Electrical test prediction using hybrid metrology and machine learning

Mary Breton [a], Robin Chao [a], Gangadhara Raja Muthinti [a], Abraham A. de la Peña [a], Jacques Simon [a], Aron J. Cepler [b], Matthew Sendelbach [b], John Gaudiello [a], Hao Tang [a], Susan Emans [b], Michael Shifrin [c], Yoav Etzioni [c], Ronen Urenski [c], Wei Ti Lee [d]

[a] IBM, 257 Fuller Road, Albany, NY, 12203

[b] Nova Measuring Instruments, Inc., 2055 Gateway Place, Ste. 470, San Jose, CA 95110, USA

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

[d] ReVera, a Nova Company, 3090 Oakmead Village Dr., Santa Clara, CA 95051

ABSTRACT

Electrical test measurement in the back-end of line (BEOL) is crucial for wafer and die sorting as well as comparing intended process splits. Any in-line, nondestructive technique in the process flow to accurately predict these measurements can significantly improve mean-time-to-detect (MTTD) of defects and improve cycle times for yield and process learning. Measuring after BEOL metallization is commonly done for process control and learning, particularly with scatterometry (also called OCD (Optical Critical Dimension)), which can solve for multiple profile parameters such as metal line height or sidewall angle and does so within patterned regions. This gives scatterometry an advantage over inline microscopy-based techniques, which provide top-down information, since such techniques can be insensitive to sidewall variations hidden under the metal fill of the trench. But when faced with correlation to electrical test measurements that are specific to the BEOL processing, both techniques face the additional challenge of sampling. Microscopy-based techniques are sampling-limited by their small probe size, while scatterometry is traditionally limited (for microprocessors) to scribe targets that mimic device ground rules but are not necessarily designed to be electrically testable.

A solution to this sampling challenge lies in a fast reference-based machine learning capability that allows for OCD measurement directly of the electrically-testable structures, even when they are not OCD-compatible. By incorporating such direct OCD measurements, correlation to, and therefore prediction of, resistance of BEOL electrical test structures is significantly improved. Improvements in prediction capability for multiple types of in-die electrically-testable device structures is demonstrated. To further improve the quality of the prediction of the electrical resistance measurements, hybrid metrology using the OCD measurements as well as X-ray metrology (XRF) is used. Hybrid metrology is the practice of combining information from multiple sources in order to enable or improve the measurement of one or more critical parameters. Here, the XRF measurements are used to detect subtle changes in barrier layer composition and thickness that can have second-order effects on the electrical resistance of the test structures. By accounting for such effects with the aid of the X-ray-based measurements, further improvement in the OCD correlation to electrical test measurements is achieved. Using both types of solution—incorporation of fast reference-based machine learning on non-OCD-compatible test structures, and hybrid metrology combining OCD with XRF technology—improvement in BEOL cycle time learning could be accomplished through improved prediction capability.

Keywords: OCD, scatterometry, machine learning, electrical test, prediction, XRF, hybrid metrology, resistance

1. INTRODUCTION

There are many different metrology techniques that are used for process control of back-end of line (BEOL) structures. It is necessary to monitor many different parameters for metal lines, including the line height, critical dimension (CD), side-wall angle (SWA), and the width of the barrier materials. Scatterometry, also referred to as Optical Critical Dimension (OCD), is a model-based technique which is capable of high throughput, non-destructive measurements. The spot size can

Metrology, Inspection, and Process Control for Microlithography XXXI, edited by Martha I. Sanchez, Vladimir A. Ukraintsev
Proc. of SPIE Vol. 10145, 1014504 · © 2017 SPIE · CCC code: 0277-786X/17/$18 · doi: 10.1117/12.2261091

capture hundreds of metal lines in a single measurement. Its other benefits include the ability to extract measurements from buried structures and to measure other parameters that cannot be captured by a top-down image-based measurement technique, like the line height mentioned above. Critical-Dimension Scanning Electron Microscopes (CD-SEM) can be used to acquire a top-down image, but these usually capture only a few metal lines. Due to the pixel sizes that are often employed to take advantage of the small spot size, large area images are too time and resource intensive to regularly acquire. Imaging and measuring metal structures can also be challenging, as the secondary electron yield and interaction energy is so great that edges are more obscured.

