first version (V1) of etch DOE increased the CD from the POR condition, and the second version (V2) of the etch DOE decreased the CD, as shown in Fig. 3.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 28 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

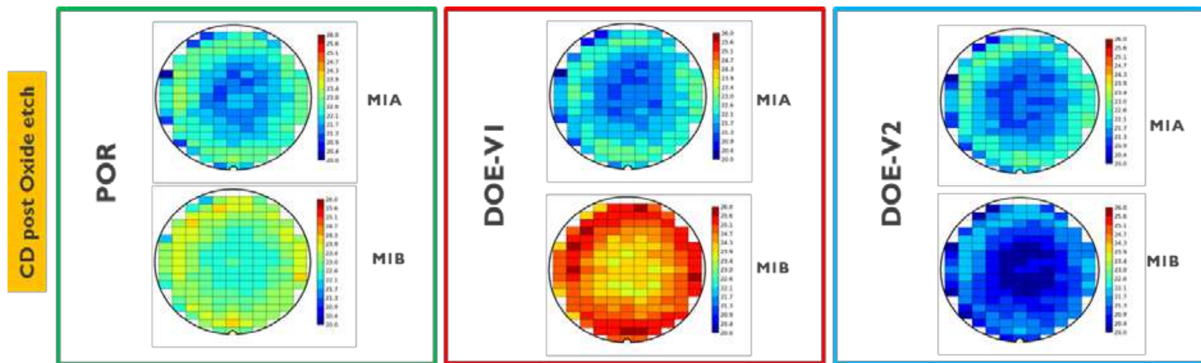


Fig. 3 CD fingerprints post Oxide etch for POR and DOE wafers.

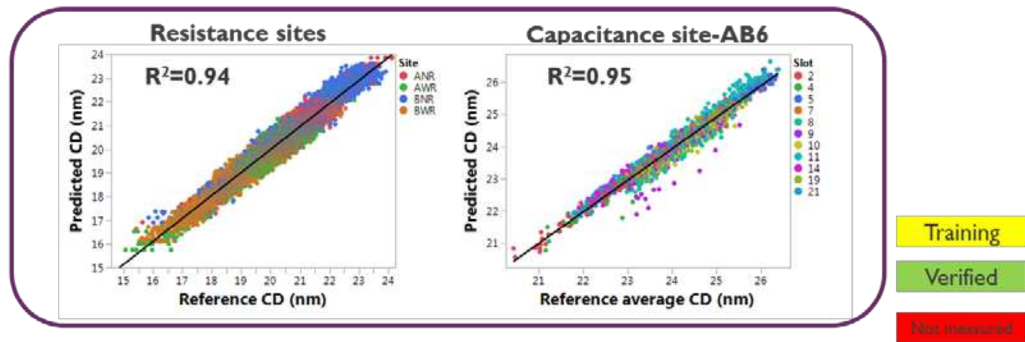
4. Results and discussion

4.1 *Modeless CD prediction from OCD spectra and its correlation to reference CD Post HM etch.*

Our goal is to measure and predict CD on electrically relevant structures. To use the machine learning functionality, we must set up wafers for a training set. Once the training is done, we can use the OCD spectra from the remaining wafers to obtain the predicted CD. For the capacitance site AB6, the average CD was calculated, based on the CDSEM measurements, and used as the reference CD. Machine learning was then used to predict the CD of the same site for other wafers. The M1A and M1B lines are practically identical, and the OCD signal, which collects the full target size (30 μ^2) in a single exposure, cannot discriminate between the two lines. For this reason, the average CD is chosen for machine learning training.

Full field OCD spectra collection and SEM measurement was carried out in this study. The SEM measurement was performed on four different resistance sites. One FEM wafer and two etch DOE wafers, as mentioned in the table, were used to perform the training for the resistance site. For the capacitance site, we chose one FEM wafer, one POR and one etch DOE wafer for the training set. The predicted CD using the machine learning solution shows excellent correlation to the reference CD, as shown in Fig. 4. The R^2 values obtained were 0.94 for the resistance site and 0.95 for the capacitance site.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 28 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use



Site	D02	D04	D05	D07	D08	D09	D10	D11	D14	D15	D16	D19	D21
Resistance	M1A-FEM	M1B-FEM	ETCH_V1	POR	POR	POG_OVL	POR	ETCH_V1	ETCH_V2	POR	POG_OVL	POR	ETCH_V2
Capacitance	M1A-FEM	M1B-FEM	ETCH_V1	POR	POR	POG_OVL	POR	ETCH_V1	ETCH_V2	POR	POG_OVL	POR	ETCH_V2

