

PROCEEDINGS OF SPIE

[SPIDigitalLibrary.org/conference-proceedings-of-spie](https://spiedigitallibrary.org/conference-proceedings-of-spie)

Implementation of machine learning for high-volume manufacturing metrology challenges

Padraig Timoney, Taher Kagalwala, Edward Reis,
Houssam Lazkani, Jonathan Hurley, et al.

Padraig Timoney, Taher Kagalwala, Edward Reis, Houssam Lazkani, Jonathan Hurley, Haibo Liu, Charles Kang, Paul Isbester, Naren Yellai, Michael Shifrin, Yoav Etzioni, "Implementation of machine learning for high-volume manufacturing metrology challenges," Proc. SPIE 10585, Metrology, Inspection, and Process Control for Microlithography XXXII, 105850X (21 March 2018); doi: 10.1117/12.2300167

SPIE.

Event: SPIE Advanced Lithography, 2018, San Jose, California, United States

Implementation of machine learning for high volume manufacturing metrology challenges

Padraig Timoney^a, Taher Kagalwala^a, Edward Reis^a, Houssam Lazkani^a, Jonathan Hurley^a
Haibo Liu^b, Charles Kang^b, Paul Isbester^b, Naren Yellai^b, Michael Shifrin^c, Yoav Etzioni^c

^a GLOBALFOUNDRIES, 400 Stonebreak Road Extension, Malta, NY 12020, USA

^b Nova Measuring Instruments Inc., 3090 Oakmead Village Drive, Santa Clara, CA 95051, USA

^c Nova Measuring Instruments, LTD, P.O. Box 266, Weizmann Science Park, Rehovot 76100, Israel

ABSTRACT

With more advanced technology nodes and device architectures being used in high volume semiconductor manufacturing, process control and metrology are facing ever-greater challenges. More advanced metrology approaches have been emerging to cope with these challenges. This study investigates machine learning prediction as a complementary method to optical critical dimension (OCD) measurements. We have evaluated the suitability of machine learning to address high volume manufacturing metrology requirements for applications in both front end of line (FEOL) and back end of line (BEOL) sectors from advanced technology nodes. In the FEOL sector, we have demonstrated initial feasibility to predict the fin CD values from an inline measurement using machine learning. In the BEOL sector, machine learning is shown to provide direct prediction of electrical resistance using spectra collected from both OCD measurement sites and electrical test (e-test) sites. The predicted resistance correlation to the actual e-test value is improved in comparison with OCD results for multiple metal levels of various products. It is also shown that e-test predictions by machine learning may assist to alleviate the challenges that conventional metrology approaches are facing when the metal linewidth becomes smaller. Furthermore, impact of number of samples in training set was investigated and it was found that reducing number of samples in training set does not degrade the e-test correlation and Total Measurement Uncertainty (TMU) significantly. This paper demonstrates that predictive metrology based on machine learning is an advantageous and complementary technique for high volume semiconductor manufacturing. Together with conventional OCD measurements, a machine learning approach could successfully overcome complex metrology challenges in advanced semiconductor manufacturing.

Keywords: Machine Learning, High Volume Manufacturing, E Test, Process Control, Optical Metrology, Metrology Budget, Model Complexity

1. INTRODUCTION

In recent years, the combination of device scaling, complex 3D device architecture and tightening process tolerances have strained the capabilities of metrology tools to meet process needs. Two main categories of approaches have been taken to address the evolving process needs. In the first category, new hardware configurations are developed to provide more spectral sensitivity. Most of this category of work will enable next generation optical metrology tools to try to maintain pace with next generation process needs. In the second category, new innovative algorithms have been pursued to increase the value of the existing measurement signal. These algorithms aim to boost sensitivity to the measurement parameter of interest, while reducing the impact of other factors that contribute to signal variability but are not influenced by the process of interest.

Machine learning enables computer systems to learn with data without being explicitly programmed. It is incredibly powerful to make predictions or calculated suggestions based on advanced algorithms as well as large amounts of data. Machine learning and analytics was used to accurately predict the electrical performance of deep trenches and metal lines in the past [1, 2]. This paper has evaluated the suitability of machine learning to address high volume manufacturing metrology requirements in both front end of line (FEOL) and back end of line (BEOL) sectors from advanced technology nodes.