X-ray techniques are useful methods of obtaining compositional and thickness information from materials. In this work, we use Low Energy X-Ray Fluorescence (LE-XRF). In XRF, incoming x-rays generate electron-hole pairs, and x-rays with characteristic energies are emitted when different electrons fill that hole. Each elemental species has a characteristic set of x-ray energies that may be emitted, so we obtain both the species information and a proportional "count" of different species within the exposed region (spot) of the sample, which has a size similar to the OCD spot. Electrical tests directly measure the device performance by probing pads attached to structures that are designed to mimic portions of the chip circuitry or detect process defects. These are high throughput tools, but they often have such dense sampling that the number of wafers per hour (WPH) that are measured can be surprisingly low. Additionally, as the probes directly contact the wafer, there are risks of localized damage to the pads which may change subsequent readouts.

A solution to the dependence on inline electrical tests (ILT) lies in a fast reference-based machine learning capability that allows for OCD measurement directly on the electrically-testable structures, even when they are not OCD-friendly (for example, in cases where the gratings are on top of extremely complex structures). Incorporating such direct OCD measurements on ILT structures significantly improves the correlation to and prediction of resistance of BEOL electrical test structures. To further improve the quality of the prediction to the electrical resistance measurements, hybrid metrology using the OCD measurements as well as X-ray metrology (XRF) is used. Hybrid metrology is the practice of combining information from multiple sources in order to enable or improve the measurement of one or more critical parameters[1,2]. Here, the XRF measurements are used to provide a second source of information about how much copper is present, as well as detect changes in barrier layer thickness that can have second-order effects on the electrical resistance of the test structures. By accounting for such effects with the aid of the XRF, further improvement in the OCD correlation to electrical test measurements is achieved. Using both types of solution—incorporation of fast reference-based machine learning on non-OCD-compatible test structures, and hybrid metrology combining OCD with XRF technology—improvement in ILT prediction capability is demonstrated.

2. DESIGN OF EXPERIMENT

2.1 Description of test structures

The selected test structures include both a "device" (i.e. electrically testable) structure, as shown in Figure 1, and a standard OCD pad. The device structure selected was from a macro family that is large enough for the OCD and x-ray spot sizes and is commonly used for electrical testing. The standard OCD pad is the same size but intentionally designed with no other patterning underneath, so as to minimize noise from structures that are not of interest at the level. As indicated in Table 1, the metrology array for this BEOL level did not share the same normalized trench width and pitch with the device structure; this was to enable comparison of two different macro designs, as it is common to have the metrology and device macros designed without the other in mind.

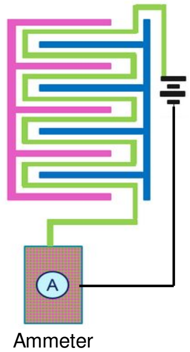

Figure 1: Schematic of device test structure.

Table 1: Test Structure Details

Array	Normalized Linewidth	Normalized Pitch
OCD	0.34	0.75
Device	0.59	1

Both test structures share the same general configuration[3], depicted in Figure 2. Trenches etched into a dielectric are lined with TaN and Co, respectively. Then copper is deposited, filling the remaining cavity, and the whole structure is polished to the desired height. After polishing, the copper is selectively-capped with Co and some films to prevent further oxidation.

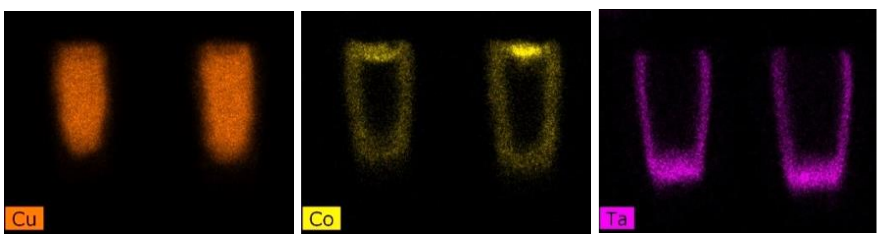

Figure 2: Schematic of test structures (images from reference 3).

