Influence of Metrology Error in Measurement of
Line Edge Roughness Power Spectral Density

Benjamin D. Bunday[a], Chris A. Mack[b]

[a]SEMATECH, Albany, NY, 12203, USA

[b]Lithoguru.com, 1605 Watchhill Rd., Austin, TX 78703, E-mail: chris@lithoguru.com

Abstract

Line-edge roughness (LER) and linewidth roughness (LWR) in lithography are best characterized by the roughness power spectrum density (PSD), or similar measures of roughness frequency and correlation. The PSD is generally thought to be described well by three parameters: standard deviation, correlation length, and roughness exponent. The next step toward enabling these metrics for pertinent industrial use is to understand how real metrology errors interact with these metrics and what should be optimized on the critical dimension scanning electron microscopy (CD-SEM) to improve error budgets. In this work, **J**ava **Mon**te Carlo Simulator of **S**econdary **El**ectrons (JMONSEL) simulation is used to better understand how various SEM parameters, beam size/shape, and sample profile influence SEM line edge uncertainty and also some of the systematic shifts in edge location assignment. A thorough understanding of the impact of the SEM on the measurement results enables better measurement of LER PSD and better interpretation of measurement results.

Subject Terms: power spectral density, PSD, line-edge roughness, linewidth roughness, LER, LWR, JMONSEL, CD-SEM

1.   Introduction

Control of roughness in photoresist and post-etch features has become more important as features continue to shrink. However, the need for control is accelerating; with planar CMOS, the roughness of the bottom edge defines a transistor's effective width, whereas with our contemporary FinFETs and tri-gates, the roughness of the entire sidewall acts as the active surface for these non-planar transistors. Metrology studies have gradually improved our ability to measure roughness, but as we approach the 10 nm node, sub-nanometer roughness control will be required, with metrology error decreased to a mere 20% of that (i.e., approaching size ranges that can be described as atomic scale). Achieving such uncertainty levels will be quite challenging [1] [2] [3].

The current SEMI standard for line-edge roughness (LER) and linewidth roughness (LWR) measurement [4] specifies the sampling intervals and length, as well as the roughness amplitude metric in terms of a standard deviation. While this standard was, at the time, a very important step in the right direction for guiding the industry toward compatible and consistent measurements that reflected industry requirements, more information is needed for fully characterizing roughness, particularly the various spatial frequency characteristics. LER and LWR in lithography are best characterized by the power spectral density (PSD) of the roughness, or similar measures of roughness frequency and correlation. The PSD, in turn, is generally thought to be described well by three parameters: the standard deviation (root mean square [RMS] roughness), correlation length, and roughness exponent. A thorough theoretical study of these metrics was recently published by one of the coauthors [5] [6]. The authors recommend that the SEMI standard should be updated along these principles.

Four major lessons were recently demonstrated in the previous work [5] [6]. First, PSD for LER is best characterized by the three parameters σ (the standard deviation), ξ (correlation length), and *H* (rolloff exponent, also called the roughness exponent). Second, since the error of a single PSD is equal to the amplitude itself at each frequency, averaging many PSDs is required to reduce the error, such as averaging 10–100 PSDs from nominally equivalent features. Third, data windowing is required for reducing spectral leakage (leakage meaning a spreading or blurring effect where the PSD components shift in frequency due to measurement of a finite line length). And fourth, the sampling distance is

optimized to balance aliasing with averaging when set to double the full width half maximum (FWHM) of the effective spot size including the scanning electron microscopy (SEM) interaction volume (best described as the Gaussian beam/resist interaction FWHM in the case of a resist line). While these lessons describe the general concepts of proper sampling, the next step toward enabling these new metrics for pertinent industrial use is to understand how real metrology errors interact with each of those metrics, and what should be optimized on the critical dimension SEM (CD-SEM) to improve the error budgets. Another previous work introduced the idea of the uncertainty of a single edge location as being a fundamental building block for how metrology error influences various CD and LER metrics [7] [8], and this edge uncertainty (σ_ε) is also directly related to the LER bias [8] and to the integral of the high frequency noise floor through the entire frequency range of a given measurement.

In this work, JMONSEL simulation [9] [10] [11] is used to better understand how various SEM parameters, beam size/shape, and sample parameters, such as materials used and profile, influence σ_ε and also some of the systematic shifts in edge location assignment. A thorough understanding of the impact of the SEM on the measurement results enables better measurement of LER PSD and better interpretation of measurement results. Eventually, these results can be input into previously-presented random edge roughness generators and solvers [6] [12] and iterated to generate the expected error responses in the various proposed roughness metrics.

2.   Simulation of Edge-Feature Linescans

As stated above, we wish to explore edge detection errors—the typical effects on accuracy offset (systematic bias) and random edge detection error of various CD-SEM and sample parameters, as both of these effects, when applied to different points along the sidewall of a measured feature, can influence the measurement of roughness. Simulations allow for more flexibility than with experiments, both in the sample and the experimental conditions and in the statistical validity and metrics that can be known. For example, accuracy can be easily measured without resorting to reference metrology, since the design of the feature is inherently the reference value, and samples can be made truly perfect, with no footing, top corner rounding (TCR) or roughness/variation, unless designed in, and sidewall angles (SWAs) are exactly definable. There are limitations, in that the results are only as good as the physical models utilized, and not all effects are necessarily represented (such as maybe charging, electronic noise of the instrument, homogeneity of the materials, etc.

Simulations in this study were performed using JMONSEL (**J**ava **Mon**te Carlo Simulator of **S**econdary **El**ectrons), a program developed at the National Institute of Standards and Technology (NIST) which has the capability to use finite element analysis in order to follow primary electrons a they enter a material, scatter, lose energy, and generate secondary/backscattered electrons.[9] By monitoring the electrons that exit the material and are captured, the electron yields can be found at each point and are plotted as line scans (often called waveforms). The physical models in JMONSEL are the best known models in the literature in the energy ranges used here, with complete transparency in their documentation, definition, and execution [10] [11]. Several different types of data are generated from the simulation. One is a waveform, which is a plot showing detected electron yields as a function of position (the electrons detected as the beam passes over each pixel). Energy histograms can also be recorded, and visualizations of the electron trajectories and resulting scattered electrons can be visualized if desired. JMONSEL is used here as a "virtual SEM", where the user can input idealized structures from a limited list of materials, with perfect user-defined geometries, including geometries that cannot be currently fabricated. Additionally, these idealized structures have zero roughness (unless roughness is added), with no process variation, and with user-defined parameters, i.e. the reference values are exactly known, meaning accuracy is a simple bias calculation as a difference between a measured value and a designed value. The user can also define SEM parameters such as the number of incident electrons per pixel, pixel size, spot size and beam energy. The program can also monitor charging phenomena in and around the sample, although most of the samples here are small in height and made of materials that are not so susceptible to charging, thus charging was judged to be a secondary effect for a single measurement; additionally, enabling charging tracking makes the calculations extremely slow, thus charging was neglected in the simulations performed for this study.

The main part of the study was a systematic DOE of dense line/space features. Feature/substrate composition, SWA, TCR, footing, beam energy and beam spot size were all varied within the DOE, not as full factorials, but as single-parameter perturbations from our chosen standard condition, which was effectively a Si CMOS gate structure—Si lines on 2 nm thin SiO_2 on Si, SWA of 89°, spot size of 0.5 nm 1σ (full width at half maximum (FWHM) ~1.2 nm, similar to current 1.4 nm resolution of current CD-SEMs), and beam energy of 500 V. All features in the DOE used linewidths of

30 nm with 60 nm pitch and 50 nm height; to avoid possible complications, it was decided to design the lines to be larger than the small feature regime, where the interaction volume is smaller than the feature and tight spaces that might approach SEM resolution limits, thus the choice of 30 nm line/space.

The TCR used had a 3 nm radius in one of the perturbations from the standard condition, and 1 nm, 2 nm and 3 nm foot with 2 nm height was added each sidewall in the conditions to explore footing effects.

