Simulation study of repeatability and bias in the critical dimension scanning electron microscope

John S. Villarrubia
András E. Vladár
Michael T. Postek
National Institute of Standards and Technology
Gaithersburg, Maryland 20899

Abstract. The ability of a critical dimension scanning electron microscope (CD-SEM) to resolve differences in the widths of two lines depends on the instrument's measurement repeatability and any sample-dependent biases. The dependence of repeatability and bias on eight different parameters is studied using the MONSEL Monte Carlo electron simulator. For each of 14,400 different combinations of values of eight parameters, three describing the sample and five describing characteristics of the instrument or measurement condition, an image is simulated, noise is added, and the edge positions are "measured" as would be done in a CD-SEM. From 100 repetitions of noise, the repeatability of such CD determinations is ascertained. Biases (i.e., average errors) are also determined. Noise amplitude, edge detection algorithm, and beam size are shown to be significant factors in measurement repeatability. The CD-SEM's measurement repeatability may be an order of magnitude better than its spatial resolution. For standard edge detection methods, the bias depends on the sample. This means that in a manufacturing environment in which the sample shape varies, there will be a random component of error that is not measured by the industry's usual same-sample tests of instrument precision. [DOI: 10.1117/1.2037447]

Subject terms: critical dimension metrology; model-based library; linewidth metrology; precision; repeatability; resolution; scanning electron microscopy.

Paper 05004 received Jan. 10, 2005; accepted for publication Mar. 7, 2005; published online Sep. 19, 2005. This paper is a revision of a paper presented at the SPIE Conference on Metrology, Inspection, and Process Control for Microlithography XVII, Feb. 2003, Santa Clara, California. The paper presented there appears (unrefereed) in SPIE Proceedings Vol. 5038.

1 Introduction

It is frequently necessary to compare widths of different features, for example chip to chip or wafer to wafer differences in transistor gate width (the critical dimension or CD). Since great efforts are made to keep manufacturing process variation to a minimum, measured differences are often small. The question of whether a particular such difference is real or a measurement artifact frequently arises, and the answer may provide the justification for a decision to alter process variables to keep the product dimensions acceptably close to their target values. The International Technology Roadmap for Semiconductors (ITRS)[1] calls "the ability of a CD measurement tool to distinguish between lines that differ in width" by the name "microscopy resolution." This ability is also sometimes called the linewidth or CD "measurement resolution," a term we prefer. We use the two terms interchangeably here.

According to the most recent ITRS, the "in-line nondestructive microscopy resolution" requirement for critical dimension (CD) measurements in 2005 is 0.29 nm, decreasing to 0.2 nm by 2008. The ITRS closely associates microscopy resolution with measurement "precision," that is, three standard deviation repeatability. Repeatability is a measure of random error associated with a measurement, the amount of spread in the distribution of measured values when a measurement on the same sample is repeated many times. The results of this study call attention to a dependence that may be less widely appreciated: the relationship between measurement *bias* and measurement resolution. Bias in a width measurement is the difference between the average measured width and the true width. The difference $w_1 - w_2$ is not affected by a bias that changes w_1 and w_2 by the same amount, but it is affected by any sample-dependent bias that affects the two widths differently.

The ITRS shows all microscopy resolution requirements (tables begin with the year 2003) in red, indicating that manufacturable solutions are not known. The question therefore naturally arises: Is microscopy resolution fundamentally limited near current levels by the physics of the critical dimension scanning electron microscope (CD-SEM) technology that is employed for these measurements? Or is current performance the result of less than optimal use of a technology that is fundamentally still good enough? Approaching fundamental physical limits to the spatial resolution of the SEM[2] are taken by some as support for the first point of view. But is spatial resolution the same as microscopy resolution? What are the other, if any, significant variables on which microscopy resolution depends?

To these questions, a simulation study has much to contribute. In such a study images are simulated for a variety

of sample shapes, instrument conditions, and noise amplitudes. The edges of these images are then located using various edge-finding algorithms. The reproducibility of repeated measurements (with varying noise) can be ascertained from such results. Unlike the situation experimentally, it is possible to hold all but one variable truly constant, thereby truly isolating the contributions of individual variables. In a simulation, unlike an experiment, one knows the true edge positions. It is therefore possible to quantify the bias and observe whether it changes as a function of sample shape or other relevant variables.

In Sec. 2 we describe the method of our simulation study. Section 3 gives the results, followed by discussion and conclusions in the sections that follow.

2 Study Approach

2.1 *Overview*

We identify measurement resolution for a linewidth measurement with the uncertainty in the measured difference $w_1 - w_2$. In the appendix in Sec. 5, under assumptions that are described there, we derive the following expression for the measurement resolution R:

$$R^2 = 2P^2 + 2\left(\frac{\partial w}{\partial \theta} U_\theta\right)^2 + \ldots. \tag{1}$$

Here P is the precision, $\partial w/\partial\theta$ measures the sensitivity of the width measurement to a parameter θ (e.g., sidewall angle), and U_θ is the 3 standard deviation uncertainty in θ for the two lines being measured.

Note that there are two types of terms that determine the measurement resolution. The first is a random component associated with the repeatability of measurements made on the same sample, the second with any measurement bias that differs between the measurements on the two samples.

Both of these types of components of measurement resolution were observed and measured in our simulation experiment. We first provide an overview of our procedure in this paragraph and in Fig. 1. This is followed by a more detailed explanation of the individual steps. For each iteration of the outermost loop, the inputs were a set of instrument, sample, noise, and data analysis parameters. Software that simulates electron trajectories and secondary electron generation was used to generate a simulated noiseless linescan. In the interior (noise) loop, simulated CD-SEM noise is generated and added to the linescan. The resulting noisy image is then measured by a procedure that depends on the initial choice of analysis parameters. This produces one width value for each iteration of the loop. After many iterations, the repeatability and bias are determined from the measured width values and the actual width. The outer loop is then repeated for other choices of input parameters. The result is a dataset showing measurement bias and repeatability under a variety of measurement conditions.