2. BACKGROUND AND MOTIVATION

Scatterometry, also called Optical Critical Dimension (OCD) metrology, is an optical measurement technique based on Rigorous Coupled Wave Analysis (RCWA), in which the light diffracted from a periodic structure is used to characterize the targeted geometric profile. Scatterometry has been proven as an advantageous method to provide the full profile information of various applications in FEOL and BEOL. Fin profile is critical for device performance and one application was investigated in this paper to obtain fin information at earlier process step by using machine learning together with OCD measurements. 3D geometric parameters of BEOL metal lines were successfully measured by OCD metrology and they demonstrated good correlation to physical cross sections and/or electrical resistance for BEOL etch and Chemical-Mechanical Planarization (CMP) processes [3]. In this work, machine learning was utilized to predict e-test resistance results based on reference data without explicit modeling, directly from reflectance spectra obtained with scatterometry tools after Cu CMP. The e-test prediction by machine learning is complementary to the OCD measurements in several aspects for BEOL Cu CMP applications. First, due to OCD modeling complexity, correlations between multiple modeling parameters can possibly cause accuracy issues. Second, to qualify an OCD model, the cross section reference is needed, which is mostly obtained by cutting wafers for Transmission Electron Microscopy (TEM). The TEM data acquisition period is usually days or even weeks, and the quantity is generally from a few dies of several wafers, very limited in comparison with e-test reference, which can be from a large number of wafers or lots. Lastly, machine learning is able to use e-test data directly as reference, providing feasibility for ultimate device performance monitoring or control. The e-test results are predicted early in the process flow (after Cu polishing) and this can assist process engineers to make actionable decisions.

However, e-test predictions by machine learning cannot provide geometric data like OCD modeling does unless there is sufficient reference (training) data regarding the parameter in question—which is usually not the case. These geometric parameters can provide detailed process information and enable more thorough process monitoring on BEOL thin film depositions, etch, and Cu polishing. This work investigates the capability of machine learning as a complementary technique to OCD metrology in a high volume manufacturing scenario. A combination of these two methods can be an advantageous approach in feedforward or feedback flows to optimize, control or monitor processes, thereby overcoming current and future semiconductor metrology challenges.

3. RESULTS

3.1 Fin CD Measurements

In this work, the fin CD was determined by using a machine learning technique. The variation of fin CD for wafers of four different products resulting from a fin hard mask Reactive Ion Etch (RIE) process was evaluated by OCD and a reference metrology source. A representative structure resulting from the fin hard mask reactive ion etch is presented in Figure 1.

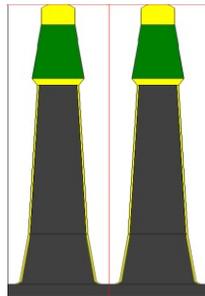


Figure 1: Representative structure from fin hard mask RIE

Wafers of four different products from eight process chambers were evaluated by OCD. As demonstrated in Figure 2, correlation of OCD and reference metrology provided linear correlation with R^2 less than 0.31 and slope less than 0.6 for each of the 4 products. Some chambers demonstrate significant variation of OCD values between wafers, suggesting a chamber related effect on the wafer structure.

