Advanced machine learning eco-system to address HVM optical metrology requirements

Timoney, Padraig, Luthra, Roma, Elia, Alex, Liu, Haibo, Isbester, Paul, et al.

Padraig Timoney, Roma Luthra, Alex Elia, Haibo Liu, Paul Isbester, Avi Levy, Michael Shifrin, Barak Bringoltz, Eylon Rabinovich, Ariel Broitman, Eitan Rothstein, Ran Yacoby, Ilya Rubinovich, YongHa Kim, Ofer Shlagman, Barak Ben-Nahum, Marina Zolkin, Igor Turovets, "Advanced machine learning eco-system to address HVM optical metrology requirements," Proc. SPIE 11325, Metrology, Inspection, and Process Control for Microlithography XXXIV, 113251H (20 March 2020); doi: 10.1117/12.2552058

SPIE.

Event: SPIE Advanced Lithography, 2020, San Jose, California, United States

Advanced Machine Learning Eco-System to Address HVM Optical Metrology Requirements

Padraig Timoney*[a], Roma Luthra[a], Alex Elia[a], Haibo Liu[b], Paul Isbester[b], Avi Levi[b], Michael Shifrin[c], Barak Bringoltz[c], Eylon Rabinovich[c], Ariel Broitman[c], Eitan Rothstein[c], Ran Yacoby[c], Ilya Rubinovich[c], YongHa Kim[c], Ofer Shlagman[c], Barak Ben-Nahum[c], Marina Zolkin[c], Igor Turovets[c]

[a]GLOBALFOUNDRIES, 400 Stonebreak Road Extension, Malta, NY 12020, USA; [b]Nova Measuring Instruments, Inc., 3342 Gateway Blvd, Fremont, CA 94538, USA; [c]Nova Measuring Instruments, LTD, P.O. Box 266, Weizmann Science Park, Rehovot 76100, Israel

ABSTRACT

Machine learning (ML) techniques have been successfully deployed to resolve optical metrology challenges in semiconductor industry during recent years. With more advanced computing technology and algorithms, the ML system can be improved further to address High Volume Manufacturing (HVM) requirements. In this work, an advanced ML eco-system was implemented based on big data architecture to generate fast and user-friendly ML predictive models for metrology purposes. Application work and results completed by using this ML eco-system have revealed its capability to quickly refine solutions to predict both external reference data and to improve the throughput of conventional Optical Critical Dimension (OCD) metrology. The time-to-solution has been significantly improved and human operational time has also been greatly reduced. Results were shown for both front end and back end of line measurement applications, demonstrating good correlations and small errors in comparison with either external reference or conventional OCD results. The incremental retraining from this ML eco-system improved the correlation to external references, and multiple retrained models were analyzed to understand retraining effects and corresponding requirements. Quality Metric (QM) was also shown to have relevance in monitoring recipe performance. It has successfully demonstrated that with this advanced ML eco-system, streamlined ML models can be readily updated for high sensitivity and process development applications in HVM scenarios.

Keywords: Machine learning, Optical Critical Dimension (OCD), big data, High Volume Manufacturing (HVM), incremental retraining, correlation, time to solution, Quality Metric (QM)

1. INTRODUCTION

1.1 Background

As semiconductor devices continue to shrink in size while more complex 3D structures are being used, process control and metrology have been facing ever-greater challenges. Advanced metrology techniques are in demands for all aspects of semiconductor R&D and manufacturing process control. On the one hand, conventional metrology technologies such as Atomic Force Microscopy (AFM), Scanning Electron Microscopy (SEM), Transmission Electron Microscopy (TEM), and scatterometry, have been further advanced to provide better accuracy and higher sensitivity to cope with these challenges. Hybrid or combined metrology of these techniques is also an important strategy that could be used to extend the applicability of current instruments [1]. On the other hand, the Machine Learning (ML) approach based on advanced data analytics and innovative algorithms has also been pursued to extract additional values of the existing measurements from these metrology techniques.

Scatterometry, an optical measurement technique based on Rigorous Coupled Wave Analysis (RCWA), can provide geometric profiles characterized from the optical diffraction or reflectance spectra from a periodic structure with nanometer-scale features — it is also called Optical Critical Dimension (OCD) metrology. Scatterometry has been proven

Metrology, Inspection, and Process Control for Microlithography XXXIV, edited by Ofer Adan, John C. Robinson, Proc. of SPIE Vol. 11325, 113251H · © 2020 SPIE
CCC code: 0277-786X/20/$21 · doi: 10.1117/12.2552058

as an advantageous method to provide the full profile information of various applications in advanced technology nodes. However, the combination of device size scaling, complex 3D architecture, and tightening process tolerances have challenged the capabilities of scatterometry. In recent years, the ML predication has been used as a complementary method to OCD metrology. ML and analytics were used to accurately predict the electrical performance of deep trenches and metal lines in the past [2-5]. In this approach, scatterometry spectra collected from relevant semiconductor device structures, together with reference data, are used for ML training to build up predictive models for inline measurements.