The lines were made of various materials including Si, poly(methyl methacrylate) (PMMA), SiO_2 and Cu, with the substrate either Si or Si coated with 2 nm thin SiO_2 or 80 nm thick PMMA. PMMA was chosen as it is believed to be a reasonable approximation of a photoresist material or organic ARC, in terms of SEM imaging. For each virtual test sample/condition in this study, many linescans (51) with N=200 incident electrons per pixel (for a total of N=10200 electrons) were calculated, with 0.5 nm pixel sizes. At each pixel, averages and standard deviations of all the N=200 values were calculated, yielding a high precision average waveform and a figure for noise. An example waveform is shown in Figure 1.
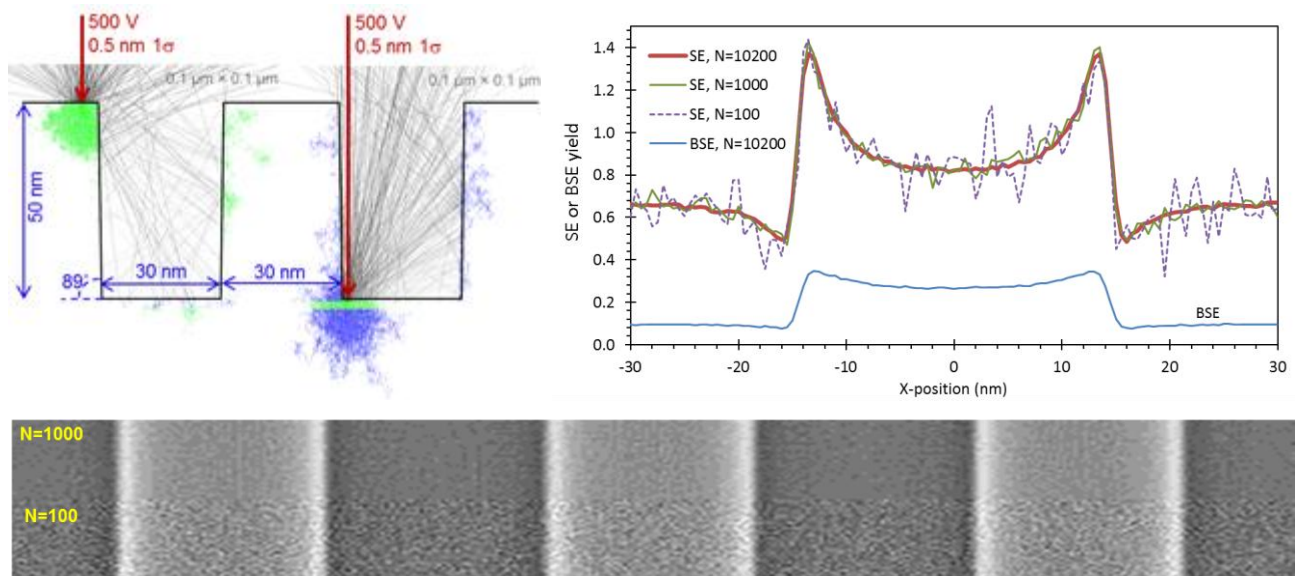

Figure 1: Example outputs from JMONSEL for DOE base condition. Top-left: Simulated trajectories in the 30 nm line/space Si-on-thin SiO_2 structure at 500 V and 0.5 nm spot, where the interaction volumes can be seen at feature tops and bottoms. Top-right: Example waveform with averaging of all data for the DOE's base condition, including waveforms for SE and BSE yields. Additionally, two versions of the SE waveforms with noise added are shown, for N=100 and N=1000. Bottom edges of the feature are exactly at ±15 nm. SE electrons are defined as having energies ≤ 50 eV, with BSE electrons defined as having energies > 50 eV. Bottom: Simulated image of same waveform with a random instance of expected noise for N=1000 (top) and N=100 (bottom).

Through averaging different subsets of the N=200 waveforms and calculating the standard deviations, it was confirmed that the noise almost exactly varied as $N^{1/2}$ (i.e., SNR ~$N^{-1/2}$) as expected, as shown in Figure 2. Kurtosis and skew, which are statistical metrics for the peakedness and symmetry of a distribution compared to a true normal distribution, were also confirmed to be small. Thus the waveform noise from the simulator can be treated as Gaussian in nature.
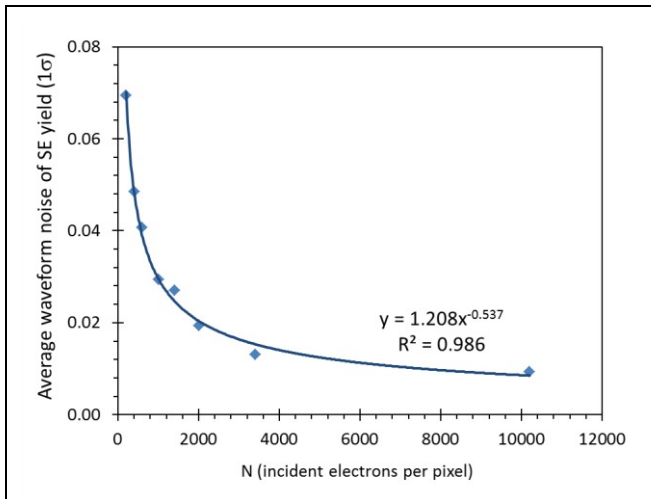
Figure 2: Example waveform noise from the main condition of the DOE, showing that waveform noise varies as $N^{1/2}$, as expected.

To turn our virtual SEM into a "virtual CD-SEM", various edge detection algorithms were made into Excel functions thru Excel VBA macro programming. These functions allow for defining a waveform function, defining whatever required thresholds and baselines, and adding a measured 1σ noise and filtering. Several edge algorithms are available, including threshold, linear regression, sigmoidal fit, maximum derivative, and Gaussian fits. Some of the more advanced fitting algorithms can reduce edge detection noise, but since the error handling of some of these are not yet perfected, it was decided to use the basic threshold algorithm for this work, as it is the most robust and also probably the most commonly used in the industry. Model-based algorithms, which we have not developed, would further reduce these errors. The edge detection noise of the threshold algorithm is typically higher than all but the max derivative algorithm, but since the nature of the edge detection noise is what is to be studied, it is an appropriate choice. For this algorithm, baselines are chosen as averages of the SE yields between the lines, and thresholds of 20% and 80% were chosen for bottom and top edge detections, respectively. Background on these general algorithm types can be found in the literature [7] [13].

With the functionalized edge detection algorithm allowing for inputting the expected waveform noise and the waveform noise shown to vary as $N^{1/2}$, a given waveform can be tested using the base average waveform, with noise at N=1000 used as a baseline and adjusted by $N^{1/2}$, to run many instances of the same edge detection calculation for nominally identical cases, to calculate the average edge detection error (the accuracy offset at one point along the line), and also as a standard deviation, which is also known as the edge detection reproducibility (σ_ε [7]) for the given case.
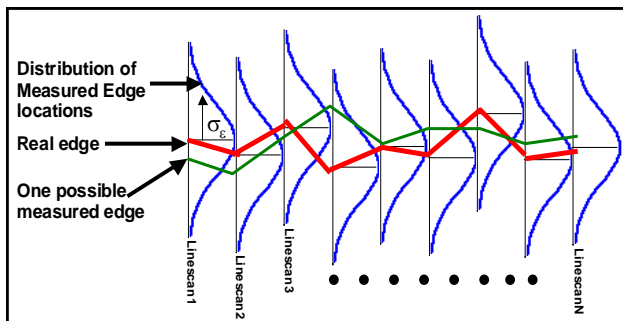

Figure 3: Schematic for problem of considering how noise effects the apparent measured position of a real line edge. The thick meandering line represents the real edge. The slimmer meandering line represents one measured edge out of infinite possibilities. The Gaussian curves overlaid on each linescan represent the distribution of measured edge locations [7].