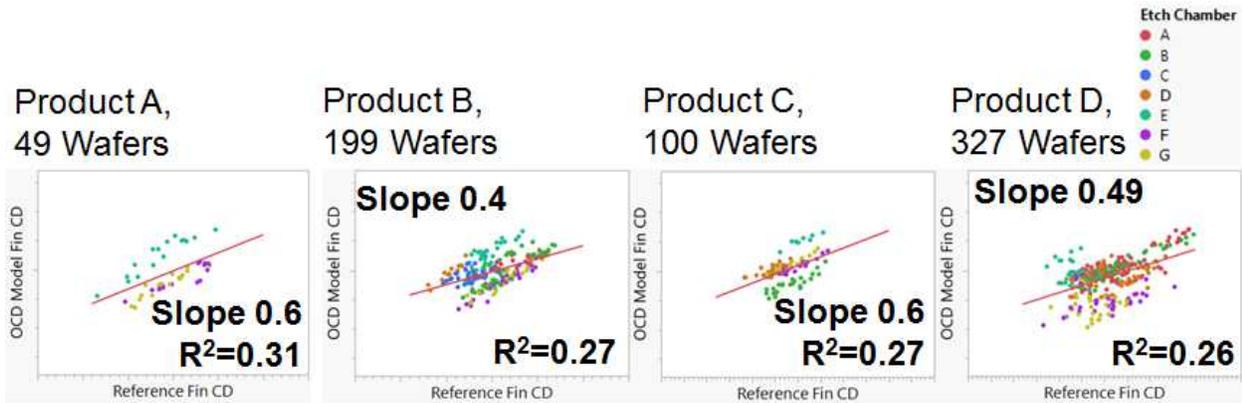


Figure 2: Correlation of reference metrology to OCD fin CD

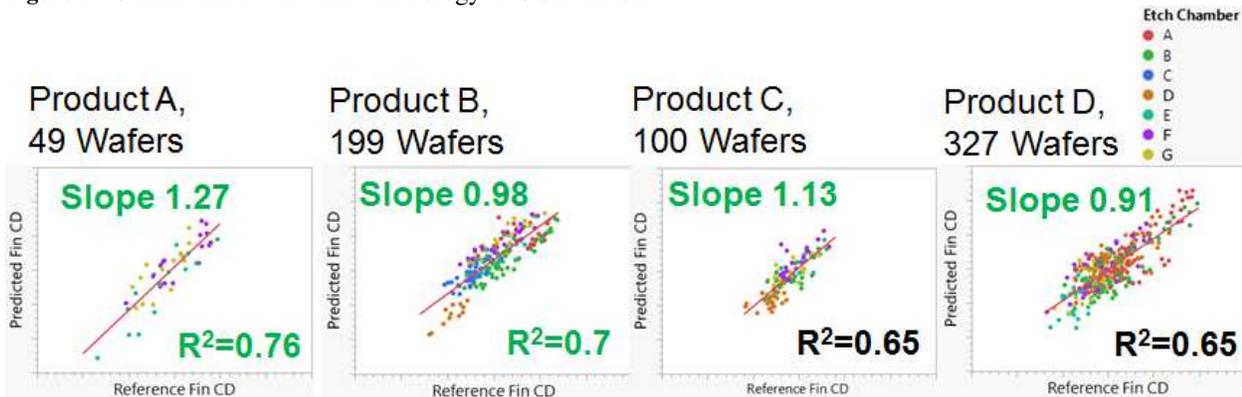


Figure 3: Correlation of reference metrology to predicted fin CD by machine learning

Product D, 327 Wafers

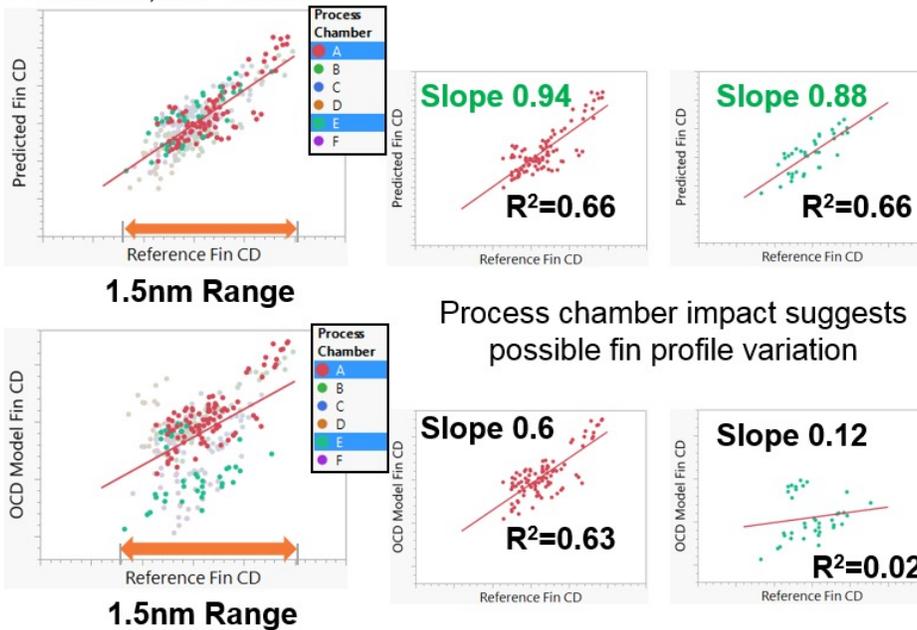


Figure 4: Product D wafers evaluated with machine learning and standard OCD.