ML gives computers the ability to learn with data without being explicitly programmed [6]. The conventional OCD modeling requires to build up a physical model with geometric profiles for CD measurements, while the ML-based predictive metrology removes necessity of building complex physical models for explicit optical simulations, instead it relies on a large set of spectra and reference for training to obtain generalized regression models. Thanks to advanced algorithm, tremendous computing power, and large amounts of data nowadays, the ML based metrology becomes incredibly powerful to make predictions or calculated suggestions. The ML enabled predictive metrology provides an alternative to the conventional metrology techniques, particularly in the regime where the direct measurements are impossible or impractical. The introduction of ML to resolve optical metrology requirements has been well documented in the past few years, accompanied by improvements in measurement quality, time-to-solution and enabling a closer link to device electrical test parameters. In our previously published work [5], improvements to sensitivity of critical geometric parameters and the prediction of future device performance were demonstrated in High Volume Manufacturing (HVM) using ML derived predictive models. Compared with conventional semiconductor metrology techniques, the ML predictive metrology is an advantageous and complementary methodology, which are more specifically reflected in the following aspects.

- Conventional OCD models have faced increasing challenges from cross-parameter correlation or inducing error when parameters causing spectral variation are presumed fixed in the model. There is also often some uncertainty in how to link variations of a particular parameter with the expected device performance impact. The benefit of ML approach was clearly established in minimizing these errors and uncertainties associated with conventional OCD models.
- For conventional scatterometry, spectra need to be collected from a periodic structure for RCWA-based OCD modeling, and such a periodic structure is usually built as a dedicated OCD test site in the scribe line. These structures often don’t have similar processing to real device structures or e-test sites, which are sometimes nonperiodic or too complex for modeling. With the availability of good quality of reflectance spectra and reference data, it is feasible for the ML predictive metrology to measure such non-periodic structures and overcome the limitations of the conventional scatterometry — this is critical for accurate process monitoring and control.
- ML techniques may help reduce cost or achieve high throughput improvements by using faster measurements instead of lower throughput and more expensive traditional metrology techniques. For example, it was reported that ML and hybrid metrology using high throughput scatterometry and lower throughput X-ray fluorescence (XRF) were applied to detect voids in copper lines, having achieved a scatterometry throughput with a high sensitivity close to that of XRF [4].

However, the ML metrology can usually be used as a complementary technique to OCD metrology because it requires a lot of reference data for training, and the OCD model can readily provide detailed geometric parameters. If combined with OCD and other metrology techniques, ML predictions can enable more sophisticated and powerful metrology to meet HVM requirements for the semiconductor industry.

1.2 Motivation

Although the ML metrology has been an interesting focus in recent years, new challenges have emerged when ML derived solutions were implemented in volume manufacturing, particularly for a sustainable methodology to train a model with sufficient sample size of the applicable “reference.” The methodology previously used to verify the accuracy of the ML output vs. the applicable reference had to rely heavily on time consuming data mining, often using very large data sets. A more automated system therefore appears desirable to make the process less labor intensive and to reduce human error.

ML community has reached a consensus that “data matters more than algorithms for complex problems”, since it was emphasized by the article “The Unreasonable Effectiveness of Data” in 2009 [7]. This big data requirement has the same significance for ML metrology in high volume semiconductor manufacturing. In addition, frequent retraining or incremental retraining is highly desirable for many applications to ensure the prompt adaption to process changes or other variations such as tool drifting, reference accuracy changes. More ML training jobs, on the other hand, require shorter time-to-solution and less resource investment. The previous manual data mining practices used to create and validate such ML models would have prevented such large-scale model developments, implementation and retraining. In order to meet these demands, an ML system with a capacity to handle big data is desirable to enable systematic analysis and information extraction, or to handle large data sets that are too large or complex to be processed by traditional dataprocessing software.

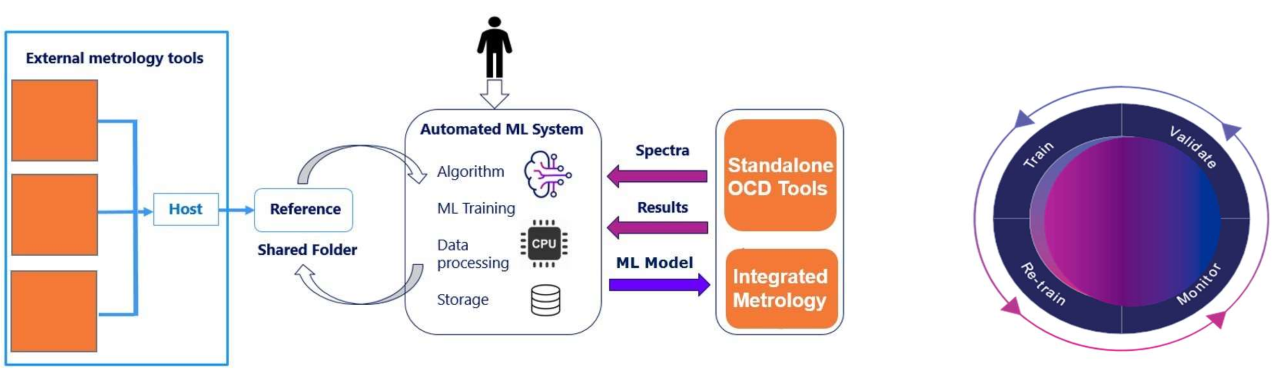
Another challenge for the ML metrology is to clean or avoid poor-quality data and reference. Most data scientists spend a significant part of their time on cleaning up the training data [6]. To achieve much more efficient data analysis, an advanced ML metrology system integrated with the existing measurement tools and reference data (or host database) is highly desirable, particularly for advanced technology nodes with complex 3D structures.