This edge detection reproducibility (σ_ε) is a fundamental building block for understanding the random aspects of dimensional metrology error, as the sum of its effect over all the edge locations of a given measurement are related to its magnitude [7]. With many linescans its effect gets averaged down for CD, LER, and LWR measurements, but it adds a bias in quadrature for LER and LWR measurements, meaning that a perfectly smooth line will still yield an LER measurement of σ_ε. The concept can be thought of as the "fuzziness of the edge", as demonstrated in Figure 3. In turn, σ_ε can be reduced as $N^{-1/2}$ by oversampling (i.e., if the measurement dose is increased). However, there are practical limits to how much this can be done, such as resist shrinkage, charging, contamination, throughput, and mechanical or environmental electromagnetic noise. Choice of edge detection algorithm is another major factor, as σ_ε is effectively a function of noise interacting with edge detection algorithm. Averaging along a waveform can also reduce σ_ε, at the expense of edge detection accuracy offset, and averaging multiple linescans (binning) can also reduce σ_ε, at the expense of the frequency bandwidth sensitivity of a roughness measurement (i.e., the Fourier window being sampled is greatly reduced). This has the effect of reducing the roughness measurement, as fewer spatial frequency components are summed. So ideally, for LER metrology at future nodes, such spatial filtering strategies should be avoided. So to summarize, σ_ε is strongly dependent on tool and recipe parameters, such as dose per pixel, imaging focus or astigmatism (which effectively means artificially increasing beam spot size), edge detection, and system electronic noise [7].

For each of the simulation conditions, 1000 edge location values for bottom and top and with different noise statistics are plotted as histograms (shown in Figure 4). In these data sets, it should be noted that since the error handling in the edge algorithm code still is not perfect, a few severe outliers were removed since those points represented the kinds of measurements that would be culled by post-processing of the data in real life utilization.
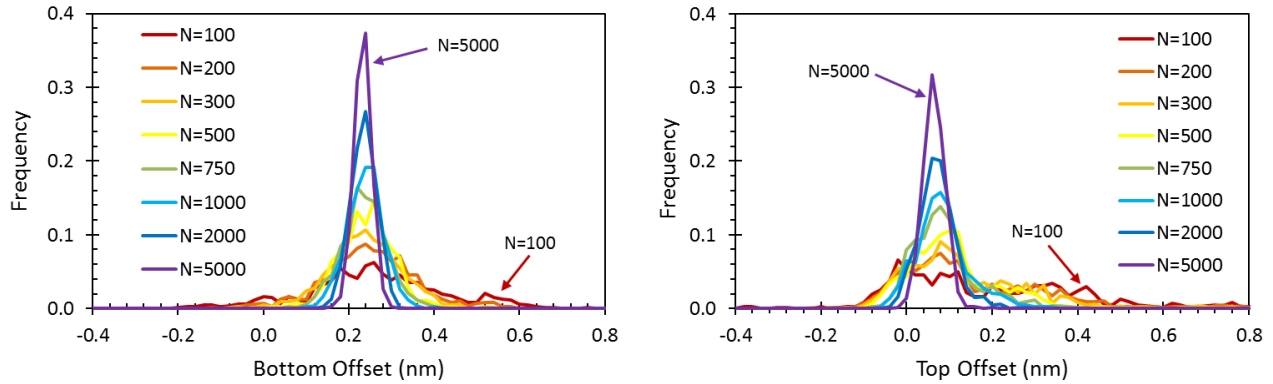

Figure 4: Example results of bottom and top edge location histograms, for the base condition of the DOE.

These same distributions can be represented by an average value and standard deviation (effectively σ_ε), as shown in Figure 5. The σ_ε metric, which represents the edge detection reproducibility, basically varies as $N^{-\frac{1}{2}}$ as did the waveform noise (as expected).
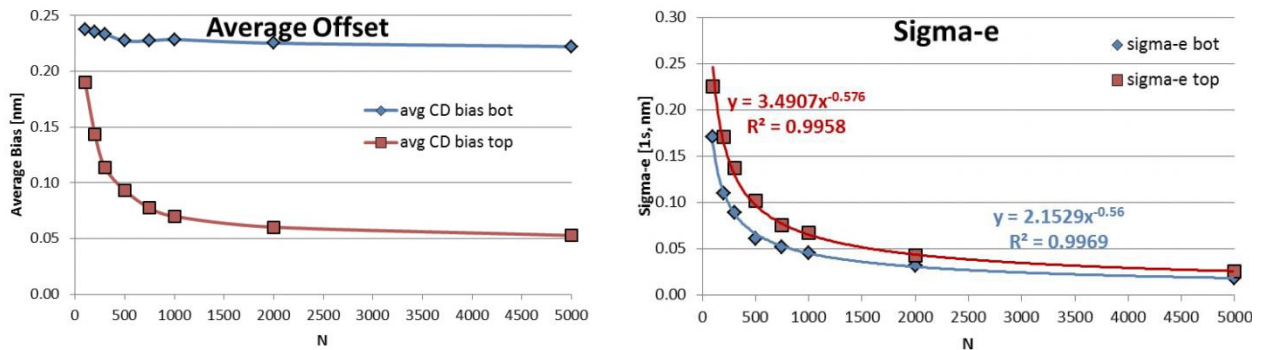
Figure 5: Example results of bottom and top edge location trends for average offset and σ_ε, for different N values, for the base condition of the DOE. Note that σ_ε decreases as $\sim N^{-\frac{1}{2}}$, as the waveform noise did.
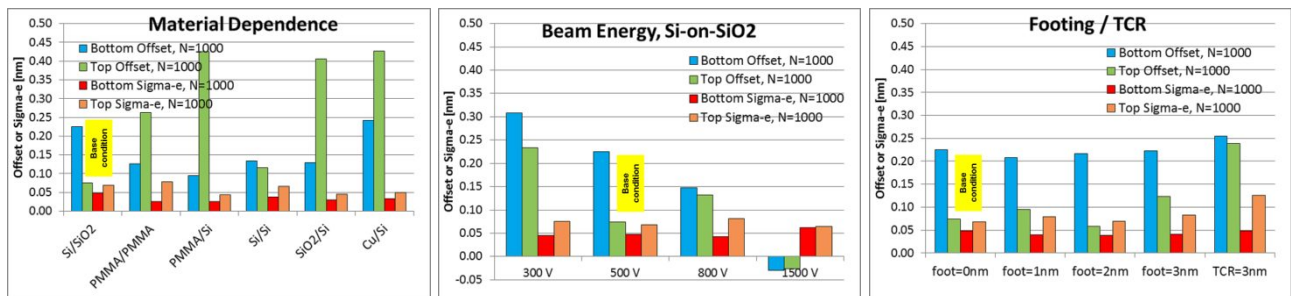

Figure 6: Dependence of average offsets and σ_ε values for the DOE for various factors. The DOE base condition is Si lines on thin SiO$_2$ on Si substrate imaged at 500 V with a 0.5 nm 1σ beam, and is marked in each of the graphs. Left: Dependence on materials, with the different x-axis categories representing different line materials (the name before the slash) on different substrates (name after the slash). Center: Dependence of beam energy on average offsets and σ_ε values for the Si-on-SiO$_2$ stack. Right: Effect of footing and top corner rounding (TCR) of the profile on average offsets and σ_ε values.

The average offset and σ_ε results of the JMONSEL DOE are shown in Figure 6. Some important and interesting trends can be seen by comparing the values of the various conditions against the base condition. For both average edge offset and σ_ε, large changes are seen with SWA, as shown in Figure 7. This was already well-known for offset, but σ_ε is strongly affected also, and seems to have a significant quadratic dependence on SWA for SWA $\leq 90°$. Material and beam voltage also have strong influences on both metrics, independent of spot size as all these values were computed with 0.5 nm spot. In addition to σ_ε, accuracy offset differences from material non-homogeneities and SWAs will appear as apparent LER. TCR has a large effect on top σ_ε, much more than does footing, which is a moderate effect. Surprisingly, the footing had minimal effect on the bottom σ_ε. Note that these σ_ε values are based only on sampling statistics (and on edge algorithm, etc.), meaning that these are best possible theoretical values, and real values from electronic noise, vibration (etc.) will be worse. Also, since σ_ε adds a bias to roughness measurement, it will always be there unless suppressed. The σ_ε values seem reasonable; real tools have been measured to have σ_ε of between 1 nm and 2 nm, but these are 3σ values. The values calculated in these simulations make up a large portion of such values, so it is likely that other noise in the tool or sample can make up the difference. Alternately, this means that while some improvement can be made, there is a significant contribution to the noise from just the scattering statistics, which will act as a limit, although one that can be suppressed by oversampling or improved edge algorithms.