2.2 Wafer splits to induce variation

Several different parameters were varied in order to produce a robust design of experiment (DOE). The lithography was done according to the same random dose across each wafer, in order to vary the linewidths of the trenches. Five different exposures were used, wider and narrower than the process of record (POR) condition, (POR, --, -, +, ++). Four different liner thicknesses were used (POR, -, +, ++) to vary the copper CD within the trench, and three different CMP times were used in order to vary the heights of the copper (POR, + 5 seconds, - 5 seconds).

2.3 Sampling

Due to capacity and throughput for these toolsets, the number of die measured varied by toolset. The entire wafer is measured with OCD, while XRF is used to measure 10 die per wafer. The electrical measurements are obtained over the entire wafer as well. OCD spectra were collected from both test structures, while electrical and XRF measurements were only completed on the device structure. All data were collected post-cap.

3. RESULTS AND ANALYSIS

3.1 Electrical treatment of test structures

The analysis presented in sections 3.2 and 3.3 rely upon the following treatment of the post-CMP test structures. For a single, uniform metal wire, the resistance can be calculated initially as:

$$R = \frac{\rho L}{A} \tag{1}$$

For each material, R is the resistance of the metal line, ρ is the resistivity, L is the total line length of the metal line, and A is the cross-sectional area of the line. The length of the metal line is approximated well with the design data, and should include all relevant connections between the array in question and the pads where the tester pins will make contact. Resistivity, generally a function of cross-sectional area, can be determined experimentally and refined by comparing results to literature models. In this work, IBM models for resistivity were employed.

Given Cu is a very low resistivity metal, it is the primary conductor in the Figure 2 configuration. Thus, a simplified relationship from Eq (1) can be used, which will be referred to as Model #1 [Eq (2)].

Model #1

$$R \propto \frac{1}{A} \tag{2}$$

In the future analysis of the structure presented in Figure 2, a key assumption is that each of the three metal components – TaN, Co, and Cu – can be treated as single resistors, in a parallel configuration. Thus, the total resistance would be:

$$\frac{1}{R_{Total}} = \frac{1}{R_{Cu}} + \frac{1}{R_{Co}} + \frac{1}{R_{TaN}} \tag{3}$$

However, given $\rho_{Ta} \gg \rho_{Co} > \rho_{Cu}$, we simplify Eq (3) to:

Model #2

$$\frac{1}{R_{Total}} = \frac{1}{R_{Cu}} + \frac{1}{R_{Co}} \tag{4}$$

3.2 Preface to Results

In the following charts, the actual ILT-measured device structure resistance will be shown on the Y-axis data. The results will progress from analysis of ILT correlations to single metrology techniques, then to hybrid metrology, and finally to hybrid metrology with machine learning. The chart titles are colored according to the source of the data on the X-axis, and the scheme is depicted in Figure 3.

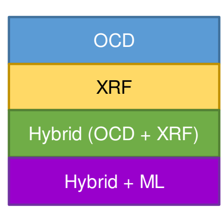

Figure 3: Results legend.

3.2 OCD-only solution

3.2a Model #1 – Cu Area

Model #1 is the basis for the initial method of predicting electrical yield of a structure. In this set of analysis, an OCD model is used to find the cross-sectional area of the copper line. A model was only built for the OCD pad (Figure 4), from which the cross-sectional area was found. For the device structure, the area was scaled based on the known design parameters.

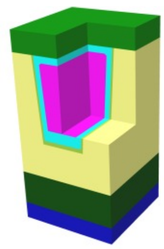

Figure 4: OCD model.

Spectra were collected on the OCD pad, as well as the device pad. Spectra from the OCD pad are interpreted based on the OCD model, allowing us to calculate the cross-sectional area of the Cu line and the liner materials. These results are scaled because the device pad has different intended dimensions than the OCD pad. The scaling speeds up the time to solution as it reduces the modeling setup time, and the same trends can be expected across structures and when compared to the ILT results. Figure 5 demonstrates the predicted metal line resistance from this initial treatment to the actual structural resistance as measured by inline test. With an R^2 of 0.5, some general trends may be able to be detected, but finer process control would not be possible only using these provided and measured values.