To overcome these obstacles and limitations, more sophisticated ML system with big data processing capability and high-level automation need to be developed. In this work, such an advanced ML eco-system was developed, installed, and evaluated, which greatly expands the capacity of ML metrology while significantly improving the time-to-solution.

2. ADVANCED MACHINE LEARNING ECO-SYSTEM

2.1 System Description

In comparison with the conventional ML system presented in the previous SPIE work [5], this advanced ML eco-system has a huge infrastructure change thereby with data handling and processing advantages based on a big data architecture and ML algorithm. In our previous ML system, spectra need to be downloaded from scatterometry tools, and reference data need to be acquired from external metrology tools to be paired with spectra for ML training. All these jobs were completed offline manually. The proposed ML eco-system is well integrated with the OCD metrology fleet, allowing reflectance spectra to be automatically streamed and stored in the system after wafer measurements. Figure 1 shows the schematic of the ML eco-system and its data process flow. Once external reference data are available, reference files are copied to a shared folder or directly imported into the ML system. Then the system can automatically pair the spectra and reference data based on wafer ID and die numbers or X/Y coordinates. ML training and cross validations can be completed for selected time span through a user-friendly interface. The ML model can be readily generated, then distributed to all metrology systems or exported for offline use. In addition, most data processing, analysis, and storage can be conducted with high speeds within the system. After implementing a ML metrology recipe, engineers can monitor the process data/chart and corresponding Quality Metric (QM) and validate new wafers when reference data are available. In this procedure, the human labor was significantly reduced — the engineer's operational time for creating a ML model was only about 5% of that for a previous ML recipe. When QM is poor or correlation to new reference is not ideal, retraining can be quickly conducted, and the recipe can be easily updated. It should be noted that all above procedures can be readily completed by a process engineer after 1-2 hours of training, unlike the many other metrological modeling techniques which usually need days or weeks of training to create measurement recipes.



(a) (b)

Figure 1. (a) Schematic of the ML eco-system to create, retrain and monitor ML metrology solutions, with significantly reduced human operations; (b) Data process flow in this ML eco-system.

2.2 Features and Advantages

This advanced ML eco-system provides superior data scalability, connectivity and retention. Its big data architecture allows the system to process large amounts of data, and to build and deploy custom ML models at the corresponding scale with a high speed. The ML eco-system is organically integrated with the OCD metrology fleet, allowing instant inline training and analysis, convenient recipe creating and distribution, as well as large data storage. These features enable the study of different retraining models and analysis of large sets of data, which could help discover patterns that were not apparent immediately, assisting “data mining” in high volume semiconductor manufacturing.

Time-to-solution is an important criterion for generating a metrology recipe, particularly for the recipe that needs to be updated frequently. With this advanced ML eco-system, the time-to-solution can be significantly improved: ML model creation in minutes instead of hours. Figure 2 shows time-to-solution comparisons between the conventional ML model creation and model creation using an advanced ML eco-system. As reflected from multiple application examples, human operational time for creating such an advanced ML recipe is about 5% of that for creating a conventional ML recipe.

Figure 2. Time-to-solution comparisons between our previous conventional ML system and the advanced ML eco-system for example applications.

*These comparisons were only based on relatively smaller training sets which the previous ML system can process. Some of the incremental retraining in this paper included many more training samples which cannot be handled through the conventional ML system due to its limited capacity.

ML models have an associated Quality Metric (QM) allowing engineers to monitor the performance of ML models. This QM is an indication of how well the spectra under interpretation is contained within a training set. When the QM becomes worse, inline results would be analyzed and a retraining could be proceeded if necessary, to prevent possible excursions.

The advanced ML system discussed in this work offers advantages to Fab engineers for easy recipe development and reducing human errors during model development, recipe preparations, and data analysis while incorporating scalability to process very large amounts of data in an automated approach, with significantly improved time-to-solution.

3. RESULTS AND DISCUSSIONS

Five different applications from Front End of Line (FEOL) and Back End of Line (BEOL) have been evaluated in this work including:

1) Fin CD measurements by predicted external reference
2) Thin film thickness measurements by predicted external reference
3) BEOL resistance measurements by predicted external reference
4) Throughput optimization for OCD thin film measurements
5) Throughput optimization for OCD BEOL etch measurements

These five applications were evaluated with different ML retraining intervals, based on either external or internal references including critical dimension (CD), thin film thickness, and e-test resistance, aiming to cover comprehensive cases to validate this methodology. Results for incoming new wafers were compared to applicable reference data, and R^2, Error %, and QM were presented for incremental training time intervals.