2.2 *Sample and Instrument Parameters*

The first step in Fig. 1 includes choice of noise and analysis parameters, but we postpone a description of these until Secs. 2.4 and 2.5, respectively. We discuss here those inputs that specify the sample and characteristics of the SEM. To

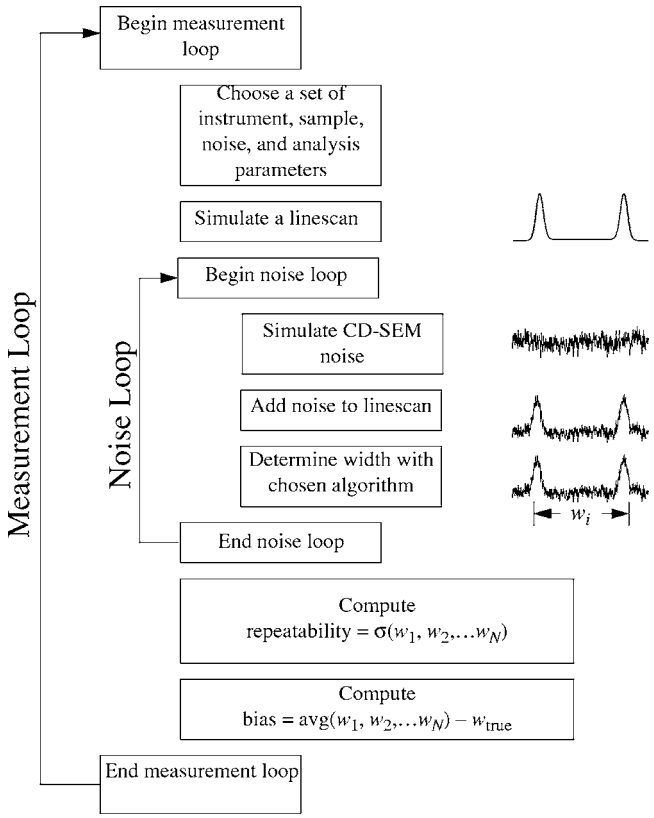


Fig. 1 Schematic of the design of the study.

make the simulation as relevant as possible to actual industrial processes, we modeled many of our sample and instrument parameters on measurements made for us on a CD-SEM at International SEMATECH. For those measurements, the samples consisted of PAR 810 UV resist lines on an antireflective layer (Brewer Science ARC 25) on silicon, as shown in Fig. 2(a). (Certain commercial equipment or materials are identified in this work to describe the experimental and analytical procedures adequately. Such identification does not imply recommendation or endorsement by NIST, nor does it imply that the items identified are necessarily the best available for the purpose.) The electron trajectory simulation used material properties (e.g., density and elemental composition) consistent with these samples. The ARC layer thickness t and the line height h were also fixed at values derived from this sample ($t=86$ nm, $h=270$ nm). The sidewall angle θ, corner radius r, and line separation s were permitted to vary as indicated in Table 1.

Fixed parameters related to the instrument settings included the electron beam energy of 800 eV, pixel size of 1 nm, and best-focus position at the top of the line. We performed simulations under two secondary electron collection scenarios. In one, secondary electrons were assumed to travel in straight lines as they leave the sample. Some of these trajectories intersect the substrate or adjacent lines and re-enter the sample. In the other, secondary electrons were assumed to be collected as soon as they emerged from the sample. Although both scenarios were calculated, in this study consideration is restricted only to results of the sec-

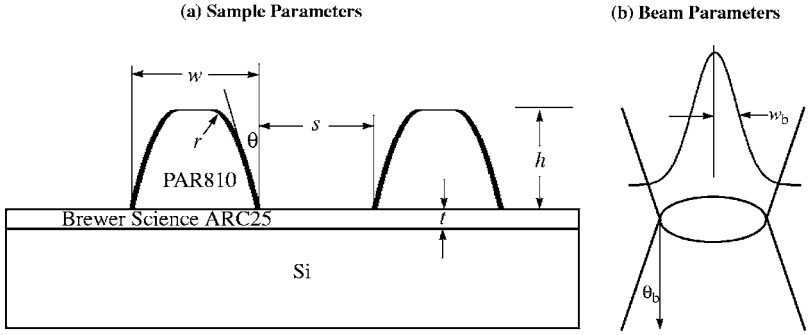


Fig. 2 Schematic of sample and beam parameters. (a) Simulated sample: resist lines on antireflective coating on a silicon substrate. (b) Model of focused electron beam: the hourglass shape (bold lines) with cone angle θ_b indicates how the electrons first converge, then diverge from a minimum spot of size w_b.

ond scenario because we deem it more relevant for a CD-SEM, which normally operates with an extraction field for efficient collection of the relatively low energy secondary electrons. In addition to these fixed parameters, there were two variable parameters associated with the instrument. These are defined in Fig. 2(b), with values listed in Table 1. One of these was the beam size w_b, the standard deviation of the assumed normal distribution of electrons within the landing spot. The second, θ_b, is the half angle of the assumed conical distribution of incoming electron trajectories. This angular distribution accounts for the microscope's depth of focus. When $\theta_b=0°$, the incoming electron trajectories are parallel and the depth of focus infinite. The other two values of θ_b listed in the table correspond respectively to spot sizes 10 nm and 20 nm larger at the bottom of the line than at the top.

Table 1 Variable input parameters. As indicated in Fig. 2, the sidewall angles represent the deviation of the sidewall from the vertical. Negative angles represent undercuts.

Parameter description	Symbol in Fig. 2	Number of values	Values
Sidewall angle	θ	8	(–3, –1, 0, 1, 3, 5, 7, 9)°
Corner radius	r	3	(0, 10, 20) nm
Line separation	s	2	Isolated (5 mm) and dense (200 nm)
Beam half angle	θ_b	3	(0, 1.06, 2.12)°
Beam half width	w_b	5	(0, 1, 3, 7, 15) nm
rms noise amplitude	—	5	0.01, 0.025, 0.063, 0.159, 0.40 in arbitrary intensity units (see Fig. 5)
Power spectrum (PSD)	—	2	Flat, sloped (see discussion)
Edge detection method	—	4	Maximum derivative, regression to baseline, sigmoidal fit, model-based library