Fig. 4 Modeless CD prediction from OCD spectra and its correlation to reference CD post HM etch

4.3 Predicting overlay from OCD spectra using machine learning

Another application of the machine learning technique is for predicting overlay values using the OCD spectra collected. Two different techniques were chosen as reference overlay values: SEM-based [3] and diffraction-based overlay (DBO). For the SEM-based overlay calculation, we used the formula below to calculate the values obtained from the CDSEM data.

Overlay by SEM = CD of (M1A+M1B)/2 +Space – 48(nominal width)

The SEM measurement and OCD spectra collection were done at the same target, while diffraction-based overlay (DBO) were collected at a different location within the field (a designed DBO target). The training set included two wafers, one POR and one Programed overlay wafer. Since the measurement performed here was only done on vertical structures, we only show the Y-overlay values here. The same prediction can be done for X-overlay using the OCD spectra from horizontal structures. Fig. 5 illustrates that an excellent correlation exists between predicted overlay and SEM based overlay techniques and a moderate correlation when comparing predicted values to the diffraction-based technique. Different measurement sites could be one reason for the low R^2 value, and there was an intra-field contribution which resulted in a lower R^2 value. However, the correlation was found to be good for the programmed overlay wafers.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 28 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

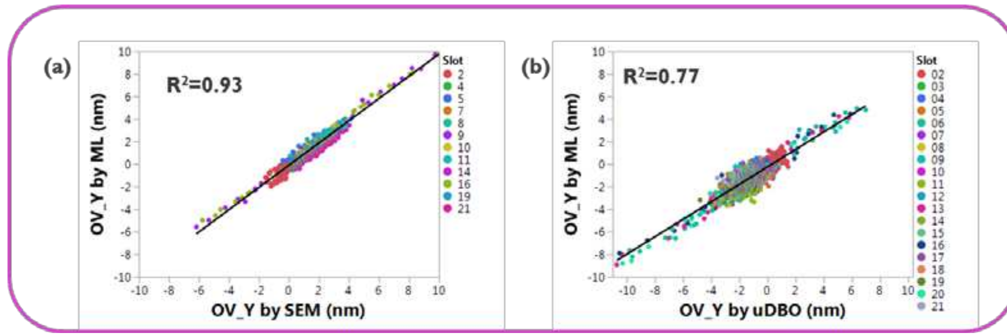


Fig 5. Correlation of overlay values predicted using machine learning technique vs (a) SEM-based overlay (b) diffraction-based overlay

4.4 Predicting line resistance from OCD spectra: OCD model vs Machine learning

The line resistance is plotted in the vertical axis for all the wafers and for three different targets. The wafers where DOE was performed can be identified with decreased and increased resistance values compared to the POR wafers. We also found that the resistance site with programed roughness (BWR compared to AWR) showed slightly higher resistance than sites without programed roughness. In the OCD model, the inverse of the Cu area is used to directly correlate with the line resistance. In Fig. 6, we have shown the result for three different sites: M1A with programmed roughness and M1B with and without roughness structures.