3.1 Fin CD Prediction

In this application, CD from an external metrology was used as reference data and predicted fin CD results were generated by the ML model for process guidance. Figure 3 compares the fin CD measurements by external reference to the data obtained from the ML model that was retrained 3 times after the initial model release. The mean and standard deviation differences of ML_Predict were +0.9% and -9.9% respectively relative to Ext_Ref. The R^2 values are shown above the respective time period and model revision. Figure 4 shows the R^2 of ML_Predict (Y axis) vs. Ext_Ref (X axis) for the four time periods, model revision combinations of Figure 3 (i.e. A1, B2, C3, D4) and the corresponding Error %, and QM trend.

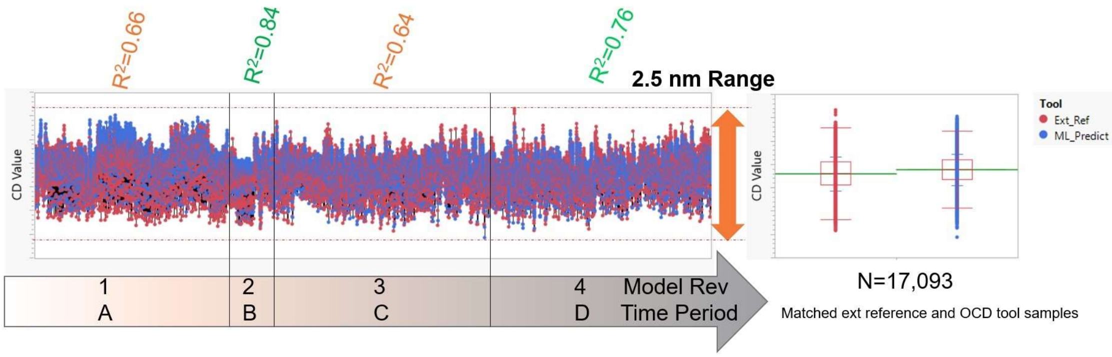


Figure 3. ML Predicted CD (labelled as ML_Predict) and external CD reference (labelled as Ext_Ref) show close baselines with a ~2.5 nm range. For 17,093 sampling dies, the averaged ML predicted CD has 0.6% higher mean, and 9.9% lower overall standard deviation than the external CD reference.

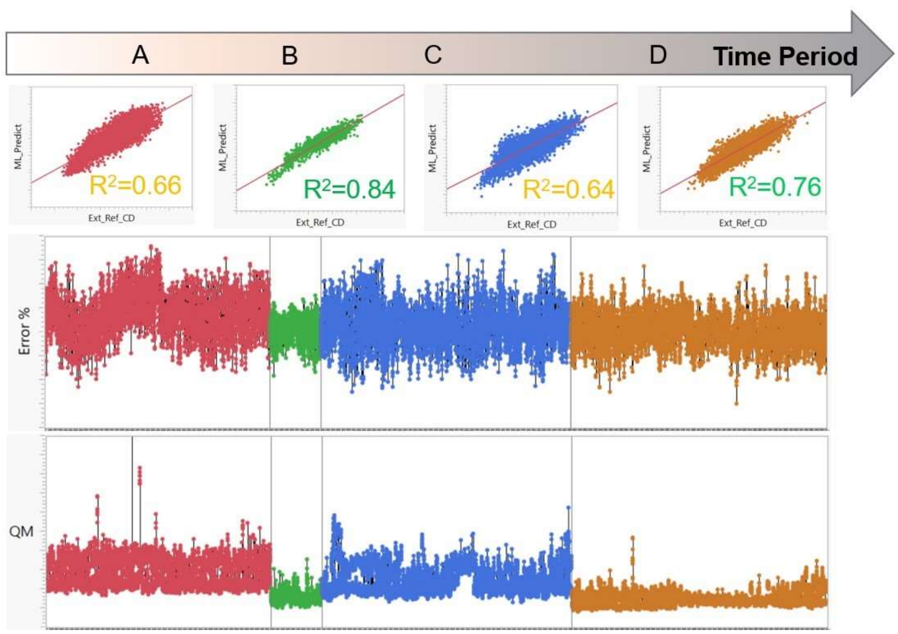


Figure 4. R^2, Error % and QM trend for fin CD prediction during blind test.

The sample size of time period B is noted to be smaller than that of time periods A, C and D and the QM performance is noted to be generally improved with time. In Figure 5, the R^2 and absolute value of Error % from the four time periods A, B, C and D were tabulated for all 4 revisions of the Model 1, 2, 3 and 4 in a matrix format. In this R^2 matrix, observations along the diagonal correspond to the values in Figures 3 and 4 coming from the first blind test of each model, values below the diagonal capture the R^2 for wafers included in the training and values above the diagonal are expected to have lower R^2 in cases where the retrain capability is improving the measurement. Since the time period B has much lower sample size, it is suspected that the R^2 values would not be reproducible for a larger sample size. Hence it is concluded that the best sustained R^2 performance is achieved in time period D, revision Model 4, as also observed in the Error % table.