2.3 *Simulating Linescans*

Given the inputs as just described, an image linescan is calculated using MONSEL,[3,4] a Monte Carlo electron trajectory simulator. MONSEL uses Browning's interpolation formula[5] for the Mott elastic scattering cross section and continuous energy losses as specified by the Bethe formula modified[6] for improved accuracy in the low energy limit. Secondary electron generation via the plasmon[7] and Möller[8] (also called Mott-Williams[9] in the nonrelativistic limit that we use) mechanisms are included. Individual incident electrons and all of their offspring (due to secondary electron generation) are followed through elastic and inelastic scattering events until they either escape the sample or their energies fall below the sample's work function, making escape impossible. For each electron beam landing position, 10,000 incident electrons were followed for the simulations in this study. A graphical representation of MONSEL trajectories for the first 200 of these is shown in Fig. 3 for an electron beam landing position near the line's edge. The dense (in this case approximately 10 nm wide) vertical column above the top of the line is mainly composed of electrons in the incident beam. These electrons scatter within the sample, forming the roughly 30 nm diameter active volume inside the line. When this active volume makes contact with the sample surface, backscattered and secondary electrons emerge. These are visible as the long straight trajectories at varying angles in the vacuum region outside of the line. The characteristic bright line edges in a secondary electron image are due to the additional secondary electrons that escape the sample when the sidewall is within the active volume. The total number of secondary electrons emitted from the sample at each beam position determines the intensity at the corresponding point on the linescan. To produce each simulated linescan, a calculation like this one was performed at 201 landing positions spaced 1 nm apart.

The relevance of a simulation study like this one depends on the degree to which the simulated instrument is a good approximation of the real one. Experimental validation of models of secondary electron imaging are in their

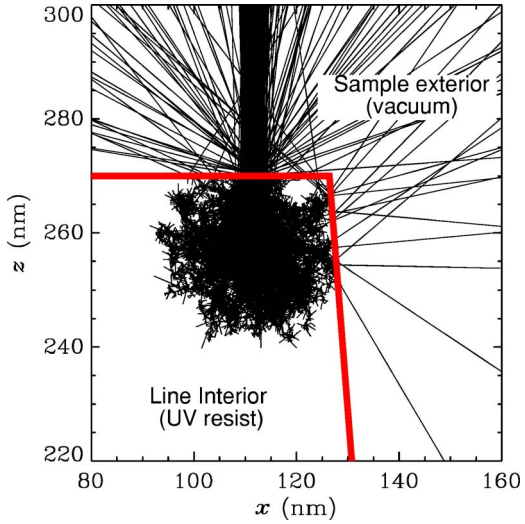


Fig. 3 Electron trajectories near the edge of a UV resist line as simulated by MONSEL.

early stages. However, the early results[10,11] are promising. Figure 4 shows a good match, apart from noise, between a simulated linescan and the experimentally measured one. The line shape that gave rise to this simulated linescan was observed in the same experiment to agree well with a cross-section of the measured line.[11]

2.4 *Simulating and Adding Noise*

After production of the linescan, the next step in Fig. 1 is the noise loop. Simulated noise was generated using software that uses a random number generator to produce noise profiles that have a power spectral density (PSD) that matches a desired one. To produce realistic simulated noise, International SEMATECH provided us with images of resist samples measured by CD-SEMs from three different manufacturers. We measured the power spectra of the image intensity in nominally featureless areas of these images. Even when normalized to the same root mean square (rms) total noise level, there were differences in the power spectra of the different tools. One had a relatively flat power spectrum up to a high frequency cutoff. Others had more noise at low frequency and correspondingly less at higher frequencies. We chose two of these power spectra, the flattest and the least flat, to represent the range of possibilities. The choice of which of these to use is labeled the "power spectrum" variable in Table 1. This variable determines the frequency distribution of the noise, but because the power spectra are normalized to the same total power, this variable has no effect on the total rms noise level. To account for different noise levels, a second variable, termed noise amplitude in Table 1, was used. After random noise with the chosen PSD is generated, it is multiplied by the chosen noise amplitude and added to the linescan determined in the last step. The range of noise amplitudes employed in our simulations is illustrated in Fig. 5. The top profile corresponds roughly to the amount of noise in a single unaveraged noisy linescan. The bottom profile corresponds to what one might expect with significant frame averaging in the CD-SEM. There are three intermediate levels as shown.

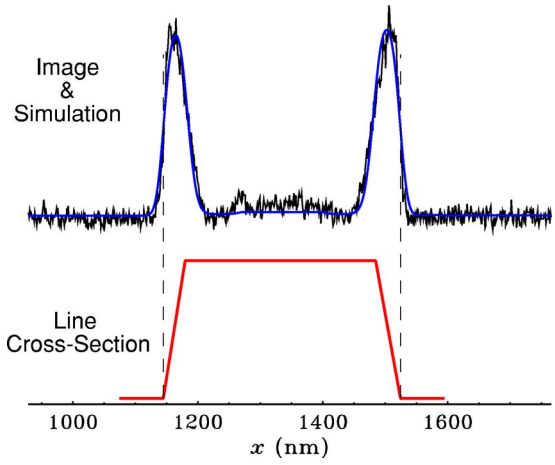


Fig. 4 Match between simulated and measured linescans. The measured linescan is the noisier of the two upper curves. The simulation is the smoother one. The line shape that produces the simulated linescan is also shown (bottom).

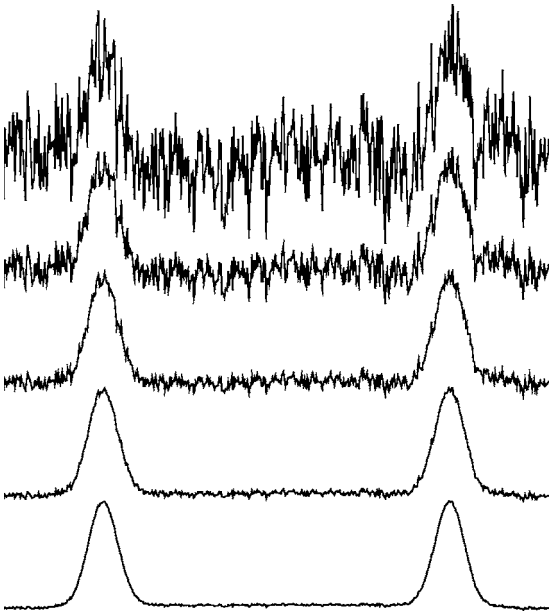


Fig. 5 Range of noise amplitudes used in the simulations.