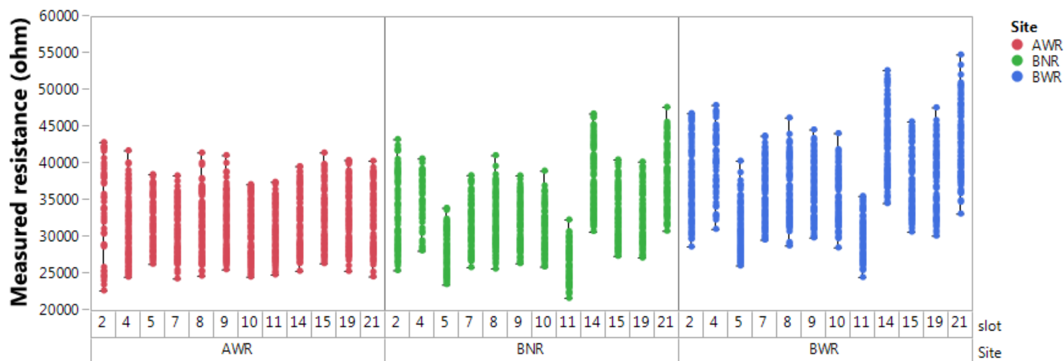
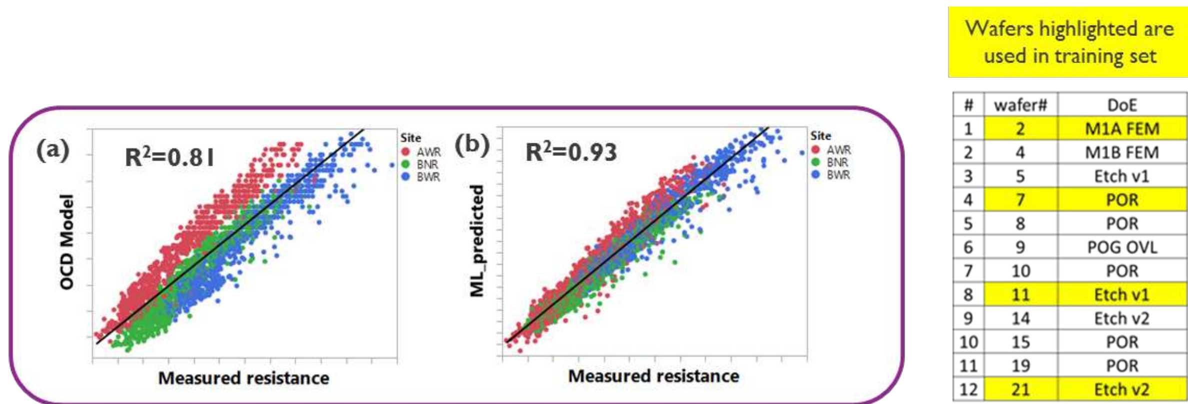


Fig 6. Measured line resistance of different wafers for different targets: AWR (M1A with programmed roughness), BNR (M1B without programmed roughness), BWR (M1B with programmed roughness)

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 28 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

The OCD model which was used to predict the line resistance of different structures shows a quite good correlation (0.81) to the measured line resistance values as illustrated in Fig. 7a. For the machine learning training set we used a FEM wafer, a POR wafer, one wafer with V1 etch recipe and another wafer from a V2 etch recipe. In Fig. 7b, it was demonstrated that by applying machine learning algorithms, the resistance value for all the wafers and different targets were predicted with an improved R^2 value when compared to the measured resistance. Moreover, all the different target types are present on the same linear fit.



#	wafer#	DoE
1	2	M1A FEM
2	4	M1B FEM
3	5	Etch v1
4	7	POR
5	8	POR
6	9	POG OVL
7	10	POR
8	11	Etch v1
9	14	Etch v2
10	15	POR
11	19	POR
12	21	Etch v2

Fig. 7. Correlation of resistance values predicted by (a) OCD model (b) machine learning technique to the measured values

4.5 Predicting capacitance from OCD spectra: OCD model vs machine learning

In Fig. 8, the measured capacitance data is plotted on the vertical axis against the different measurement sites (explained in section 3.2) on the horizontal axis. It shows our expected target design trend of increased capacitance when the distance between the metal Line A and B decreases.

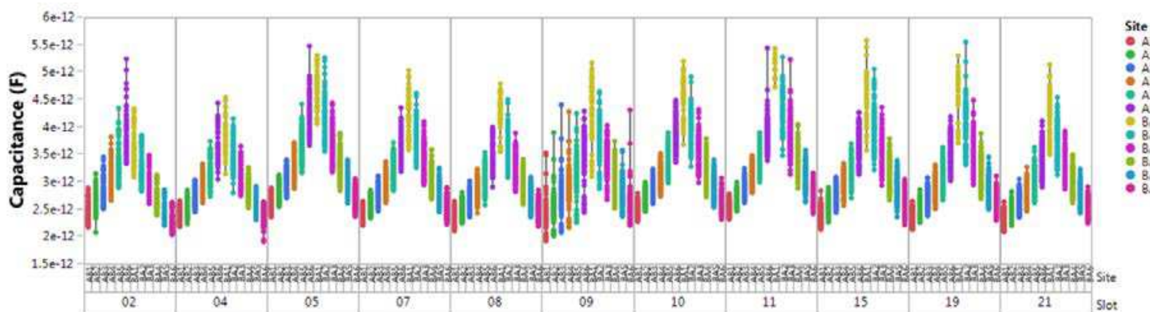


Fig. 8. Measured capacitance in vertical structures for different sites and wafers.