Some chambers were observed to have wide wafer to wafer variability of fin CD as reported from the OCD model. The reference metrology did not show a chamber dependent distribution. Applying machine learning to the same wafers

provided significant improvement in the linearity as demonstrated through R^2 and slope as shown in Figure 3. Figure 4 further demonstrates that chamber E provided significant variations of OCD that were not observed in the reference metrology. Applying machine learning to the optical spectra provided a similar response which removed the chamber dependence on the variability of the OCD.

3.2 BEOL E-Test Prediction

In this paper, two groups of products with different BEOL measurement scenarios were investigated: for product group A, the OCD measurement site being used for OCD modeling, was the same site as e-test site; for product group B, the OCD site and e-test site are different and located at the two distant pads within die.

For product group A, scatterometry spectra and e-test data were collected from the same site and e-test data were used as reference for machine learning. A solution was built up from machine learning and used for subsequent e-test data predictions. With the machine learning solution, untrained spectra can be interpreted to predict e-test data, either inline from the tools or offline from a computer with the appropriate software. Figure 5(a) shows the correlations of OCD Trench Height to actual e-test, and predicted e-test to actual e-test data. Cu Trench geometric results such as Trench Height and Trench BCD were obtained from a Nova OCD model, as shown in Figure 5(a). The key parameter Cu Trench Height is usually used for Cu polishing process monitoring and control. As it can be seen, OCD Trench Height has good correlation to the actual e-test Rs, while predicted e-test Rs has even better correlation to the actual e-test. R^2 is better for machine learning and the scatter of the data about the straight line fit was seen to improve for Machine Learning also relative to OCD data. These results can be well explained by various factors such as possible poor quality of TEM reference, correlations in the OCD model, and much larger quantity of e-test reference data than TEM reference being used.

In product group B, OCD measurement site and e-test site were different. Scatterometry spectra were measured at both sites, for OCD modeling (OCD site) and machine learning (OCD site and e-test site). Figure 5(b) shows the correlations of OCD Trench Height to actual e-test, and predicted e-test to actual e-test results. For product group B, because OCD Trench Height was obtained from a different site where e-test was conducted, the correlation between Trench Height and actual e-test is worse than that of product group A. We believe this was caused by site-to-site Cu trench height difference within the die. This within die variation could be potentially gauged by e-test prediction at two sites, which will be our investigations in the future. By using machine learning, the predicted e-test results from OCD site have improved correlation to the actual e-test in comparison with OCD Trench Height. When machine learning was applied on the spectra collected from the e-test site, the correlation of the predicted e-test to the actual e-test was further improved, very close to that of product group A.

For both product groups, the predicted e-test data by machine learning demonstrate better correlations to the actual e-test than the conventional OCD measurement. Furthermore, we investigated a large number of samples (sites) from different Cu metal levels, which is shown in Table 1. Figure 6 presents correlation results for all these metal levels and we can observe similar results like those in Figure 5, except for M2 of product group A, in which the correlation R^2 of the predicted e-test to the actual e-test was about 0.79, not as good as other e-test predictions (further analysis and optimizations will be explored for this application).

It can also be seen that for product group A, the higher the metal level, the better the correlations of both OCD trench height and predicted e-test to the actual e-test. This results from the fact that if the metal level is higher, the top Cu trench height is larger, which blocks more light and reduced the spectral contributions of underneath Cu lines, so both OCD trench height and predicted e-test is more accurate because the top Cu lines dominate the sensitivity of scatterometry spectra. In product group B, due to site-to-site Cu trench height difference, the correlation of Trench Height to the actual e-test is worse than that of the predicted e-test, particularly for metal level C5 & G1. As stated, a large number of untrained sites or wafers were investigated, and after the training was done, there were no observations of training becoming invalid over 3 month interval. We expect this will continue to work if there is no process change.