A few other runs of the simulator were also performed to understand how beam spot size influences these metrics, with results shown in Figure 8. This is equivalent to a SEM resolution improvement on the low end of the spot size scale, and similar to an out of focus beam or an older tool on the higher end of the spot size scale; $1\sigma = 0.5$ nm is thought to be roughly equivalent to the performance of contemporary tools. SEM spot size is related to, but not exactly the same as, image resolution; a rough guideline is to use a spot's FWHM as close to the *image resolution*, which by definition is 2.35σ for a Gaussian spot. It should be noted that while spot size is a measurable metric and should be sample independent, image resolution is a term which is ill-defined and inconsistent among the literature and the industry, with no standard, and seems quite sample and SEM-condition dependent. We include a value for an unobtainable, *perfect zero spot size* (actually 0.01 nm to avoid possible strange boundary cases of a delta function-like beam impacting an edge exactly), impossible due to diffraction limits, but still of interest, as results at that condition demonstrate the effects of the scattering and interaction volume only, not convoluted with the spot profile. Note that offset and σ_ε are both strongly dependent on spot size. For offset, the offset has a strong linear dependence on spot size, and top likely does also although it is likely that the waveform at the top of the feature used here has some kind of second order non-linear interaction with the edge algorithm. The edge detection repeatability σ_ε has a strong quadratic dependence on spot size. Both fits are included, with the quadratic fit having significantly higher correlation, with enough data points to demonstrate this significance.
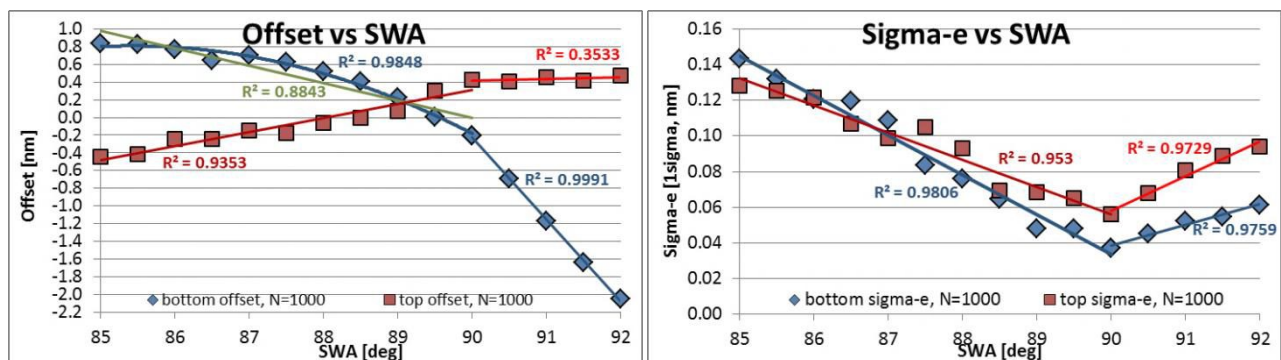

Figure 7: Influence of sample SWA on average offset and σ_ε, using the Si lines on thin SiO_2 virtual sample set. Note the response is roughly linear (maybe quadratic, but need more points to confirm). For both offset and σ_ε, the data points at SWA $\geq 90°$ are treated with a separate line fits as all these edge detection algorithms are discontinuous in behavior due to ceasing to image the feature bottom of a re-entrant profile (N=1000).
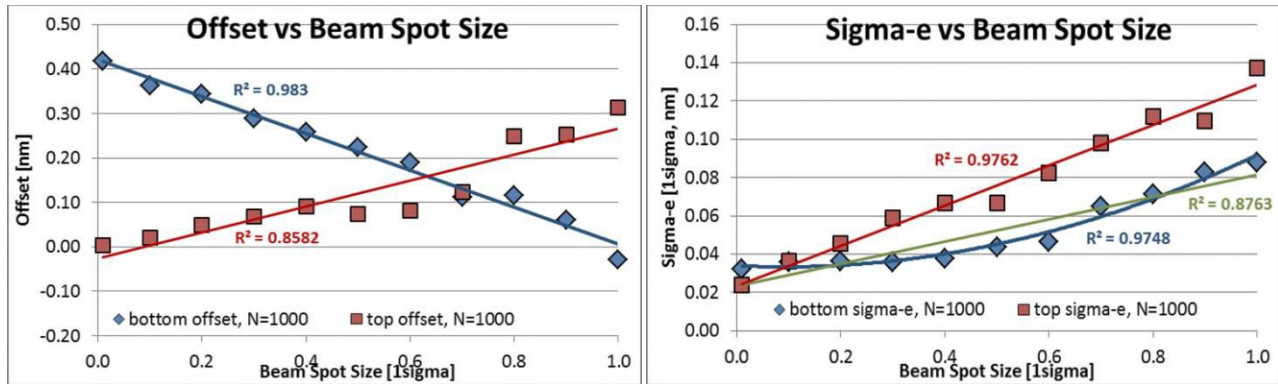
Figure 8: Influence of SEM beam spot size on average offset and σ_ε, using the Si lines on thin SiO$_2$ virtual sample set, at the DOE's base condition at 500 V. Note the response is roughly linear, although a quadratic dependence is more likely for bottom σ_ε (N=1000).

Aside from the DOE, another more basic feature explored in this simulation study was a single 50 nm step of 50 nm height on a Si surface. The resulting waveforms, collected with 0.5 nm Gaussian spot size (1σ) and a *perfect* 0.01 nm spot size at various beam energies (300 V, 500 V, 800 V) and with 0.1 nm pixel size and N=25000 incident electrons, are used next to study functional forms of the resulting waveform from such a simple perfect target. Materials of the step were also varied, although to a lesser extent, with both Si and PMMA.

3.   Simplified Linescan Model

While rigorous 3D Monte Carlo simulations of SEM linescans can be extremely valuable, they are often too computationally expensive for certain applications. Here, our goal is to understand how the systematic response of the SEM measurement tool impacts statistical LER measures such as the PSD. Rather than applying rigorous Monte Carlo simulations to a rough feature, we will instead develop a simplified, analytical linescan model that will be more computationally appropriate to the task. This analytical linescan expression will then be fit to the rigorous Monte Carlo simulations. Finally, in the next section, the analytical linescan model will be applied to the problem of measuring a rough line edge.

The analytical linescan model will be developed using the simplest possible feature: a silicon edge on a silicon wafer. A point beam of electrons will be scanned in the *x*-direction perpendicular to the edge feature, which is placed at *x* = 0 (later a PMMA edge will be used, see Figure 9). The linescan, corresponding to the detected secondary electrons, will be *SE*(*x*). Consider first a bare silicon wafer (corresponding to when the scan beam is a long way from the edge). Inside the silicon we can assume that the energy deposition profile takes the form of a double Gaussian, with a forward scattering width and a fraction of the energy forward scattered, and a backscatter width and a fraction of the energy deposited by those backscattered electrons. We will also assume that the number of secondary electrons that are generated within the wafer are in direct proportion to the energy deposited per unit volume, and the number of secondary electrons that escape the wafer (and so are detected by the SEM) are in direct proportion to the number of secondaries near the very top of the wafer.

The secondary electrons that reach the detector will emerge some distance *r* away from the position of the incident beam. From the assumptions above, the number of secondaries detected will be a function of

$$f(r) = ae^{-r^2/2\sigma_f^2} + be^{-r^2/2\sigma_b^2} \qquad\qquad \text{Eq. (1)}$$

where σ_f and σ_b are the forward and backscatter ranges, respectively. The integration of this function over all space will be the linescan signal at *x* = -∞.
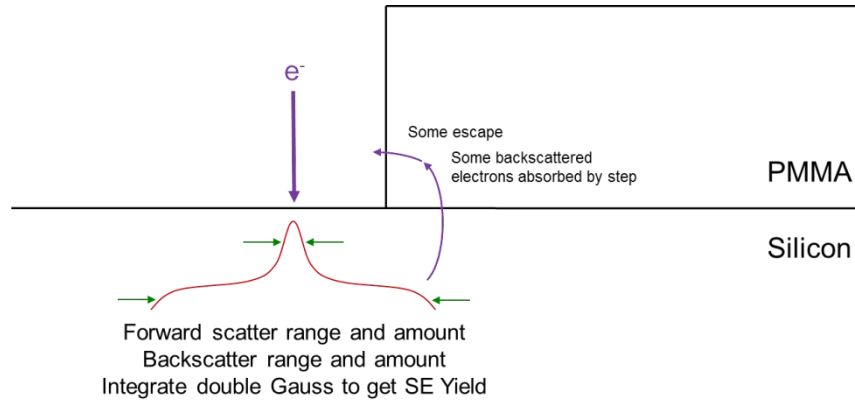
Figure 9: Geometry and physical interpretation of the analytical linescan model derived in this paper, showing a PMMA edge on a silicon wafer.