Time Period	A	B	C	D
# Samples	1,997	1,241	7,157	6,698
Model 1	0.66	0.77	0.61	0.67
Model 2	0.76	0.84	0.65	0.76
Model 3	0.81	0.83	0.64	0.74
Model 4	0.82	0.86	0.7	0.76

Time Period	A	B	C	D
# Wafers	1,997	1,241	7,157	6,698
Model 1	1.84	1.50	1.30	1.26
Model 2	1.43	1.33	1.37	1.23
Model 3	1.03	0.94	1.24	1.10
Model 4	1.00	0.78	1.10	1.03

Figure 5. R^2 (left) and absolute Error % (right) matrices for ML predicted fin CD results.

3.2 Thin Film Thickness Prediction

Figure 6 compares the ML predicted film thickness to the external reference, similar to the format shown in Figure 3. The average from all ML_Predict samples is 0.13% higher than that of the Ext_Ref with 7.9% higher standard deviation. Figure 7 illustrates the R^2, Error % and QM trend. Revision Model 1 was found to have high QM and Error % on some samples measured in time period A. It's also found that revision Model 4 had a bimodal distribution in R^2, Error % and QM. Revision Model 1 is carefully examined in Figure 8, where an additional population of wafers, Training 0, are added. The observation is that these wafers with higher (poorer) QM and larger Error % are also found to have a different population in the R^2 plot which was not observed during model training.

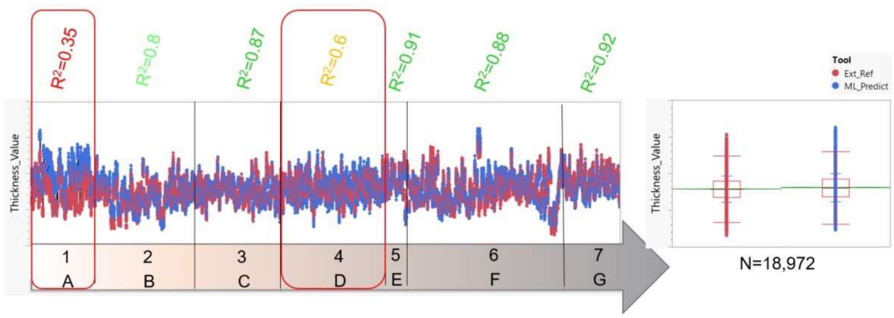


Figure 6. Comparison of thin film thickness results from external reference (Ext_Ref) and ML model (ML_Predict). For 18,972 sampling dies, the averaged ML predicted thickness has 0.13% higher mean, and 7.9% higher overall standard deviation than the external reference.

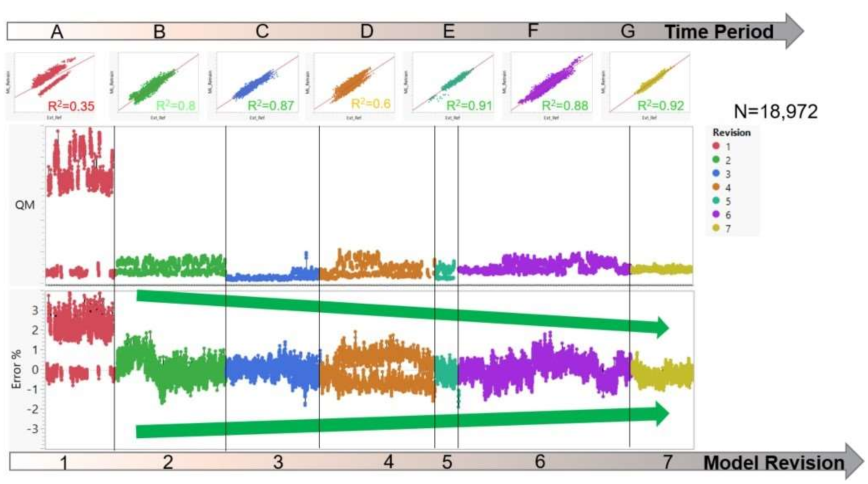


Figure 7. R^2, Error % and QM trend for thin film thickness prediction during blind test

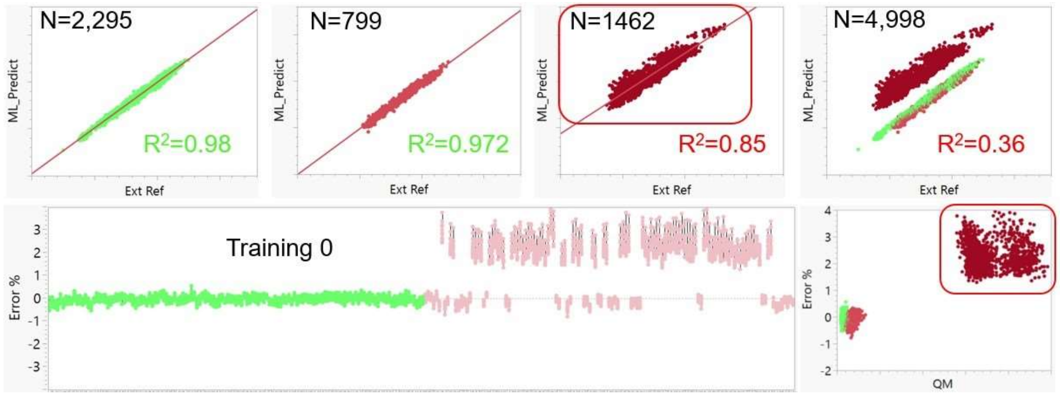


Figure 8. R^2, Error % and QM for thin film thickness prediction from Model 1 applied on training and blind test samples.