A – Products with meas site at e test site

(a)



B – Products with different e test site from meas site

(b)

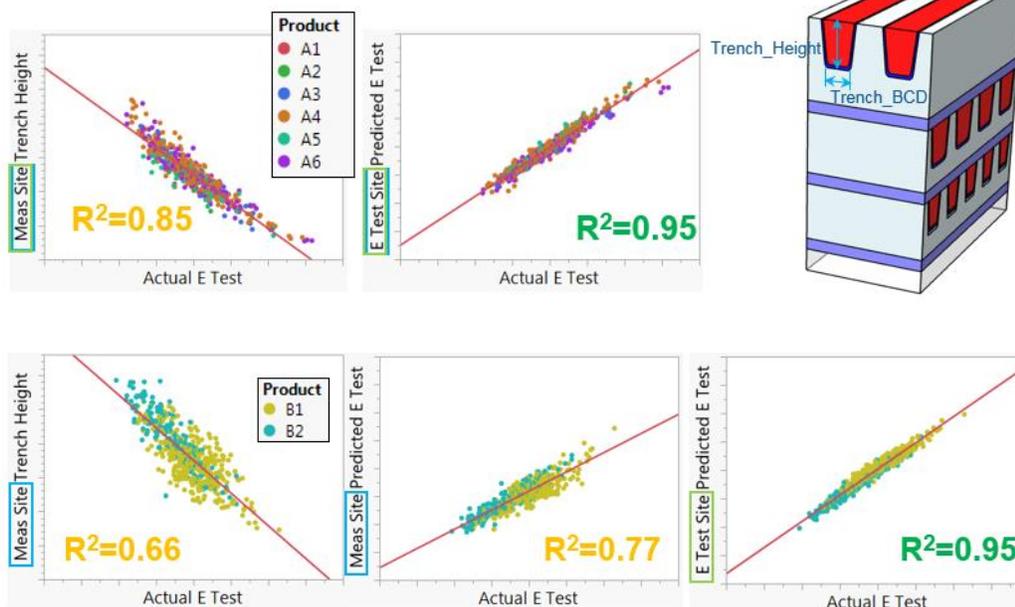


Figure 5: The correlations of OCD Trench Height to actual e-test, and predicted e-test to actual e-test data for product group A (a) and B (b). OCD measurement site or “meas site” is the same as the e-test site within die for product group A, and different for product group B.

Product Scenario	Metal level	M1	M2	C5	G1
A – Products (Meas at E Test Site)	# Samples	6760	6816	6708	1573
B – Products (Meas Site not at E Test Site)	# Samples	3575	3484	6832	1911

Table 1: The numbers of samples (sites) for different metal levels being used in this work.

Based on the results from Figure 6, the correlation R^2 of the predicted e-test to actual e-test for different metal levels were plotted against the normalized Cu line pitch, as shown in Figure 7. Here the R^2 values were calculated based on the best data for both OCD measurement site Trench Height or e-test site predicted e-test data stream from all products. As it can be seen, the R^2 improvement in predicted e test is more significant for decreasing metal linewidth, indicating conventional OCD measurements face more challenges when the metal linewidth becomes smaller, while e-test predictions by machine learning can alleviate this issue. As OCD measurements can provide geometric profiles for Cu line and dielectric stack structure, which are missing in e-test predictions based on machine learning, incorporating two techniques will enable more comprehensive and accurate outputs for thorough process monitoring or control. This is becoming particularly significant as process technologies advance to 7nm process node and below.

When machine learning is used in high volume manufacturing metrology, the number of samples that are needed for training becomes an important question. Using metal level C5 case as an example, we built up different machine learning libraries with different numbers of training samples, then apply them on the same untrained sites to investigate the impact of different training sets on the correlation to actual e-test.