$$SE(-\infty) = \int_0^{2\pi} d\theta \int_0^{\infty} f(r)r\,dr = 2\pi\left(a\sigma_f^2 + b\sigma_b^2\right)$$

Eq. (2)

Now consider the effect of the step. Before looking at a silicon step, imagine a completely absorbing step that does not allow the release of secondary electrons. In other words, all electrons that travel up into the step material simply disappear. For $x < 0$, we can calculate the reduction in the SE signal by calculating the number of absorbed secondaries as

$$SE_{absorbed}(x) = 2\int_0^{\pi/2} d\theta \int_{r_0}^{\infty} f(r)r\,dr$$

Eq. (3)

where $r_0 = x/\cos(\theta)$. Carrying out the integration,

$$SE_{absorbed}(x) = \pi\left(a\sigma_f^2\, erfc\left(\frac{x}{\sigma_f\sqrt{2}}\right) + b\sigma_b^2\, erfc\left(\frac{x}{\sigma_b\sqrt{2}}\right)\right)$$

Eq. (4)

The linescan signal for $x < 0$ will be

$$\frac{SE(x)}{SE(-\infty)} = 1 - \frac{SE_{absorbed}(x)}{SE(-\infty)} = \left(\frac{1}{2} - \beta\right)erfc\left(\frac{x}{\sigma_f\sqrt{2}}\right) + \beta\, erfc\left(\frac{x}{\sigma_b\sqrt{2}}\right)$$

Eq. (5)

where $\beta = \dfrac{\pi b\sigma_b^2}{SE(-\infty)}$.

In what will follow, the complimentary error functions will prove cumbersome. However, since they are only used for $x < 0$ they can be approximated with exponentials with good-enough accuracy.

$$\text{For } x < 0, \quad erfc\left(\frac{x}{\sigma\sqrt{2}}\right) \approx 2 - e^{x/\sigma} \qquad \text{Eq. (6)}$$

giving

$$\frac{SE(x)}{SE(-\infty)} = 1 - \left(\frac{1}{2} - \beta\right)e^{x/\sigma_f} - \beta e^{x/\sigma_b} \qquad \text{Eq. (7)}$$

Let us see how this model matches the rigorous Monte Carlo simulations of the linescan for the case of a silicon edge at 500 V landing energy, using a point incident beam. Figure 10 shows the best fit of Eq. Eq. (7) to the simulation results. The best fit parameters are $\sigma_f = 1.73$ nm, $\sigma_b = 40.4$ nm, and $\beta = 0.22$. The equation matches the simulations for the most part within the noise of the simulations, and the resulting parameters make intuitive sense for the scattering of 500 V electrons in silicon.
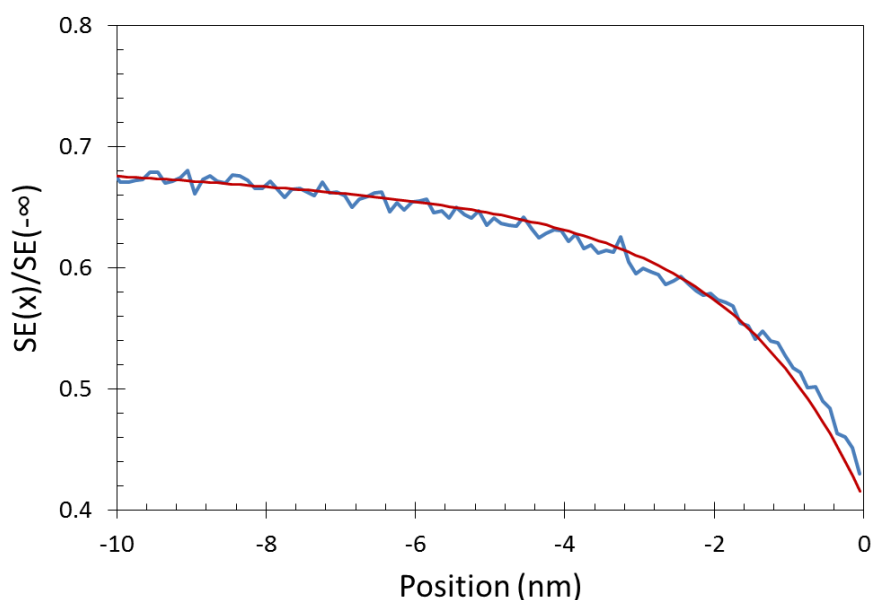
Figure 10. Best fit of Eq. Eq. (7) to rigorous Monte Carlo simulations of a silicon edge at 500 V landing energy, using a point incident beam, when $x < 0$ (that is, on the bottom half of the step). The smooth (red) line is the equation and the jagged (blue) line is the Monte Carlo simulation.

The assumption of an absorbing step is not completely accurate for silicon, and even less so for a step made of photoresist. In reality, backscatter electrons travel up into the step and generate secondaries that can escape from the sidewall of the step. As a result, the step may not absorb all of the electrons entering, and in fact may generate more secondaries than if no step were present. This can be accommodated by simply modifying the forward scatter and backscatter step absorption terms in Eq. Eq. (7).

$$\text{For } x < 0, \quad \frac{SE(x)}{SE(-\infty)} = 1 - \alpha_f e^{x/\sigma_f} - \alpha_b e^{x/\sigma_b} \qquad \text{Eq. (8)}$$

where α_f is the fraction of forward scatter secondaries absorbed by the step and α_b is the fraction of backscatter secondaries absorbed by the step. Note that if the presence of the step results in more secondary electrons, then the result will be a negative value of α_b.

The $x < 0$ linescan model of Eq. (8) was compared to the Monte Carlo simulations described in the previous section (using 5000 electrons per pixel) for the cases of an isolated silicon step on silicon and for a PMMA step on silicon, with landing voltages of 300, 500, and 800 V. The fits of the model to the simulations are shown in Figure 11, with the best-fit parameters shown in Table 1. As can be seen, the fits are within the random variations present in the Monte Carlo results.
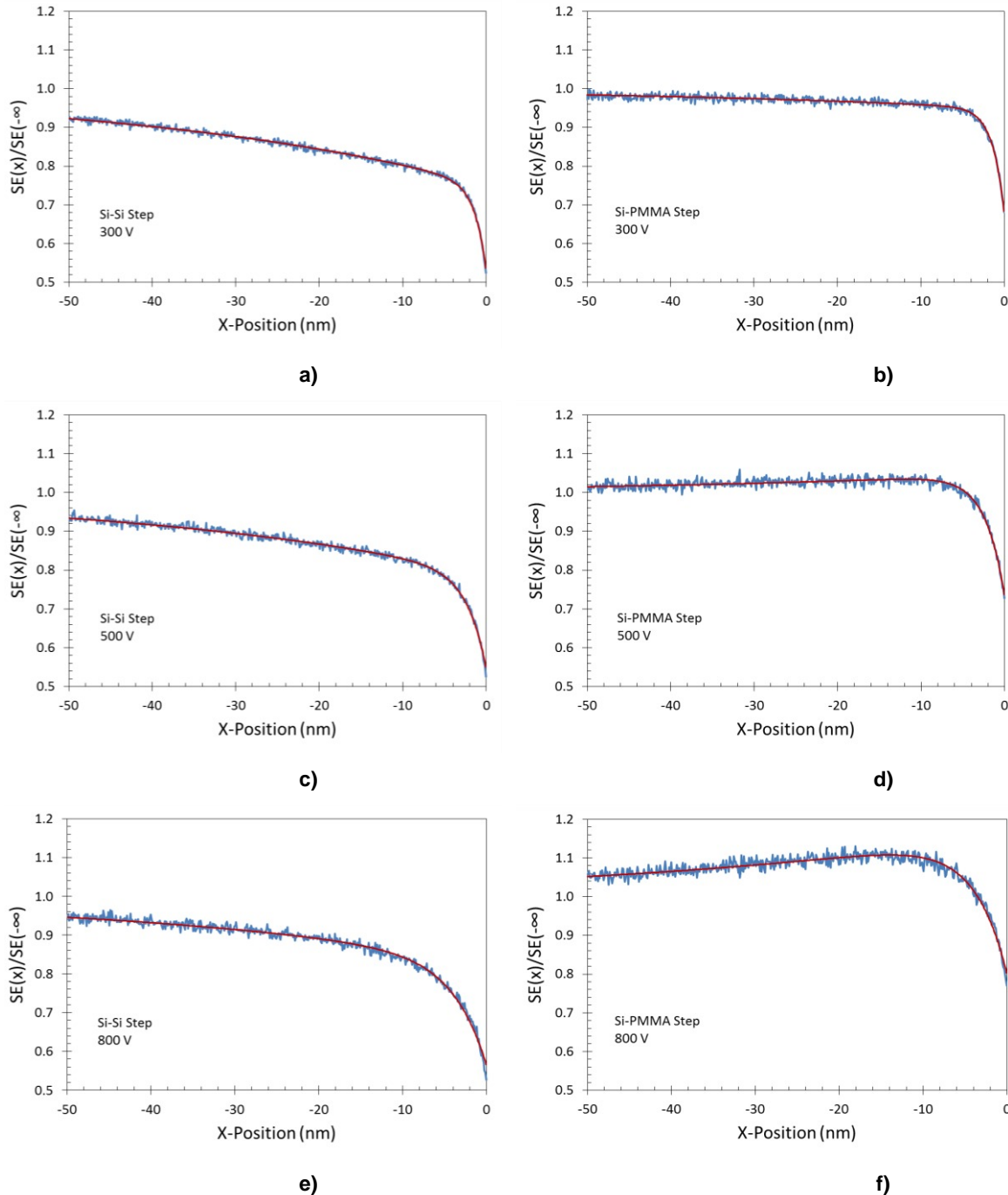