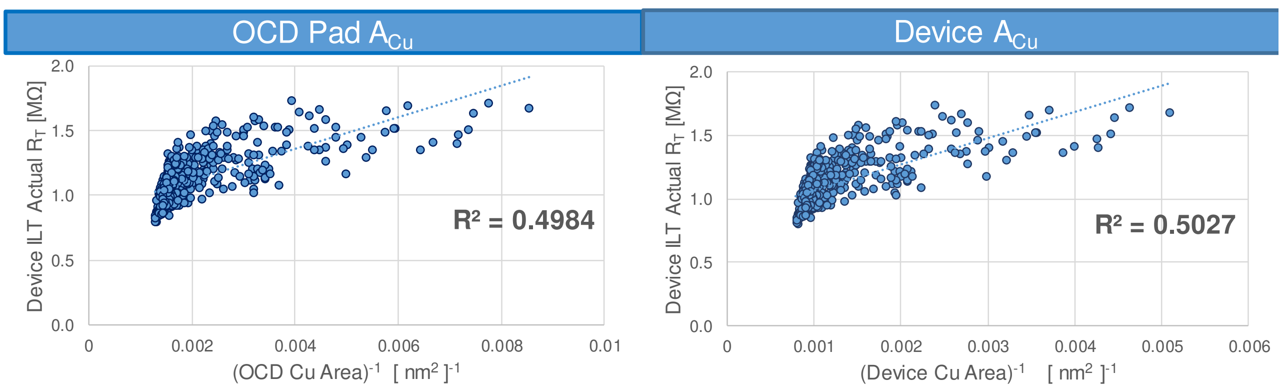

Figure 5: Comparing cross-sectional area of Cu to measured resistance.

In Figure 6, we take the calculated cross-sectional area of both the copper line and the cobalt liner and use this information to calculate the expected resistance, based on equation (4). We again see evidence of a trend, but with R^2 around 0.44, more improvement is needed.

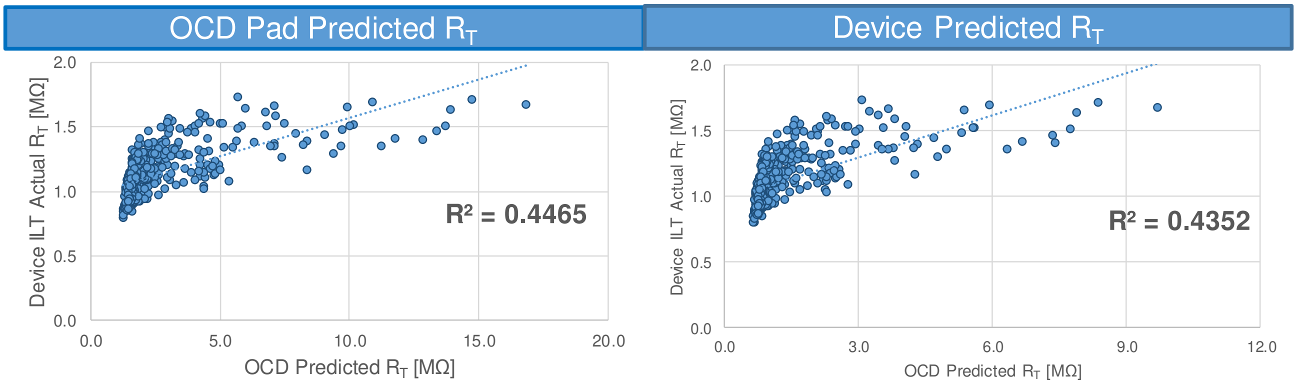

Figure 6: Calculating resistance based on OCD's measured cross sectional area of Co liner and Cu and comparing to measured resistance.