Once the space width, or *d* value (described in Fig. 9), was obtained from the model, the capacitance was calculated using the formula shown in Fig. 9b. It was found that there is a good correlation between the capacitance obtained from the OCD model and the measured capacitance value, although the slope is slightly high. The value of slope is dependent on parameters such as the relative

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 28 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

permittivity (k) and the area of the plates (A). Thus, any error occurring while calculating the theoretical value of capacitance could affect the slope.

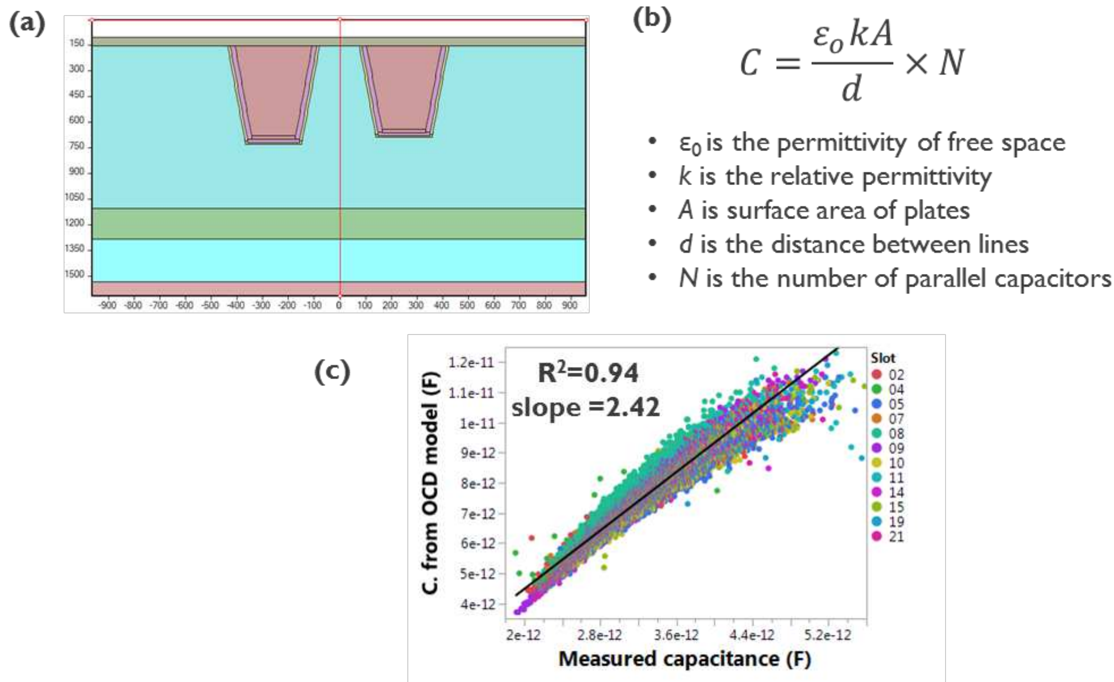


Fig. 9. (a) Illustration of the OCD model (b) formula to calculate the theoretical capacitance values (c) correlation plot between predicted capacitance (from OCD model) and measured capacitance

For the machine learning solution, we used three wafers for training (wafers 7, 9 and 11) and we obtained a training score of 0.99. The solution thus generated was used to predict the capacitance of other wafers in the lot and it still showed a good correlation, as shown in Fig. 10. In the histogram in Fig. 11, we compare the correlation between predicted and measured capacitance for different capacitance sites. Values predicted from the machine learning solution showed higher correlation as compared to OCD modelling.