In Figure 9, selected wafers from time period D, and revision Model 4 are similarly highlighted to compare R^2, Error % and QM. Generally high QM wafers are observed to have high Error % and follow a different R^2 trend. The wafers are also suspected to have variability modes not being captured during model training. The R^2 variations from time period A/Model 1 and time period D/Model 4 are addressed by Model 6 and 7, where the R^2 and absolute Error % are optimal throughout time periods A through G as shown in Figure 10. It is also noted as an unexpected finding that the R^2 and absolute Error % of model revisions 2, 3 and 4 recovers in time periods F and G.

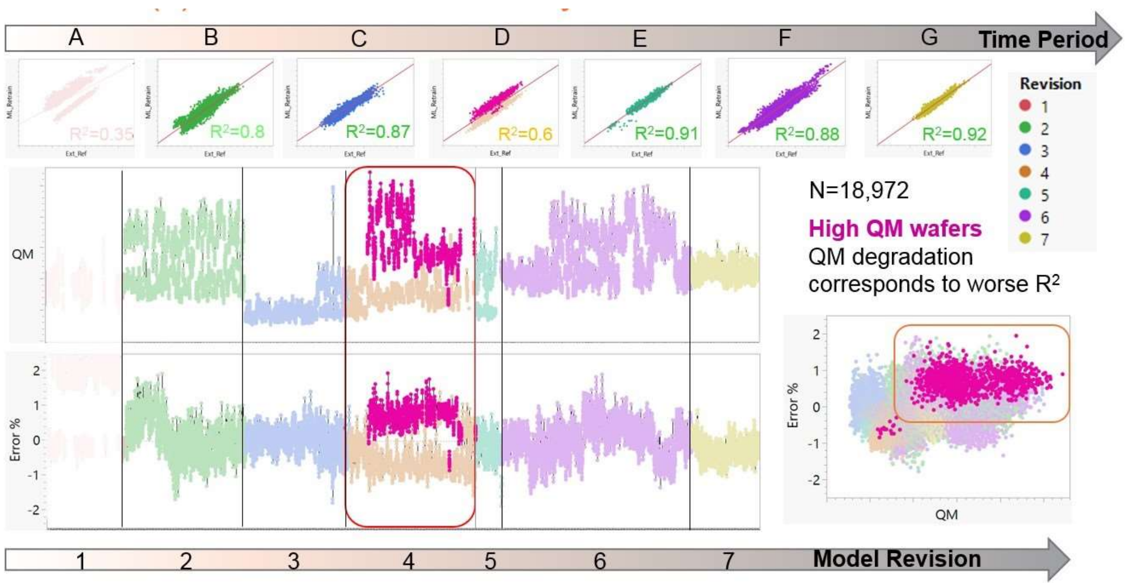


Figure 9. R^2, Error % and QM for thin film thickness prediction highlighting high QM wafers in time period D by Model 4.

Time Period	A	B	C	D	E	F	G
# Samples	2,193	3,145	2,771	3,213	510	5,202	1,938
Model 1	0.36	0.53	0.34	0.33	0.52	0.58	0.53
Model 2	0.90	0.80	0.84	0.67	0.76	0.88	0.94
Model 3	0.90	0.93	0.86	0.61	0.58	0.84	0.87
Model 4	0.93	0.90	0.89	0.56	0.65	0.86	0.92
Model 5	0.95	0.93	0.94	0.95	0.87	0.84	0.90
Model 6	0.94	0.91	0.93	0.87	0.95	0.87	0.93
Model 7	0.93	0.91	0.93	0.95	0.95	0.95	0.92

Time Period	A	B	C	D	E	F	G
# Samples	2,193	3,145	2,771	3,213	510	5,202	1,938
Model 1	1.55	1.74	1.01	0.90	0.68	1.45	0.56
Model 2	0.35	1.66	0.35	0.56	0.44	0.37	0.24
Model 3	0.30	0.34	0.27	0.73	0.63	0.45	0.42
Model 4	0.34	0.09	0.24	0.64	0.56	0.42	0.34
Model 5	0.22	0.11	0.19	0.26	0.28	0.47	0.34
Model 6	0.25	0.07	0.20	0.17	0.17	0.40	0.30
Model 7	0.25	0.05	0.20	0.17	0.17	0.23	0.31

Figure 10. R^2 (left) and absolute Error % (right) matrices for thin film thickness prediction.

3.3 Throughput Optimization from OCD Thin Film Measurement

Figure 11 summarizes the R^2, Error % and QM results from the ML prediction of an OCD thin film thickness measurement using a reduced incident angle spectral acquisition process as compared to the OCD model. The R^2 is noted to be already at 0.996 value with the first model. This ML prediction enabled about 40% throughput improvement.