3.3 XRF vs eTest

Figure 7 shows the correlation between the XRF measurement of Cu and Co quantity on the device structure to the measured ILT resistance from the same structure. The purpose of investigating these correlations is to confirm that there is, in fact, information in the XRF signal that potentially can be used to further improve predictions of ILT resistance.

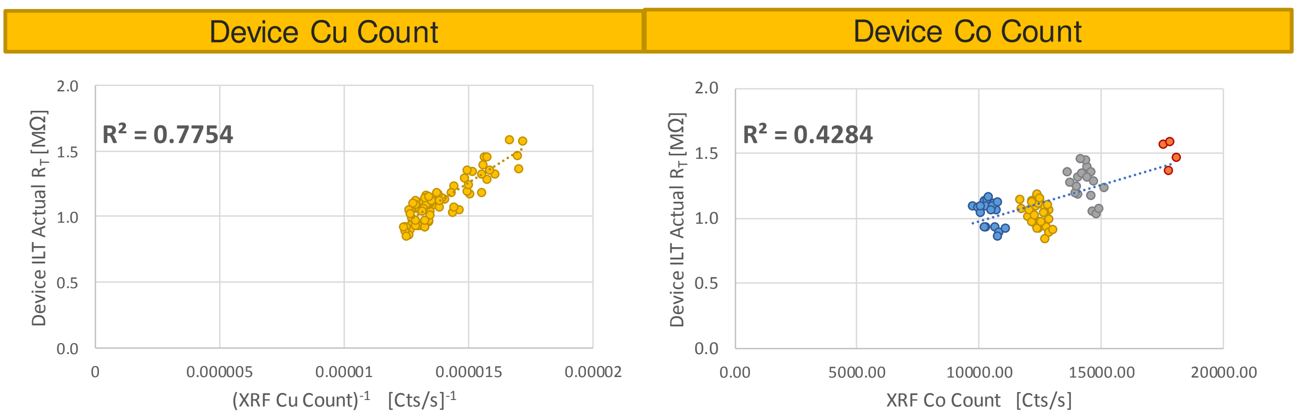

Figure 7: Comparing XRF data to measured resistance. For the Co count, the different Co liner thickness splits are shown in different colors. The presence of a nominally-constant Co cap on top of the Cu line is adding some noise to the Co measurement, likely reducing the quality of the correlation.

3.4 Hybrid Prediction vs eTest

A hybrid scheme is explored in order to combine inputs from both the OCD and XRF and compare that to the ILT. In previous work[2], this hybrid configuration has been shown to improve sensitivity to structural parameters. As compared to the OCD-only results (using only the die that are also measured with XRF, as opposed to the results shown in Figure 6 which use all die on the wafer), the R^2 improves considerably, showing a large benefit to the hybridization, as shown in Figure 8.

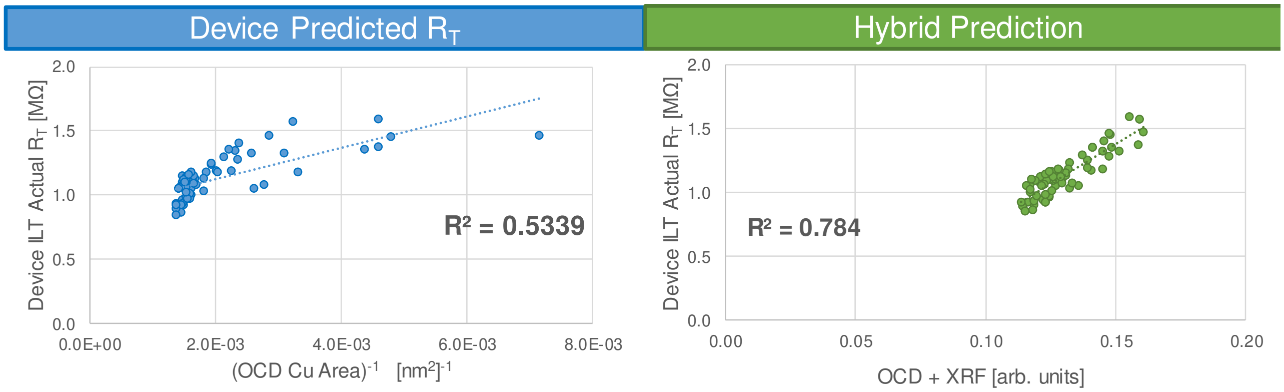

Figure 8: Hybrid Predicted Resistance (on right) compared to ILT. On the left are the results without the use of XRF. In both sets of data, the same die and the same measurement target (device) were used.