Figure 11. Best fit of Eq. Eq. (8) to rigorous Monte Carlo simulations (N = 5000) of an isolated silicon step (parts a), c), and e)) and PMMA step (parts b), d), and f)) on a silicon wafer at 300, 500, and 800 V landing energy, using a point incident beam, when $x < 0$ (that is, on the bottom half of the step). The smooth (red) line is the equation and the jagged (blue) line is the Monte Carlo simulation.

Table 1. Best fit parameters of Eq. Eq. (8) to rigorous Monte Carlo simulations of an isolated silicon step and PMMA step on a silicon wafer at 300, 500, and 800 V electron energy, using a point incident beam, for $x < 0$.

	Electron Energy		
	300 V	**500 V**	**800 V**
Si wafer background signal, $SE(-\infty)$	1.070	0.818	0.597
Si forward scatter range, σ_f (nm)	1.37	2.31	3.96
Si backscatter range, σ_b (nm)	43.5	43.0	43.7
Si step forward scatter absorption, α_f	0.223	0.243	0.267
Si step backscatter absorption, α_b	0.251	0.212	0.171
PMMA step forward scatter absorption, α_f	0.272	0.316	0.363
PMMA step backscatter absorption, α_b	0.055	-0.047	-0.162

When $x > 0$, the incident beam is on top of the step. A model similar to that described for the bottom of the step can be applied to the top, but with some differences. The edge of the step does not absorb the long-range backscattered electrons, and in fact enhances the release of secondaries created by the forward scattered electrons. Thus, we will add a positive term $\alpha_e e^{-x/\sigma_e}$ to account for the enhanced escape of forward-scattered secondaries where σ_e is very similar to the forward scatter range of the step material. When the incident beam is very close to the step, however, the interaction volume of the forward-scattered electrons with the material is reduced, causing the generation of less secondaries. Thus, we subtract a term $\alpha_v e^{-x/\sigma_v}$ where $\sigma_v < \sigma_e$. This gives our linescan expression for the top of the step:

$$\text{For } x > 0, \quad \frac{SE(x)}{SE(\infty)} = 1 + \alpha_e e^{-x/\sigma_e} - \alpha_v e^{-x/\sigma_v} \quad \text{Eq. (9)}$$

Combining the $x < 0$ and $x > 0$ linescan expressions into one expression (using the unit step function, $u(x)$), we have a final linescan expression for the case of a point incident beam.

$$SE(x) = SE(-\infty)\left[1 - \alpha_f e^{x/\sigma_f} - \alpha_b e^{x/\sigma_b}\right]u(-x) + SE(\infty)\left[1 + \alpha_e e^{-x/\sigma_e} - \alpha_v e^{-x/\sigma_v}\right]u(x) \quad \text{Eq. (10)}$$
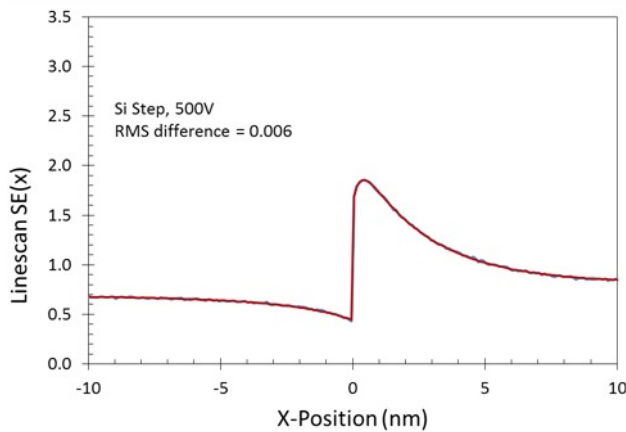
Figure 12 shows the best fit of Eq. Eq. (10) to the Monte Carlo simulations for both silicon and PMMA steps on a silicon wafer at 300, 500, and 800 V electron voltages. In addition to the parameters shown in Table I (which were not modified), the new best-fit parameters of Eq. Eq. (10) are shown in Table 2.
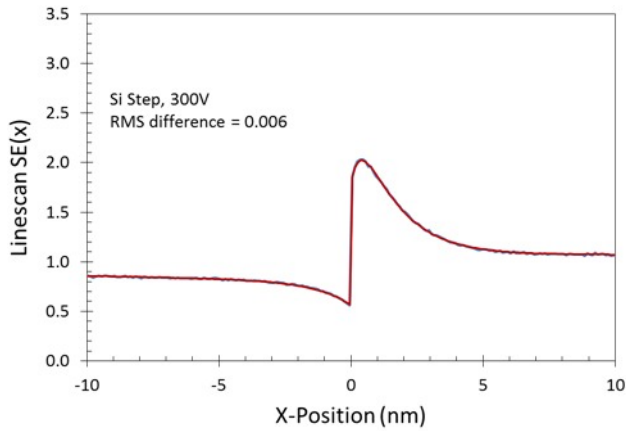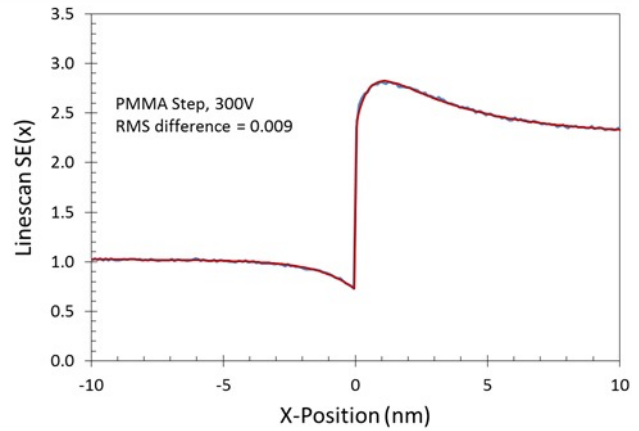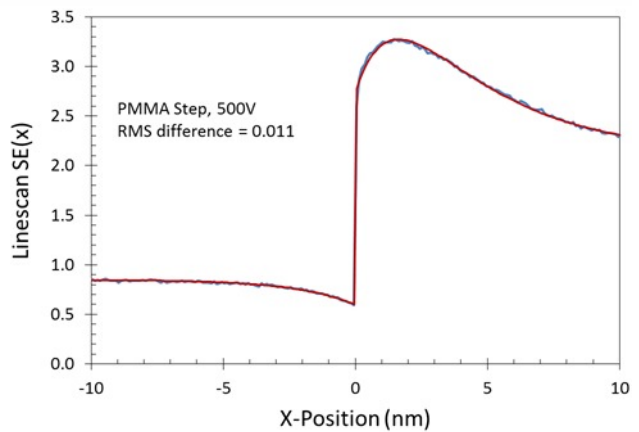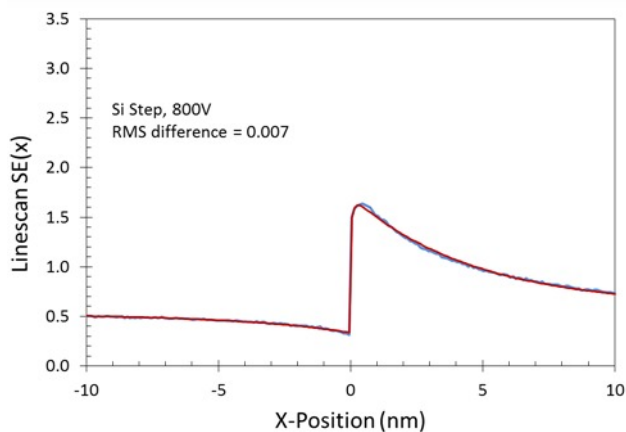
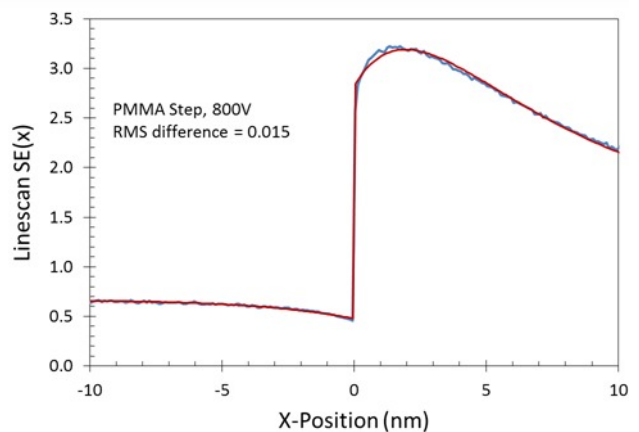Figure 12.   Best fit of Eq. Eq. (10) to rigorous Monte Carlo simulations of an isolated silicon step (parts a), c), and e)) and PMMA step (parts b), d), and f)) on a silicon wafer at 300, 500, and 800 V landing energy, using a point incident beam. The smooth (red) line is the equation and the jagged (blue) line is the Monte Carlo simulation.