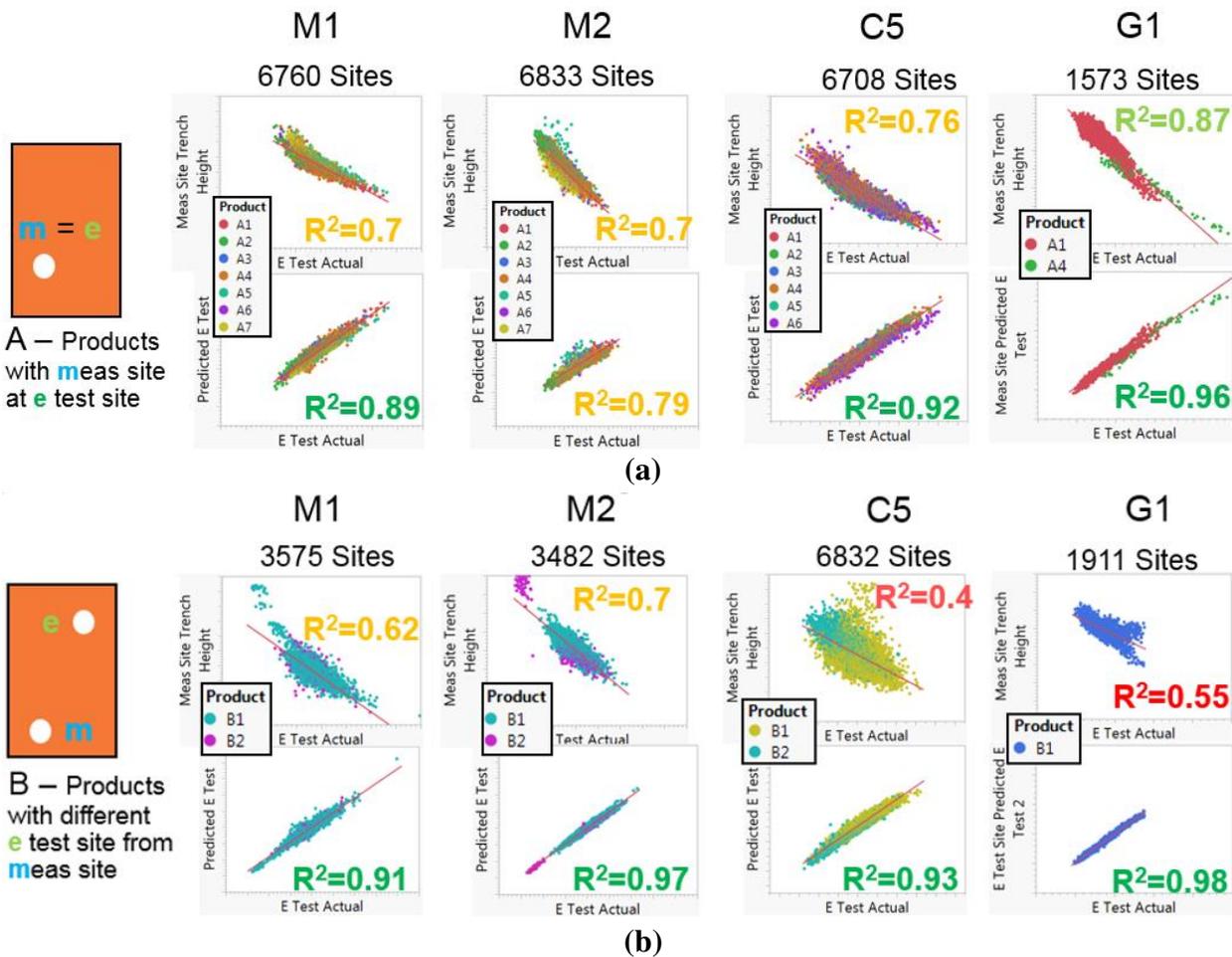


Figure 6: The correlations of OCD Trench Height to actual e-test, and predicted e-test to actual e-test data for different metal levels of product group A (a) and B (b). Note all predicted e-test results were obtained from the scatterometry spectra collected at the e-test site. OCD measurement site or “meas site” is the same as the e-test site within die for product group A, and different for product group B.