2.5 *Determining the Edge Positions*

As is evident in the Figs. 4 and 5 linescans, the bright image peaks produced by the line's edges have a finite width. This width is large compared to the desired measurement resolution, so the measured value of the linewidth depends, in an important way, on how within this peak the edge is assigned. Commercial CD-SEMs may make this assignment in a number of ways. Instruments generally offer the metrology engineer a choice of algorithm. The principles by which the methods operate are generally known, but the details are proprietary. Accordingly, we developed our own software to implement four methods: maximum derivative, regression to baseline, sigmoidal fit, and model-based library (MBL).

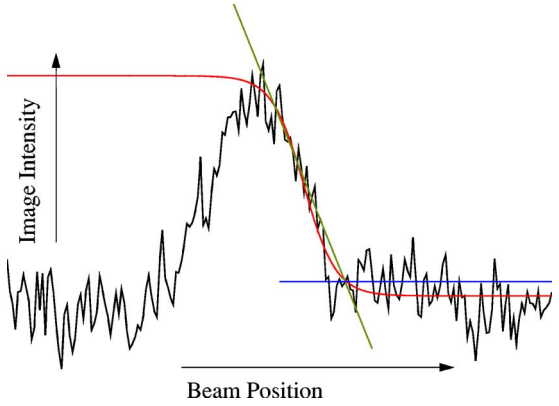


Fig. 6 Various methods for edge assignment. The sloped line is tangent to the outside edge of the intensity peak at the position of the maximum slope. The intersection of this line with the horizontal baseline is the position assigned by the regression to baseline method. The sigmoidal fit method fits an s-shaped curve through the data, and takes the edge position from a parameter of this fit.

The first three of these are based on the same principles as methods with similar names available on commercial CD-SEMs. These principles are illustrated in Fig. 6. The noisy peaked curve is a linescan centered on the right edge of a line. The maximum derivative method assigns the edge position on the point on the outside half of the peak (i.e., the half farthest from the center of the line) where the magnitude of the slope is greatest. Since derivatives tend to exaggerate noise, in practice algorithms smooth the data before differentiation. Generally, the more smoothing, the better the repeatability; but too large a smoothing window distorts the peak shape and changes the measurement bias. This leads to an important choice for this study. Some of the variables in our study affect the width of the peak that is being fit. For example, an edge with a vertical sidewall exhibits a narrower intensity peak than does an edge with a sloped sidewall. We can choose either to smooth all images with smoothing windows of the same size or to smooth each image with the largest window that does not significantly distort the peak shape. The former choice results in very little smoothing because it must treat all edges the same as the narrowest ones. We believe that a metrology engineer would more likely choose an amount of smoothing based on the characteristics of the samples produced by the manufacturing process, with more smoothing when the images are amenable to it. Therefore, this is the choice that we made.

If we draw a line tangent to the edge at the maximum derivative point (the straight, sloped line in Fig. 6) and draw a horizontal line at the background level (the baseline, also shown), their intersection is the position of the edge as assigned by the regression to baseline method.

The s-shaped (or sigmoid) curve in the figure is a function of the form

$$y = a + \frac{b-a}{1+\exp[-c(x-x_0)]}, \tag{2}$$

where a, b, c, and x_0 are parameters. This function is fit to the outer edge of the peak by least squares, and the best value of x_0, which corresponds to an intensity level midway between the sigmoid's maximum and minimum values, is assigned as the edge location. Smoothing is implicit in the fitting process, so this method should not benefit much from smoothing the data. Accordingly, we smoothed much less to avoid changing the peak shape. On the other hand, there are other choices that must be made. In Fig. 6, it is obvious that the sigmoid curve does not even approximately fit the data to the left of the peak. To obtain a good fit on the right, the data on the left must therefore not be included in the fit. In our implementation, the fit window includes the point of maximum intensity and beyond. The small amount of smoothing that we employ allows several data points near the maximum to be averaged for determining the maximum intensity level and its position. The absence in the data of the sigmoid's plateau on the left may, because of noise, result in some spurious fits in which the best sigmoid plateaus well above the intensity maximum in the data. To prevent this, the value of a is fixed at the maximum intensity, and the least squares procedure is permitted to vary only the other three parameters in Eq. (2).

The MBL method is an experimental method that was developed relatively recently. In this method, a range of edge shapes is simulated in advance using an electron trajectory simulator like MONSEL. These shapes and their images form a database, or library. Measured images are compared to the library, which may be interpolated, to find the parameters, including edge position, that produce the best match. Like the sigmoidal fit, smoothing is implicit in the fitting process, so images were not otherwise smoothed. Unlike the sigmoidal fit, the library entries are peaks; instead of starting the fitting window at mid-peak, all of the data near the edge can be included in the fit. More details are available elsewhere.[11–13]

2.6 *Computing Repeatability and Bias*

After noise was added to the simulated linescan and the edge positions were determined by the chosen algorithm, the linewidth was determined as the difference between the left and right edge positions. The noise loop was repeated 100 times, resulting in 100 linewidth determinations. The standard deviation of these is the recorded repeatability. The difference between their average and the actual width of our simulated sample is the recorded bias.

3 Results and Discussion

The summary of input parameters in Table 1 includes the number of different values for each of the parameters. There are 28,800 possible combinations of the different parameter values. However, MONSEL does not permit nonzero corner radii to be combined either with undercut wall angles or nonzero θ_b, so simulations for these combinations were not performed. Excluding these, we simulated 14,400 different combinations of sample, instrument, noise, and data analysis parameters. For each of these, 100 different random instances of noise were added and two edge positions (left and right) were determined, for a total of close to three million edge determinations. The results constitute a multidimensional data set giving the repeatability and bias as functions of the eight parameters listed in the table. We do not attempt to summarize all of the results here. Instead, we limit discussion to three important results: repeatability as a function of noise amplitude, repeatability as a function

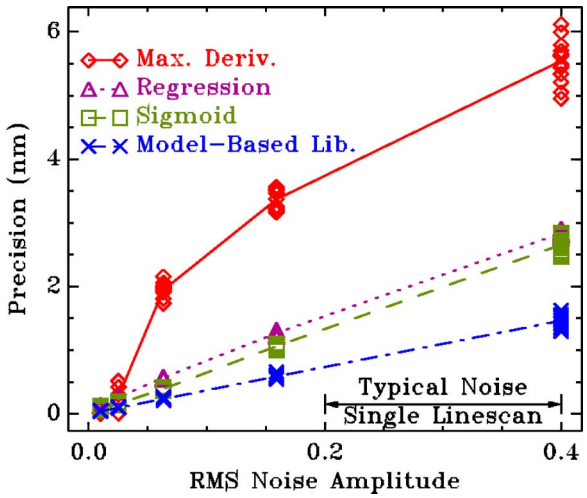


Fig. 7 Precision (three standard deviation repeatability) versus noise amplitude for four measurement algorithms. For this plot, the sidewall angle was held constant at θ=3 deg and the beam size was w_b=3 nm.