Table 2. Best fit parameters to rigorous Monte Carlo simulations of an isolated silicon step and PMMA step on a silicon wafer at 300, 500, and 800 V electron energy, using a point incident beam.

	Electron Energy		
	300 V	**500 V**	**800 V**
Si step forward scatter range, σ_e (nm)	1.47	2.64	4.65
Si step volume loss range, σ_v (nm)	0.40	0.27	0.11
Si step edge enhancement factor, α_f	1.61	1.65	1.87
Si step volume loss factor, α_v	0.924	0.660	0.543
PMMA step forward scatter range, σ_e (nm)	3.42	3.54	5.28
PMMA step volume loss range, σ_v (nm)	0.53	1.43	2.39
PMMA step edge enhancement factor, α_f	0.382	1.427	2.784
PMMA step volume loss factor, α_v	0.347	1.143	1.968
PMMA background signal, $SE(\infty)$	2.286	2.132	1.552

Real scanning electron microscopes do not have point beams of electrons impinging on the sample. Instead, the beam is approximately a Gaussian owing to the finite resolution of the microscope and other beam non-idealities. Thus, the expected linescan will be Eq. Eq. (10) convolved with a Gaussian. Carrying out this convolution gives

$$
SE(x) = \frac{SE(-\infty)}{2}\left[erfc\left(\frac{x}{\sqrt{2}\sigma_p}\right) - \alpha_f e^{\sigma_p^2/2\sigma_f^2} e^{x/\sigma_f} erfc\left(\frac{\sigma_p^2 + x\sigma_f}{\sqrt{2}\sigma_p\sigma_f}\right) - \alpha_b e^{\sigma_p^2/2\sigma_b^2} e^{x/\sigma_b} erfc\left(\frac{\sigma_p^2 + x\sigma_b}{\sqrt{2}\sigma_p\sigma_b}\right) \right]
$$
$$
+ \frac{SE(\infty)}{2}\left[erfc\left(\frac{-x}{\sqrt{2}\sigma_p}\right) + \alpha_e e^{\sigma_p^2/2\sigma_e^2} e^{x/\sigma_e} erfc\left(\frac{\sigma_p^2 - x\sigma_e}{\sqrt{2}\sigma_p\sigma_e}\right) - \alpha_v e^{\sigma_p^2/2\sigma_v^2} e^{x/\sigma_v} erfc\left(\frac{\sigma_p^2 - x\sigma_v}{\sqrt{2}\sigma_p\sigma_v}\right) \right]
$$

Eq. (11)

where σ_p is the standard deviation (width parameter) of the Gaussian electron beam probe. Monte Carlo simulations using a Gaussian beam with $\sigma_p = 0.5$ nm can now be compared with Eq. Eq. (11). Without any adjustment of the parameters given in Tables 1 and 2, excellent fits were obtained. A representative example is shown in Figure 13 for the PMMA step and a 500 V beam (RMS error of fit to Monte Carlo simulation is 0.01).

The final linescan model of Eq. Eq. (11) has 11 physically meaningful parameters (or 10 for the case of a silicon step on a silicon wafer since SE(∞) = SE(-∞)). For example, the pair σ_e, α_e represent the range and amount of forward-scattering generated secondaries in the PMMA step that escape from the step edge. They control the height of the secondary electron peak near the step edge and the rate of fall-off as the beam moves to the right.
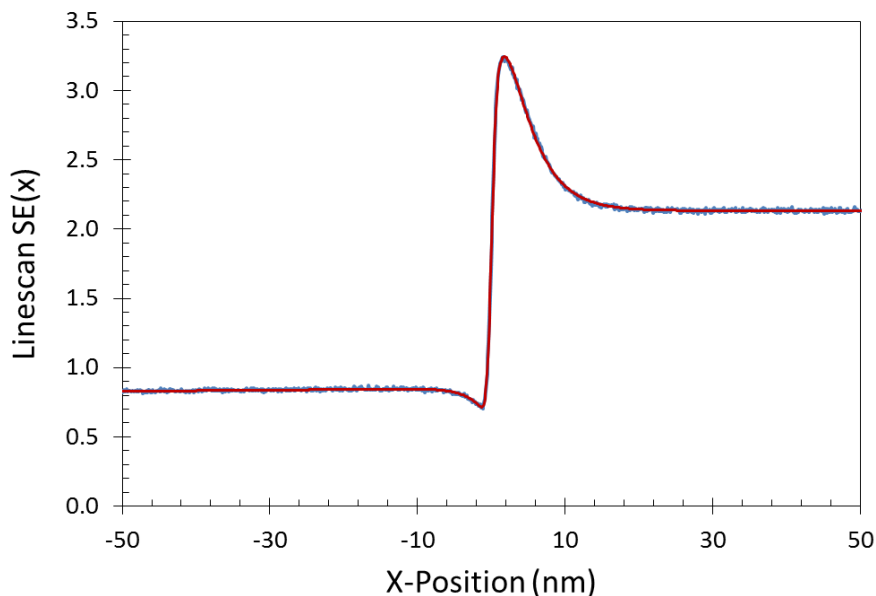
Figure 13.   Comparison of the linescan model of Eq. Eq. (11) to rigorous Monte Carlo simulations (N = 5000) of an isolated PMMA step on a silicon wafer at 500 V landing energy, using a 0.5 nm Gaussian incident beam. The smooth (red) line is the equation and the jagged (blue) line is the Monte Carlo simulation. RMS error of fit to Monte Carlo simulation is 0.01.