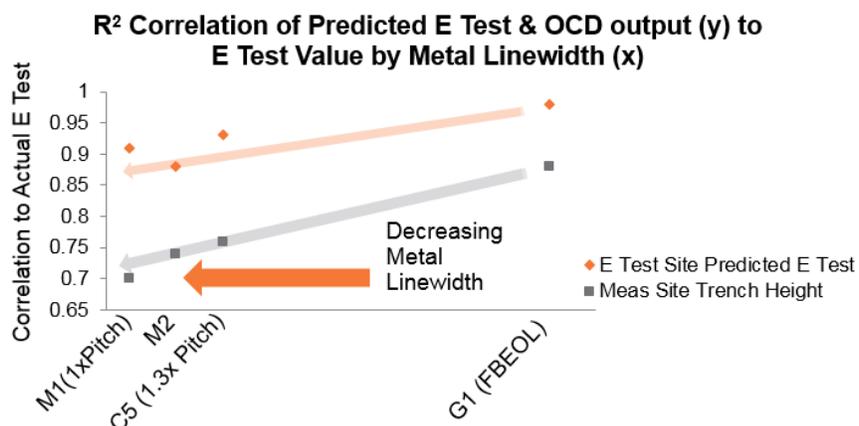


Figure 7: The correlation R^2 of the predicted e-test to actual e-test for different metal levels versus the normalized Cu line pitch (FBEOL - far back end of the line).

The results are shown in Figure 8. It’s noted that reasonable correlation was obtained even the verification set has wider range (~50% wider) than the training set, which means extrapolation was successful for this machine learning. It can be

seen that good correlation was obtained even only 20% of the original training set was used. If the process variability is well covered by training set, the reducing number of lots in the training set does not significantly degrade the correlation and Total Measurement Uncertainty (TMU), therefore in high volume manufacturing we might not need a large number of reference data for training set. In practice, as a rule of thumb, a certain number of lots were selected for training set based on reference ranges to cover the process variations.