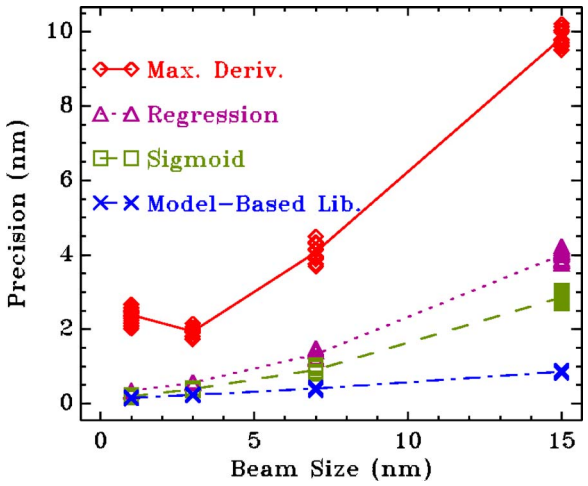


Fig. 8 Precision (three standard deviation repeatability) versus electron beam size for four measurement algorithms. The sidewall angle was θ=3 deg and the noise amplitude corresponded to the middle curve in Fig. 5.

of beam size, and bias as a function of sidewall angle. All of these are shown for each of the four edge-determination algorithms.

The precision as a function of noise amplitude is shown in Fig. 7. The figure summarizes a large amount of simulation data. To produce this plot, we extracted from the simulation results all of those linewidth determinations for which the beam spread was Gaussian with 3-nm width (w_b=3 nm), and the line's sidewalls were angled at 3 deg from the vertical (θ=3 deg). For each noise amplitude, there were a number of sets of simulation results satisfying these conditions, corresponding to the various values of the other parameters (e.g., power spectrum shape, corner radius, whether the lines were isolated or dense). These other parameters had relatively small effects on the measurement repeatability, allowing them all to be plotted on the same graph. This is the reason each curve has many symbols clustered at each noise amplitude. The lines are drawn through the average position within each cluster. The four curves correspond to the four different edge algorithms. Unsurprisingly, the curves all have positive slopes, indicating that the precision gets worse as the noise amplitude increases. It is worth noting that the noise amplitude is to a certain extent under the control of the metrology engineer. It may be reduced by increasing pixel dwell times or the amount of frame averaging, thereby improving measurement repeatability up to a limit imposed by throughput requirements or avoidance of sample damage.

The magnitude of the repeatabilities derived from our simulation study is comparable to that observed under manufacturing conditions. A notable feature of this result is that the repeatability depends strongly on the choice of measurement algorithm. Those algorithms that are based on fitting curves to the image have better repeatabilities than those that depend on differentiation. Between the two methods that use fitting, it appears that the more data are included in the fit, the better the repeatability. The most repeatable of these methods was a factor of 3 to 8 better than the least repeatable, depending on the noise amplitude. Attention to the measurement algorithm therefore offers the potential for improved repeatability without the lower throughput or increased potential for sample damage that accompanies longer pixel dwell times or more frame averaging.

The precision as a function of beam size is shown in Fig. 8. This result sheds light on the question we raised in the introduction about the relationship between spatial resolution and microscopy resolution, as defined in the ITRS. Consider first the relationship between the beam size and the instrument's spatial resolution. The resolution may be worse than the beam size if other phenomena (e.g., scattering of the electrons within the sample as illustrated in Fig. 3) degrade it by increasing the part of the sample with which the instrument interacts, but it is hard to imagine the interaction region being smaller than the beam size. Thus, the beam size plotted along the horizontal axis in Fig. 8 represents a reasonable lower bound on the CD-SEM's spatial resolution. The precision, plotted along the vertical axis, is implicitly (but incorrectly) identified by the ITRS with microscopy resolution. (The values tabulated in the ITRS are said to correspond to a specified precision to tolerance ratio, and they agree quantitatively with the precision values tabulated in another ITRS table.) The four curves have positive slope, indicating that measurement resolution declines as spatial resolution declines. This in itself is unsurprising. Poorer spatial resolution corresponds to more smearing of the edge in the image. Finding its location in the presence of noise becomes more subject to random error. However, the relative magnitudes immediately indicate that precision is not the same as spatial resolution. At w_b=15 nm for example, $P \approx 3$ nm for the sigmoidal fit method and $\approx$1 nm for the MBL method. Thus, for these methods at this noise amplitude w_b/P equals 5 and 15, respectively. This indicates that the precision can be more than an order of magnitude better than the spatial resolution. (If the spatial resolution had been worse—assuming it is as good as the spot size is an ideal case—this ratio would have been even larger. Similarly, more signal

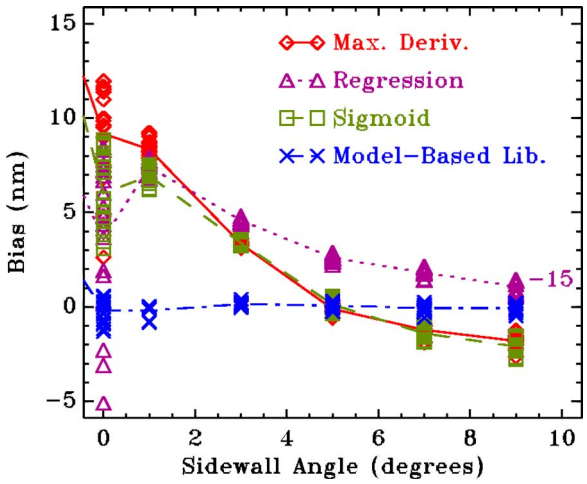


Fig. 9 Linewidth measurement bias as a function of sidewall angle for four measurement algorithms. For this plot, the beam size was w_b=7 nm and the noise amplitude corresponded to the second curve from the top in Fig. 5. The regression to baseline curve was shifted down by 15 nm.

averaging would decrease the noise and increase the ratio.) Thus, spatial resolutions on the order of a few nanometers need not pose an inherent limitation on achieving the ITRS's desired levels of precision.