4.   Application of Linescan Model to PSD Measurement

In a previous publication on PSD measurement [5], the impact of the electron beam spot and its interaction with the materials on the wafer was assumed to be a form of metrology averaging over y (along the line length). A Gaussian-shaped beam of electrons was assumed to interact with the feature being measured to produce a Gaussian-shaped measurement signal (wider than the incident beam) of FWHM η. The impact of this averaging can be seen in Figure 14 using the metrology simulator from Ref. [5], and is a function of $\eta/\Delta y$ (where Δy is the sampling distance for the measurement). For no averaging ($\eta = 0$), aliasing makes the measured PSD higher at the high frequencies. Averaging lowers the measured PSD at high frequencies, thus reducing the impact of aliasing. However, for $\eta > \Delta y/2$ the impact of averaging is greater than aliasing, and the measured PSD is suppressed at high frequencies. (Note that the simulations shown in Figure 14 assumed no metrology noise, and thus $\sigma_\varepsilon = 0$.)

A goal of this work is to understand the whether the assumption of a Gaussian-shaped averaging along y is a good one, and if so, what is an appropriate Gaussian FWHM (η). As the above work has shown, a point electron beam incident on the top of a PMMA step produces essentially a Gaussian sampling of the feature with a width parameter σ_e, the PMMA forward scattering range. This Gaussian feature interaction range is then convolved with the Gaussian beam with a width parameter σ_p. Thus, the FWHM of the Gaussian averaging function will be

$$\eta \approx 2.355\sqrt{\sigma_e^2 + \sigma_p^2} \qquad\qquad \text{Eq. (12)}$$

For example, the 500 V PMMA step modeled above showed $\sigma_e = 3.5$ nm (Table II). If $\sigma_p \approx 1.0$ nm (a typical value), the resulting FWHM of the Gaussian averaging will be $\eta = 8.6$ nm. Since this value is larger than half of the sampling distance typically used for LER measurement, the result will be suppression of high frequency roughness, as seen in Figure 7. In fact, an optimum sampling distance for this case will be about $\Delta y = 2\eta \approx 17$ nm.
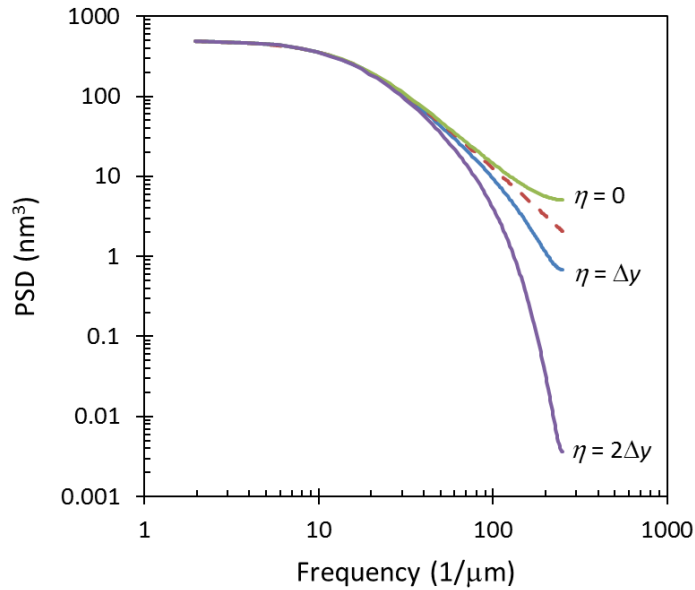
Figure 14. Simulations of the impact of averaging on the measured PSD (number of measurement points $N = 256$, correlation length $\xi = 10$ nm, roughness exponent $H = 0.5$, sampling distance $\Delta y = 2$ nm, rectangular measurement window). The FWHM of the Gaussian measurement signal (η) is varied from 0 to twice the sampling distance. The continuous PSD (without aliasing, leakage or averaging) is shown as the dotted line. [Figure from Ref. 5].

5.   Conclusions and Future Work

In this work, JMONSEL simulations with edge algorithms were used to explore random and systematic error sources in CD-SEM metrology. Offset and σ_ε were found to vary strongly with SEM beam spot size and energy, materials used, feature sidewall angle (SWA), TCR, but interestingly, not footing. The Monte Carlo simulations allow us to explore σ_ε, although noise will need to be supplemented with a term to represent other real-life noise sources such as electronic noise, environmental electromagnetics, vibrations, contamination, charging, or others. Additionally, offsets were also shown to vary strongly with many of these parameters, and if some of these parameters vary along the length of a line, such as SWA, TCR or material homogeneity, additional apparent LER might result after edge topography points are assigned.

An analytical linescan function was derived based on physical parameters, which fits the Monte Carlo results very well for the Si or PMMA step on Si. These best-fit parameters allowed us to extract an estimate of the beam/resist interaction width, giving $\eta \approx 8.6$ nm at 500 V. These findings are important considerations in properly analyzing PSD. Knowing FWHM of beam/resist interaction allows us to set up an optimal LER/LWR sampling distance. The ideal sampling distance should be double the beam/resist interaction width. For the cases explored here (an ideal 50 nm resist edge on silicon) the optimal LER sampling distance was found to be about 16–25 nm, depending on the beam energy. This is bigger than expected, and definitely larger than the current value recommended in the SEMI standard, which is 10 nm.

The work presented here is not complete. Detailing the error sources in SEM metrology is an important step in ultimately determining error estimates for all of the parameters extracted from a PSD. Other plans for this work include extension of the analytical linescan function from a single step-edge to a full line, the impact of resist thickness and sloped profile, and exploration of use of this function as an improved edge detection algorithm tailored for such cases.

6.   Acknowledgements

7.   References

[1]   The International Technology Roadmap for Semiconductors, 2012 (San Jose: Semiconductor Industry Association); http://www.itrs.net.

[2]   B. Bunday, T. Germer, V. Vartanian, A Cordes, A. Cepler & C. Settens. "Gaps Analysis for CD Metrology Beyond the 22 nm Node", Proc. SPIE, v8681, pp 86813B (2013).

[3]   Benjamin Bunday, Victor Vartanian, Abraham Arceo, and Aaron Cordes. "Evolution or Revolution: Defining the Path for Metrology Beyond the 22nm Node". Solid State technology, vol 55, issue 2, March 2012. Available online at http://www.electroiq.com/articles/sst/print/vol-55/issue-2/features/metrology/evolution-or-revolution.html

[4]   SEMI Standard P47-0307, "Test Method for Evaluation of Line-Edge Roughness and Linewidth Roughness", SEMI, San Jose, CA (2006).

[5]   C. A. Mack, "Systematic Errors in the Measurement of Power Spectral Density", *Metrology, Inspection, and Process Control for Microlithography XXVII*, *Proc*., SPIE Vol. 8681 (2013).

[6]   C. A. Mack, "Systematic Errors in the Measurement of Power Spectral Density", *Journal of Micro/Nanolithography, MEMS, and MOEMS*, Vol. 12, No. 3, p. 033016 (Jul-Sep, 2013).

[7]   B. Bunday, J. Villarrubia, A. Vladar, R. Dixson, T. Vorburger, N. Orji, and J. Allgair, "Determination of Optimal Parameters for CD-SEM Measurement of Line Edge Roughness." SPIE 2004, v5375, pp 515-533, 2004.

[8]   Villarrubia, J. and Bunday, B. "Unbiased Estimation of Linewidth Roughness". Proceedings of SPIE 2005, v5752, pp 480-488.

[9]   Villarrubia, J. S. , Ritchie, N. W. M., and Lowney, J. R. "Monte Carlo modeling of secondary electron imaging in three dimensions," Proc. SPIE **6518**, 65180K (2007).

[10]   Villarrubia, J. S., and Ding, Z. J. "Sensitivity of SEM width measurements to model assumptions," J. Micro/Nanolith. MEMS MOEMS **8**, 033003 (2009).

[11]   Aron Cepler, Benjamin Bunday, Bradley Thiel, John Villarrubia. "Scanning electron microscopy imaging of ultra-high aspect ratio hole features". Metrology, Inspection, and Process Control for Microlithography XXVI. Proceedings of the SPIE, Volume 8324, pp. 83241N-83241N-14 (2012).

[12]   C. A. Mack, "Generating random rough edges, surfaces, and volumes", *Applied Optics*, Vol. 52, No. 7 (1 March 2013) pp. 1472-1480.

[13]   Villarrubia, J. S., Vladár, A. E., and Postek, M. T. "Simulation study of repeatability and bias in the critical dimension scanning electron microscope," J. Microlith., Microfab., Microsyst. 4(3), 033002(Jul–Sep 2005).