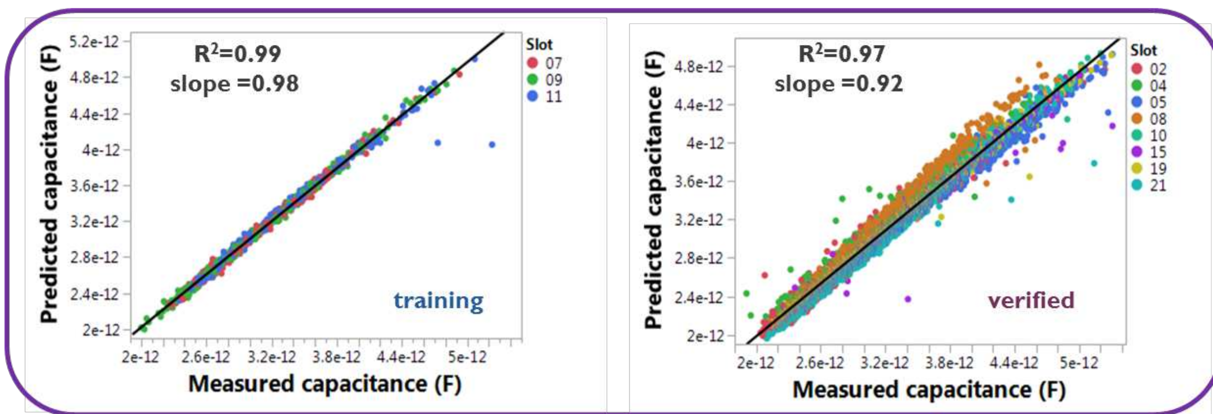


Fig. 10. Correlation plots between predicted capacitance using the machine learning technique and measured capacitance

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 28 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

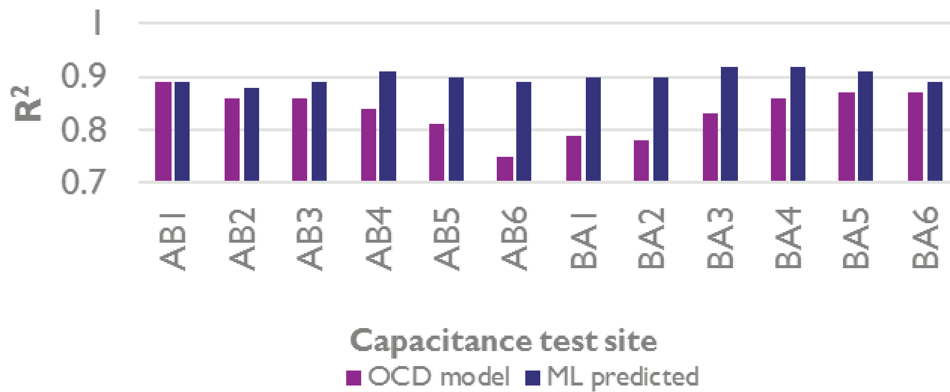


Fig. 11. Complementary machine learning solution using OCD spectra showing better R^2 values compared to OCD model

5 Conclusions

There is a motivation to introduce machine learning solutions for R&D and HVM for advanced technology nodes, which require multiple metrology steps for process monitoring and control—the increasing number of OCD steps can require a longer time to solution. Moreover, tighter process control budgets could be compromised by model complexity. The OCD technique is based on a rigorous coupled-wave analysis (RCWA), which is designated for periodic structures. Applying this technique for non-periodic structures, such as electrical structures, is a challenge. Using machine learning algorithms can overcome these three challenges and become a complementary approach, in addition to the RCWA modeling technique, for future process control and monitoring purposes. This work shows that machine learning using OCD spectra can predict electrical performance with high R^2 values, enable good correlation to reference CD, demonstrate good correlation to diffraction-based overlay and SEM-based overlay and improved correlation between measured and predicted resistance and capacitance compared to the OCD model. To complete this study, and qualify the machine learning approach, we will need to run a validation set over a longer time period, with multiple wafer batches.

Acknowledgement

This project has received funding from the Electronic Component Systems for European Leadership Joint Undertaking under grant agreement No. 737479. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation program in Netherlands, Belgium, Czech Republic, Hungary, France, Germany, Israel, Italy.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 28 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

References

1. A. Vaid *et al* "A holistic metrology approach: Hybrid Metrology Utilizing Scatterometry, CD-AFM and CD-SEM" Proc. of SPIE Vol. 7971 797103-18 (2011).
2. R. W. Collins and J. Koh, "Dual rotating-compensator multichannel ellipsometer: Instrument design for real-time Mueller matrix spectroscopy of surfaces and films," J. Opt. Soc. Am. A **16**, 1997–2006 (1999).
3. K. Hasumi et al. "SEM based overlay measurement between Via patterns and buried M1 patterns using high voltage SEM" Proc. of SPIE Vol. 10145 101451J-1 (2017)

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 28 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use