The measurement bias as a function of sidewall angle is shown for positive angles in Fig. 9. The most biased method for this sample was regression to baseline. Its curve was shifted down by 15 nm to get it on the graph. The maximum derivative and sigmoidal fit methods were unbiased for sidewall angles near 5° but had biases of 6 nm to 12 nm for vertical sidewalls (θ=0°). As we have seen, however [Eq. (1)], for a comparative measurement it is the slopes that are important, not the absolute bias levels. Process variation can produce a range of sidewall angles, and it is generally not known that the angles of lines being compared are the same. Under this circumstance, the biases are not known to cancel, and there is a corresponding level of uncertainty in the width difference. The slopes are generally steepest for sidewalls closest to vertical, but are in the range of 1 nm/1° to 2 nm/1° for all but the MBL method. These slopes are consistent with experimental results[14] that were obtained contemporaneously with our study. For the methods in common use, uncertainty of 2° in the sidewall angle becomes an uncertainty of 2 nm to 4 nm in the linewidth. If the precision is around 1 nm, then clearly the measurement resolution in this example is dominated by the differential bias term in Eq. (1), not the precision term. As long as $\partial w/\partial\theta$ remains at 1 nm/1° to 2 nm/1°, the differential bias term continues to dominate until the sidewall angle process variation becomes smaller than 0.5° to 1° (three standard deviations).

Measurement errors are either varying or constant. Sometimes it is assumed that the former are measurable using the industry's standard test of precision and the latter are irrelevant. Constant errors are assumed to be either corrected by building in a constant offset (determined, for example, to force agreement with cross section for a test measurement) when reporting measurement results, or if not corrected at least of no consequence for measuring changes in linewidth. As Fig. 9 and our example illustrate, this view overlooks another possibility, namely that some components of bias, while constant for a given sample, vary from one sample to another. This variation cannot be measured in the industry's standard same-sample precision test, but it is potentially more significant than precision as a constraint on measurement resolution for real manufacturing processes in which different lines are being compared.

As shown in Fig. 9, the MBL method had slope close to zero for all positive angles. The model employed to generate the model-based library was the same MONSEL simulator used to produce our test images, except that the library did not include corner radius or depth of focus variation. Consequently, our library is a closer match to our test images than it is likely to be to real unsimulated images. The results show the kind of bias and repeatability that can be achieved by an MBL method when the model is closely matched to what nature does in the CD-SEM. They are not meant to show, indeed a simulation *cannot* show, that the particular model we used is itself such a close match. For that, one needs experimental evidence, to which we have made reference in Sec. 2.3. In practice, we expect that the better the model approximates nature, the nearer the MBL method's performance will approach that exhibited here.

4 Summary and Conclusions

Measurement repeatability (or precision) and measurement bias are both important components of a CD-SEM's ability to distinguish between lines with different widths. We describe a simulation study of these phenomena. For this study, image linescans are simulated for particular sample and instrument operating conditions. Random noise is added to these images, which are then analyzed to determine the linewidths.

Simulations are a complement to, not a replacement for, actual measurements. Without measurements, the relevance of the simulation to an actual manufacturing environment would be in doubt. On the other hand, simulation can explore a vast array of sample and instrument conditions that would be prohibitively time consuming to duplicate experimentally, and some things, e.g., the determination of bias, are more difficult experimentally than in simulations, where everything is known about the sample and measurement conditions. To keep the study as relevant as possible to current semiconductor manufacturing applications, the sample is chosen to be UV resist on an antireflective layer on a silicon wafer. Simulations are performed with a Monte Carlo electron trajectory simulator that has been demonstrated to produce image linescans closely resembling experimentally measured ones. The random noise added to the images is chosen to have the same power spectrum as noise measured in commercial CD-SEMs in a fabrication facility environment.

We determine measurement bias and repeatability for 14,400 combinations of values of eight variables: sidewall angle, top corner radius, line separation (isolated or dense), beam half angle (i.e., depth of focus), beam width, noise amplitude, the shape of the noise power spectrum, and the measurement algorithm. Among the more important results to emerge are these:

1. Microscopy resolution, defined by the ITRS as the ability to distinguish between linewidths that are nearly the same, does not depend only on instrument precision, but also on any sample-dependent measurement biases. We express this relationship in Eq. (1); which is derived in the appendix in Sec. 5.
2. Instrument precision gets worse as noise amplitudes get worse (e.g., as the amount of averaging is decreased), but for a given noise amplitude, there are significant differences in the precision achieved by different analysis algorithms. The model-based library algorithm was best in our simulations, with the sigmoidal fit second. These algorithms use more of the data within a given linescan to form the estimate of edge position than do many other industry-standard algorithms, such as maximum derivative.
3. Instrument precision gets worse as the spot size gets larger (i.e., spatial resolution gets worse). However, it is not inconsistent to have microscopy resolution a factor of 10 or more better than the beam size. Microscopy resolution under the ITRS definition is not the same as a microscope's spatial resolution. Because of the potential for confusion, we prefer the term "measurement resolution" to "microscopy resolution."
4. Sample-dependent measurement biases exist for the three common measurement algorithms we tried, and probably exist in all those in common use by the industry today. These have the potential to significantly degrade the ability to distinguish one linewidth from another, but they are an artifact of the measurement algorithm, not an inherent limitation of the CD-SEM. A model-based algorithm, to the extent the model is a good approximation, can be significantly less biased.

5 Appendix: Relationship between Measurement Resolution, Precision, and Measurement Biases

We derive the relationship between measurement resolution, which we take to be the uncertainty in the measurement of a width *difference*, $\Delta w = w_1 - w_2$, the measurement precision, and sample-dependent biases. The difference is normally obtained by subtracting the individually measured values of w_1 and w_2. Let us suppose that our measurement is sensitive to a number of parameters, such that in a series of measurements

$$w_{1i} = w_{1t} + \varepsilon(q_i, \theta_1). \tag{3}$$

Here w_{1i} is the i'th measured value of the width of the first line, w_{1t} is the true value of the width, and ε is the difference between the two (the error). For the sake of simplicity, we limit consideration to two parameters, on which ε (and hence w_1) depend. We label them q and θ. The generalization to more than two will be obvious. We choose these parameters to be representative of two different types. The parameter q represents parameters, the variations of which are sampled during repeated width measurements on the same sample, i.e., during the course of the semiconductor industry's defined precision measurement. Parameters of this type would be, for example, relative position of the sample and the electron gun, the electric and magnetic fields through which the beam must pass to reach the sample, voltage levels in the electronics that process the signal, etc. These vary due to vibration, electrical noise, shot or thermal noise in electronics, etc., and will as a rule differ from one measurement to the next [hence q_i instead of simply q in Eq. (3)]. The parameter θ represents parameters that are not expected to vary during the precision measurement, but which may differ between measurements on samples 1 and 2 (hence θ_1 rather than simply θ). For the sake of concreteness, we may think of θ as the sidewall angle, but any other sample-dependent property (e.g., corner radii, proximity of neighboring lines, sample composition) will be represented by a parameter like this one. Non-sample parameters (e.g., beam size) might be of this type under some circumstances, for instance if w_1 and w_2 are measured on different instruments that are not known to be perfectly matched.