3.5 Machine Learning Results

To continue to improve the results described, Nova's machine learning capability was employed. The results are shown in Figure 9. This results in another large improvement in R^2, showing that this technique of combining hybrid metrology (OCD and XRF) with machine learning is effective in predicting resistance. Such predictive capability can enable a reliable dynamic ILT sampling. For example, for those wafers or lots where the hybrid + machine learning results are passing, ILT sampling can be reduced, but for those results that are failing, the ILT sampling can be increased in order to take a "closer look"—or, the wafers can be scrapped to avoid unnecessary processing. Although for this work the ILT, OCD, and XRF measurements were done at the same point in the process, making dynamic ILT sampling less appealing, the concept can be applied to other applications where there is significant processing between the OCD/XRF measurements and the ILT measurements.

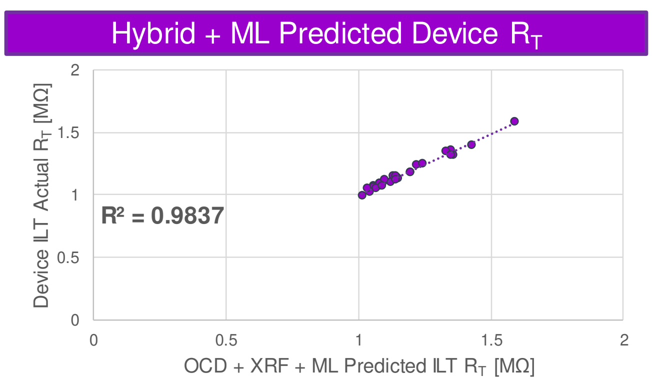

Figure 9: Hybrid + Machine Learning Predicted Resistance compared to ILT. The same sampling was used here as was used for the data in figure 8.

4. CONCLUSIONS

It is an important goal to predict electrical test results using in-line metrology. Initial results from the OCD model show insufficient match to the electrical test data, while the combination of OCD and XRF data (hybrid metrology) show much better performance. Combining data from OCD and XRF with a machine learning technique developed by Nova is shown to have excellent predictive capability of the ILT data. Such results can enable further improvement of yield learning, process control, and early detection of problematic wafers/lots, as well as the implementation of a more dynamic sampling for electrical test measurements.

ACKNOWLEDGEMENTS

The authors would like to thank Roger Quon, Nicholas Lanzillo, Prasad Bhosale, Han You, Theo Standaert, Ted Levin, Kee Won, Swathi Kamat, Kyle Warnacke, Jeffrey Laprade, Michael Callahan from IBM, and Dana Bernstein from Nova.

REFERENCES

[1] Vaid, A., Yan, B. B., Jiang, Y. T., Kelling, M., Hartig, C., Allgair, J., Ebersbach, P., Sendelbach, M., Rana, N., Katnani, A., Mclellan, E., Archie, C., Bozdog, C., Kim, H., Sendler, M., Ng, S., Sherman, B., Brill, B., Turovets, I., and Urensky, R., "A holistic metrology approach: hybrid metrology utilizing scatterometry, CD-AFM, and CD-SEM," Proc. SPIE **7971**, 797103 (2011).

[2] L'Herron, B., Chao, R., Kim, K., Lee, W. T., Motoyama, K., et al. "Hybridization of XRF/XPS and scatterometry for Cu CMP process control," Proc. SPIE 9424, 94241M (2015).

[3] Standaert, T., Beique, G., Chen, H.-C., Chen, S.-T., et al., "BEOL Process Integration for the 7nm Technology Node," Proc. IEEE International Interconnect Technology Conference / Advanced Metallization Conference (IITC/AMC), (2016); doi:10.1109/IITC-AMC.2016.7507636