Time Period	A	B	C	D
# Wafers	1,173	6,460	3,179	2,142
Model 1	0.996	0.974	0.97	0.988
Model 2	0.982	0.996	0.996	0.997
Model 3	0.981	0.987	0.997	0.997
Model 4	0.981	0.988	0.995	0.997

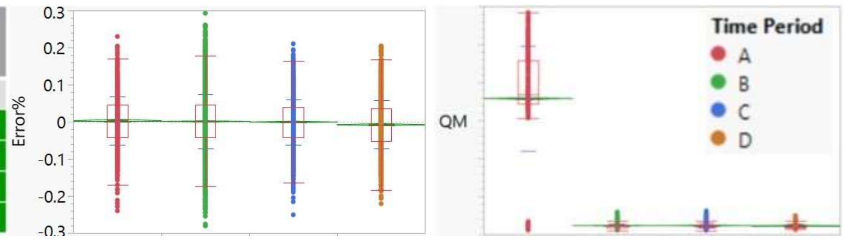


Figure 11. R^2 matrix, Error % and QM trend for thin film thickness measurement by improved throughput ML model.

No significant differences are evident in the Error % metric for the 4 blind test intervals. These ML models appeared to give adequate similar performance from the initial training. However, the QM is noted to be higher than normal for time period A by Model 1. This finding indicates that although QM may highlight cases where the ML model needs to be retrained, the R^2 and Error % data is also needed to make a full decision.

3.4 Throughput Optimization from OCD BEOL Etch Measurement

Figure 12 summarizes the R2, Error % and QM from a ML prediction of an OCD measurement of a dielectric etch depth parameter using a reduced incident angle spectral acquisition process as compared to the OCD model.

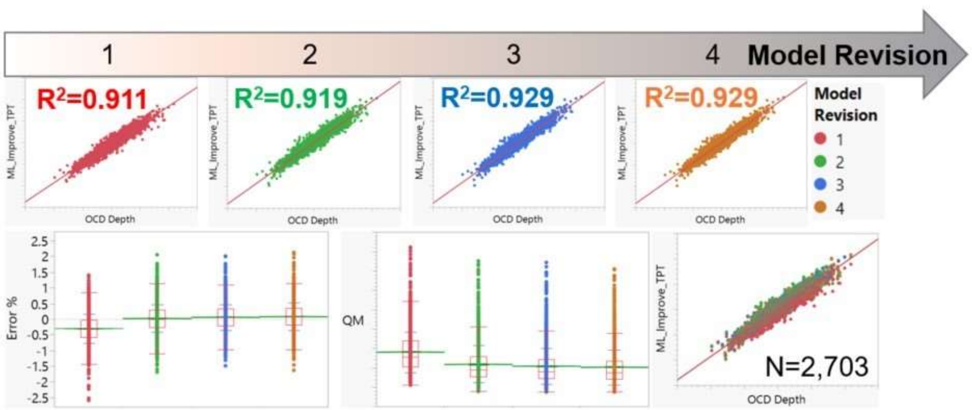


Figure 12. R^2, Error % and QM trend for thin film thickness measurement by throughput optimization ML model.

This ML prediction enabled about 40% throughput improvement. For these results, all 4 model revisions were tested on a common blind test time period after all 4 model training intervals to ensure a common sample size across 4 model revisions. A small R^2 increase was observed from Model 1 through 4 as well as a minor improvement of Error % (converges towards 0%) and gradual reduction in QM.

3.5 BEOL E-Test Prediction

Figure 13 shows the R^2 matrix obtained from 3 model revisions of predicted BEOL e-test by ML as compared to the actual e-test value. Although a surprising lower R^2 result is observed for model 3 applied to time period B samples, model revision 3 is shown to have optimal R^2 Error % and QM trend, particularly given the significantly larger sample size of time period C as compared to time period A and B. The results are in very good agreement with previously published work evaluating the impact of training size on the R^2 for applications such as BEOL e-test prediction [5].

Time Period	A	B	C
# Samples	368	551	9,864
Model 1	0.89	0.88	0.92
Model 2	0.94	0.88	0.93
Model 3	0.97	0.83	0.91

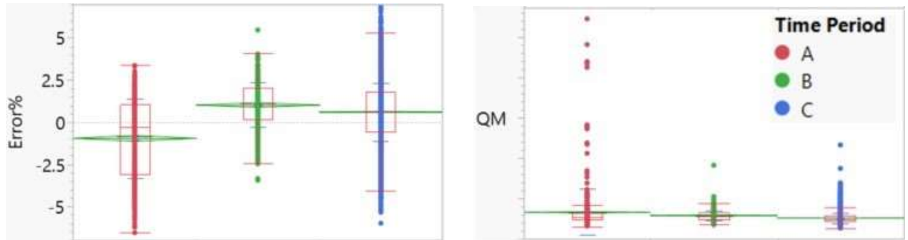


Figure 13. R^2 matrix, Error % and QM blind test trend for BEOL E-test Prediction model.