Subtracting from Eq. (3) the analogous expression for sample 2, the i'th measurement of Δw can be written as

$$\Delta w_i = w_{1t} + \varepsilon(q_i, \theta_1) - [w_{2t} + \varepsilon(q_i', \theta_2)]. \tag{4}$$

We put a prime on q_i the second time it appears because the measurements of w_{1i} and w_{2i} are not simultaneous and the q's do not therefore have the same value. Expand the right side of Eq. (4) around the average value of q during a measurement. Then to first order

$$\Delta w_i = \Delta w_t + \varepsilon(\langle q\rangle, \theta_1) - \varepsilon(\langle q\rangle, \theta_2) + (q_i - \langle q\rangle) \left.\frac{\partial \varepsilon}{\partial q}\right|_{\langle q\rangle, \theta_1} - (q_i' - \langle q\rangle) \left.\frac{\partial \varepsilon}{\partial q}\right|_{\langle q\rangle, \theta_2}. \tag{5}$$

We have defined the true value of the width difference by $\Delta w_t = w_{1t} - w_{2t}$ and used $\langle\rangle$ to denote the average. Obtaining the average and standard deviation of the Δw_i is straightforward. The average is obtained by summing both sides over all i and dividing by the number of measurements. The first three terms do not depend on i, so they survive unscathed. The last two terms have averages of 0, so

$$\langle \Delta w\rangle = \Delta w_t + \varepsilon(\langle q\rangle, \theta_1) - \varepsilon(\langle q\rangle, \theta_2). \tag{6}$$

The standard deviation is defined such that $\sigma_{\Delta w}^2 = \langle (\Delta w_i - \langle \Delta w_i\rangle)^2\rangle$. Using Eqs. (5) and (6) to evaluate this yields

$$\sigma_{\Delta w}^2 = \left\langle (q_i - \langle q\rangle)^2 \left(\left.\frac{\partial \varepsilon}{\partial q}\right|_{\langle q\rangle, \theta_1} \right)^2 \right\rangle + \left\langle (q_i' - \langle q\rangle)^2 \left(\left.\frac{\partial \varepsilon}{\partial q}\right|_{\langle q\rangle, \theta_2} \right)^2 \right\rangle. \tag{7}$$

The cross-term vanishes because q_i and q_i' are uncorrelated. Calling the first term on the right σ_1^2 and the second one σ_2^2, we can write

$$\sigma_{\Delta w}^2 = \sigma_1^2 + \sigma_2^2. \tag{8}$$

The two terms on the right are the repeatabilities of measurements on a single sample. That is, they are the standard deviations that would be measured on the two samples during the industry's same-sample precision measurement. [To see this, expand Eq. (3), which describes a single-sample width determination rather than a width difference, to first order and then calculate its standard deviation, analogously to what we have just done starting with Eq. (4). The result would be identical to the first term in Eq. (7). The second term follows in the same way from measurements on the second sample.]

If, in Eq. (6), the angles θ_1 and θ_2 differ by only small amounts from the average $\langle\theta\rangle$ for samples produced by the manufacturing process, then we can expand the error terms around $\langle\theta\rangle$ and solve for Δw_t to get to first order

$$\Delta w_t = \langle \Delta w \rangle - (\theta_1 - \langle\theta\rangle) \left.\frac{\partial \varepsilon}{\partial \theta}\right|_{\langle q\rangle,\langle\theta\rangle} + (\theta_2 - \langle\theta\rangle) \left.\frac{\partial \varepsilon}{\partial \theta}\right|_{\langle q\rangle,\langle\theta\rangle}. \tag{9}$$

This is our best estimate for the true value of the width difference. In this form it contains three terms, two of which constitute a correction for measurement bias. In practice, we usually do not have independent information about θ_1 and θ_2. In that case our best estimate will be that they are equal to $\langle\theta\rangle$, since that is the average value produced by the manufacturing process. With this, the bias terms in Eq. (9) are each 0, and our best estimate of the width difference becomes

$$\Delta w_t = \langle \Delta w \rangle. \tag{10}$$

Our uncertainty in the estimates of θ_1 and θ_2 will be equal to the standard deviation σ_θ of our process. The remaining term on the right side of Eq. (9) is the measured width. If it is determined by a single measurement of w_1 and w_2, this uncertainty is just the standard deviation, Eq. (8). The total uncertainty $u_{\Delta w}$ in our estimate of Δw_t is obtained in the usual way[15] by adding the uncertainties of the individual terms in quadrature.

$$u_{\Delta w}^2 = \sigma_{\Delta w}^2 + 2\left(\frac{\partial \varepsilon}{\partial \theta}\sigma_\theta\right)^2 = \sigma_1^2 + \sigma_2^2 + 2\left(\frac{\partial \varepsilon}{\partial \theta}\sigma_\theta\right)^2. \tag{11}$$

The derivatives are evaluated at the average values, $\langle q\rangle$ and $\langle\theta\rangle$, as in Eq. (9), though we have dropped the explicit specification of where the derivatives are to be evaluated to simplify the notation. It remains only to clean this up a bit for the sake of clarity. Since ε differs from the measured width only by a constant term [the true width, see Eq. (3)], we can replace $\partial\varepsilon/\partial\theta$ with $\partial w/\partial\theta$. Since we have assumed θ_1 and θ_2 are both close to $\langle\theta\rangle$, we expect $\sigma_1 \approx \sigma_2$. The semiconductor industry multiplies standard deviations by three to obtain precisions ($P = 3\sigma_1 \approx 3\sigma_2$), so let us use the same factor for all uncertainties. We let $U_\theta = 3\sigma_\theta$ and identify the measurement resolution R with $3u_{\Delta w}$. Then Eq. (11) becomes

$$R^2 = 2P^2 + 2\left(\frac{\partial w}{\partial \theta}U_\theta\right)^2 + \ldots. \tag{12}$$

Obviously, if there are other differences between the measurements of the two widths that are not sampled by the precision measurement, there would be additional terms of the same form as the second one. We have appended the … to denote such other possible terms. This result is the same as Eq. (1).