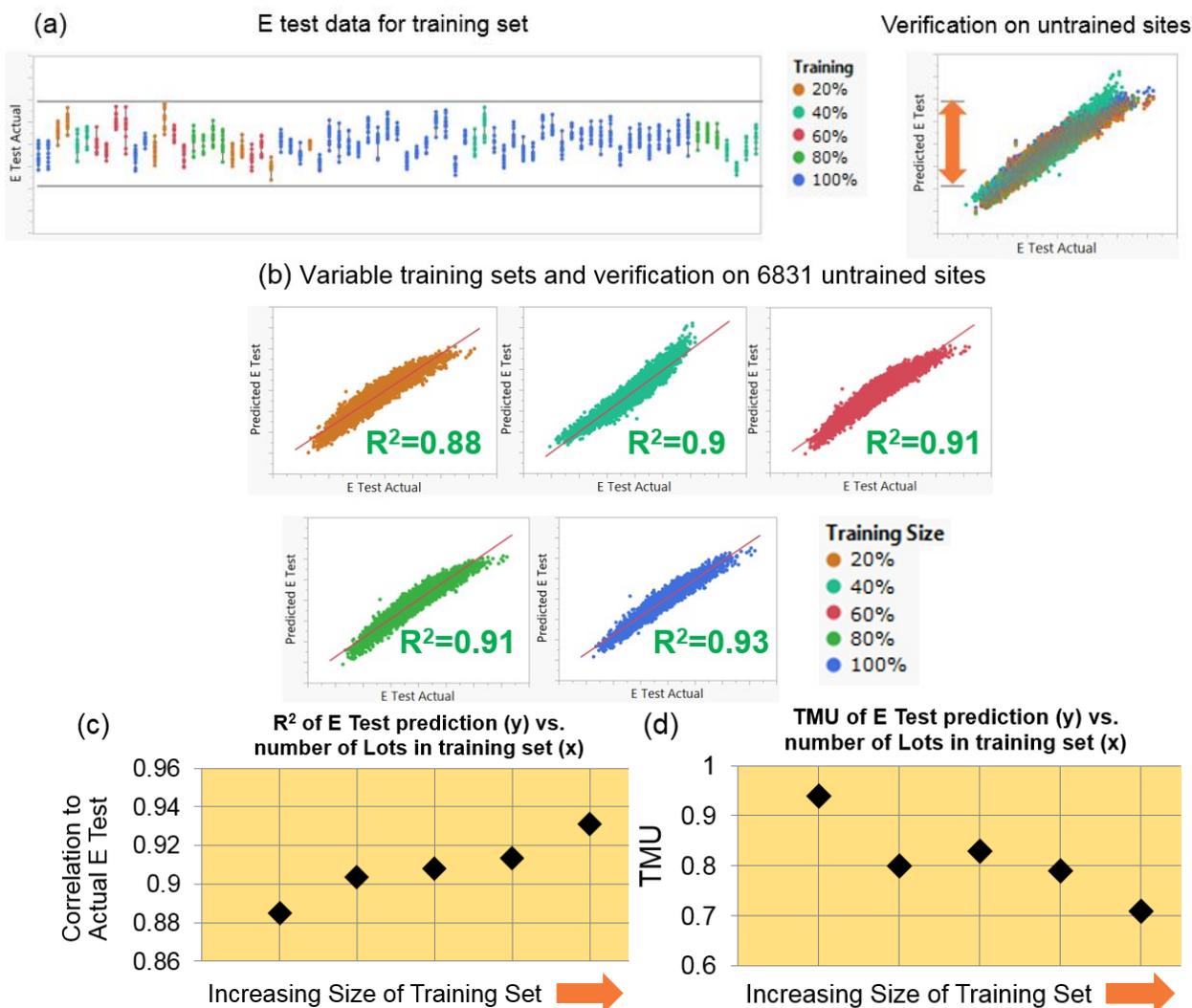


Figure 8: Impact of number of samples in training set. (a) Variability chart of e-test reference data (left) and predicted e-test correlation to actual e-test on untrained sites. Plots for 100% means all training Lots were used in training; plots for 20% means 20% of whole training lots was used for training. Note in e-test reference variability chart, training sets with larger lot numbers cover all training sets with smaller lot numbers; (b) Separated e-test correlation plots when the machine learning solution was trained by different number of lots; (c) Correlation R^2 of e-test prediction plotted against the number of lots in training set; (d) TMU of e-test prediction plotted against the number of lots in training set.

4. CONCLUSIONS

In this work, the value of machine learning based prediction is demonstrated, leveraging OCD spectra from FEOL and BEOL applications and applicable reference metrology. In FEOL the feasibility of fin CD prediction using machine learning is established with improved correlation to reference metrology than from the fin CD measurement by OCD. For

BEOL, we have shown that predicted e-test results by machine learning have better correlations to the actual e-test data than conventional OCD measurements across multiple products and metal levels, particularly for decreasing metal linewidth. Multiple metal levels and spectral acquisition scenarios with large data sets are presented, verifying that predictive metrology based on machine learning can be an advantageous technique in high volume semiconductor manufacturing. Furthermore, the impact of training set size was evaluated and it was found that reducing number of lots in the training set did not significantly degrade the correlation and TMU, suggesting that a large volume of reference data may not be required for the training set. As OCD measurements can provide detailed geometric profiles for complex device structures, which are missing in e-test predictions based on machine learning, incorporating both techniques will enable more comprehensive and accurate outputs for thorough process monitoring and control in semiconductor manufacturing.

REFERENCES

- [1] N. Rana, Y. Zhang, D. Wall; B. Dirahoui, "Predictive data analytics and machine learning enabling metrology and process control for advanced node IC fabrication", 2015 26th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC)
- [2] M. Breton, R. Chao, G. Muthinti, A. de la Peña, J. Simon, A. Cepler, M. Sendelbach, J. Gaudiello, H. Tang, S. Emans, M. Shifrin, Y. Etzioni, R. Urenski, W. T. Lee, "Electrical test prediction using hybrid metrology and machine learning", Proceedings of SPIE, Volume 10145, Metrology, Inspection, and Process Control for Microlithography XXXI (2017)
- [3] P. Timoney, J. Tsai, S. Baral, L. Economikos, A. Vaid, H. Liu, B. Kang, P. Isbester, P. Dasari, R. Koret, N. Yellai, "Comprehensive BEOL Control using Scatterometry and APC", Proceedings of SPIE, Volume 9424, Metrology, Inspection, and Process Control for Microlithography XXIX, 2015