3.6 General Observation

In Figure 14, R^2 of 4 applications in this work are plotted on the Y-axis as a function of time along the X-axis. Although the maximum achievable R^2 and number of samples measured varied among these applications, it is interesting to observe that the fin CD and thin film thickness applications benefit more significantly from model retraining than the throughput optimization thin film measurement and the BEOL e-test prediction application.

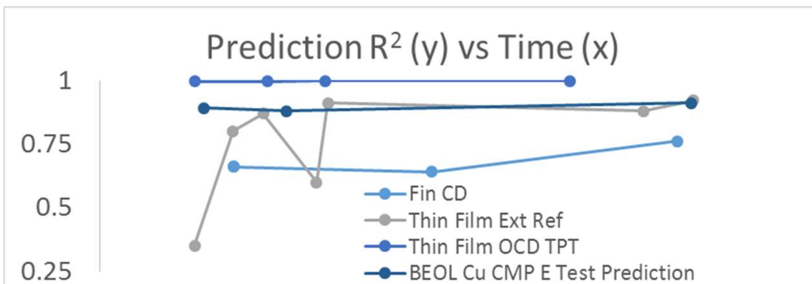


Figure 14. R^2 of the ML model (Y-axis) vs. time period (X-axis) for 4 applications studied in this work.

4. CONCLUSION AND FUTURE WORK

An advanced ML eco-system integrated with the OCD technique developed and implemented for fast, user friendly ML model creations to address HVM metrology requirements in semiconductor industry. This system allows for streamlined ML OCD model updates for high sensitivity and process development applications. Results from five applications over 6-8 month of period were demonstrated.

Two applications (fin CD and thin film thickness) showed benefits of model retraining to achieve best correlation R^2 and minimized Error % to the reference. The QM signal provided visibility of relatively low R^2 and high Error %, which is particularly valuable where the outcome and verification have long Mean Time to Detect (MTTD), a key performance indicator (KPI) for HVM management. Other applications had small or almost no correlation R^2 improvements but appear to be driven only by number of training samples. Those ML models provided adequate performance from the initial training: R^2, Error % and QM were close for all retrained models — we can say these ML models were saturated. Two throughput optimization applications showed up to 40% OCD throughput improvement compared to the reference OCD recipe, while keeping excellent correlations to the OCD reference.

In conclusion, it has been successfully demonstrated that with this advanced ML eco-system, streamlined ML models can be readily updated for multiple applications in HVM scenarios. Compared with the previous ML model creation procedure, this advanced ML eco-system has the advantage of being able to process very large amounts of data in an automated approach with significantly improved time-to-solution. The incremental retraining improved the correlation to reference for specific applications, multiple retrained models were analyzed to understand retraining effects, and QM was used to monitor recipe performance.

During this work, one-way incremental retraining strategy has been tried for different applications. Specific applications were evaluated using large amounts of data for analysis to identify appropriate ML solutions. This is a reasonable first step for a new application where the performance and details of each application were uncertain at the beginning. To spare time and resource further while keeping satisfactory performance and better automation. More efficient ML retraining strategies can be evaluated on different applications based on the available results and analyses to improve training efficiency while maintaining performance requirements. A final observation from this paper is that the exploration of OCD-based ML metrology continues to evolve and shows a very promising trend in the future semiconductor manufacturing metrology roadmap.

REFERENCES

[1] Orji, N.G., Badaroglu, M., Barnes, B.M., Beitia, C., Bunday, B.D., Celano, U., Kline, R. J., Neisser, M., Obeng, Y., Vladar1, A. E., "Metrology for the next generation of semiconductor devices," Nature electronics 1 (10), 532-547 (2018).

[2] Rana, N., Zhang, Y., Wall, D., Dirahoui, B., “Predictive data analytics and machine learning enabling metrology and process control for advanced node IC fabrication,” 26th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC) (2015).

[3] Breton, M., Chao, R., Muthinti, G., Peña, A., Simon J., Cepler, A., Sendelbach, M., Gaudiello, J., Tang, H., Emans, S., Shifrin, M., Etzioni, Y., Urenski, R., Lee, W., “Electrical test prediction using hybrid metrology and machine learning”, Proc. SPIE 10145 (2017).

[4] Kong, D., Motoyama, K., Huang, H., Mendoza, B., Breton, M., Muthinti, G. R., Shobha, H., Gaudiello, J., Cepler, A., Sendelbach, M., Emans, S., Wolfing, S., Ger, A., Shah, K., “Machine learning and hybrid metrology using scatterometry and LE-XRF to detect voids in copper lines,” SPIE 109590A (2019)

[5] Timoney, P., Kagalwala, T., Reis, E., Lazkani, H., Hurley, J., Liu, H., Kang, C., Isbester, P., Yellai, N., Shifrin, M. , Etzioni, Y., "Implementation of machine learning for high-volume manufacturing metrology challenges," Proc. SPIE 10585 (2018).

[6] Géron, A., [Hands-on Machine Learning with Scikit-Learn and TensorFlow], O'Reilly Media (2017).

[7] Halevy, A., Norvig, P., Pereira, F., "The Unreasonable Effectiveness of Data," IEEE Intelligent Systems, vol. 24, 812 (2009).