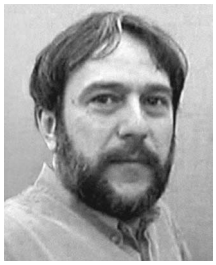
Acknowledgments

The idea for this study was suggested to us by Chas Archie of IBM, to whom belongs credit for the inspiration but no blame for any errors of execution. Ben Bunday and Michael Bishop of International SEMATECH supplied us with samples of CD-SEM images, from which we derived the noise characteristics. The authors gratefully acknowledge International SEMATECH and the National Semiconductor Metrology Program for financial support.

References

1. *International Technology Roadmap for Semiconductors* (Semiconductor Industry Association, 2003), Table 116a, see http://public.itrs.net.
2. D. C. Joy, "Some issues in SEM-based metrology," *Proc. SPIE* **3332**, 102–109 (1998).
3. J. R. Lowney, "Application of Monte Carlo simulations to critical dimension metrology in a scanning electron microscope," *Scanning Microsc.* **10**, 667–678 (1996).
4. J. R. Lowney, A. E. Vladár, and M. T. Postek, "High-accuracy critical-dimension metrology using a scanning electron microscope," *Proc. SPIE* **2725**, 515–526 (1996).
5. R. Browning, T. Eimori, E. P. Traut, B. Chui, and R. F. W. Pease, "An elastic cross section model for use with Monte Carlo simulations of low energy electron scattering from high atomic number targets," *J. Vac. Sci. Technol. B* **9**(6), 3578–3581 (1991).
6. D. C. Joy and S. Luo, "An empirical stopping power relationship for low-energy electrons," *Scanning* **11**, 176-180 (1989).
7. M. Kotera, R. Ijichi, T. Fujiwara, H. Suga, and D. Wittry, "A simulation of electron scattering in metals," *Jpn. J. Appl. Phys., Part 1* **29**(10), 2277-2282 (1990).
8. C. Möller, "Zur theorie des durchgangs schneller elektronen durch materie," *Ann. Phys.* **14**, 531 (1932).
9. R. D. Evans, *The Atomic Nucleus*, McGraw-Hill, New York (1955).
10. J. S. Villarrubia, A. E. Vladár, B. D. Bunday, and M. Bishop, "Dimensional metrology of resist lines using a SEM model-based library approach," *Proc. SPIE* **5375**, 199–209 (2004).
11. J. S. Villarrubia, A. E. Vladár, J. R. Lowney, and M. T. Postek, "Scanning electron microscope analog of scatterometry," *Proc. SPIE* **4689**, 304-312 (2002).
12. M. P. Davidson and A. E. Vladár, "An inverse scattering approach to SEM line width measurements," *Proc. SPIE* **3677**, 640-649 (1999).
13. J. S. Villarrubia, A. E. Vladár, J. R. Lowney, and M. T. Postek, "Edge determination for polycrystalline silicon lines on gate oxide," *Proc. SPIE* **4344**, 147–156 (2001).
14. V. A. Ukraintsev, "Effect of bias variation on total uncertainty of CD measurements," *Proc. SPIE* **5038**, 644–650 (2003).
15. *Guide to the Expression of Uncertainty in Measurement*, International Organization for Standardization, Geneva, Switzerland (1993, corrected and reprinted in 1995). Also available as a U.S. National Standard: NCSL Z540-2-1997.

John S. Villarrubia received his PhD in physics from Cornell University in 1987. He was a visiting scientist for two years at IBM, where he did scanning tunneling microscopy of silicon surfaces. At the National Institute of Standards and Technology since 1989, he has been working on aspects of nanometer-scale dimensional metrology, particularly modeling the instrument function for atomic force and scanning electron microscopes and using these models of the instrument function to correct images for measurement artifacts.



András E. Vladár is the leader of the Scanning Electron Microscope Metrology Project at the National Institute of Standards and Technology (NIST). He is an expert in scanning electron microscope-based dimensional metrology and one of the best known research scientists and a technical leader of this field. He received his MS in 1977 and his PhD in 1984 in electrical engineering from the Technical University in Budapest, Hungary. He was a research fellow at the Research Institute for Technical Physics of the Hungarian Academy of Sciences, Budapest until he joined the Nano-Scale Metrology Group in the Precision Engineering Division at NIST as a guest researcher in 1991. From 1995 to 1999, he was a metrology engineer and member of the technical staff of the Hewlett-Packard ULSI Research Laboratory, Palo Alto, California. He returned to NIST as a permanent member of the staff in 1999. He has been a key member to the work of the Metrology Council of International SEMATECH and has contributed to the microscopy and metrology sections of the International Technology Roadmap for Semiconductors (ITRS). His is also engaged in the development of various dimensional metrology methods for nanotechnology.

Michael T. Postek is the leader of the Nano-scale Metrology Group and program manager of the Nanomanufacturing Program at the National Institute of Standards and Technology (NIST). He is a nationally and internationally recognized expert in scanning electron microscope (SEM) critical dimension (CD) metrology as a research scientist and a technical leader. His current research areas include nanotechnology and nanometrology. He has been involved with micro and nano-scale scanning electron microscope-based length metrology for over 20 years and has published over 120 papers, articles, and book chapters on this and related topics. He is a charter member of the International SEMATECH Advanced Metrology Advisory Group and a contributor to the International Technology Roadmap. His work centers on improvements in SEM relative to semiconductor and nanotechnology related metrology. Among his credits are: NIST bronze medal, two NIST silver medals, 1998 R&D 100 Award, Federal Laboratory Consortium Award for Technology Transfer, and NIST Top 100 Papers Award 2000. He received his BA from the University of South Florida in 1973, his MS from Texas A&M University in 1974, his PhD from Louisiana State University in 1980, and an Executive MS in technology management from the University of Maryland in 1997.