NIST Simulation of E-beam Inspection and CD-SEM in-line metrology

Final Report

International Sematech Manufacturing Initiative

Agreement Number: 100193-MM

Project Number: LITG440-N2

Project Title: Defect Metrology

Project Subtitle: Simulation study for CD-SEM and EBI tools

NIST Simulation of E-beam Inspection and CD-SEM in-line Metrology, Final Report*

International Sematech Manufacturing Initiative

Agreement Number: 100193-MM

Project Number: LITG440-N2

Project Title: Defect Metrology

Project Subtitle: Simulation study for CD-SEM and EBI tools

December 10th, 2010

Executive summary:	This report summarizes results from a two-year project to develop a simulator for electron beam inspection and critical dimension scanning electron microscope (SEM) inline metrology tools. The development attempts to improve on prior simulators and develop a tool that can be used to explore the fundamental limits of e-beam instruments for imaging and measuring small features. The work covered by this report includes improvements to the secondary electron generation model, assembly of a database of the material properties and computation of scattering tables required for use of the improved model, improvements to simulation algorithms for non-conducting materials, and the addition of a capability to model the effects of charging in insulating samples.
Keywords:	e-beam, charging, critical dimension, modeling
Author:	John Villarrubia (NIST)

* Certain commercial equipment and software is identified in this report to adequately describe the experimental procedure. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that the equipment or software identified is necessarily the best available for the purpose.

Table of contents

[Create TOC based on final draft of report]

1.0 Introduction

Electron-beam-based tools are used by the semiconductor electronics industry for failure analysis, critical dimension and contour metrology, defect inspection and classification, and lithography (particularly for mask production). Measuring techniques based upon other physical principles, e.g., atomic force microscopy (AFM) and optical methods (microscopy and scatterometry), are also put to good use in many applications. It is sometimes implied by an enthusiastic proponent of one or another of these methods that the favored method will supplant the others entirely. Such claims seem to us injudicious. These methods appear to have distinct advantages: The inherent three spatial dimensions in an AFM image and the relative simplicity (at length scales above few nanometers) of its interaction mechanism have given AFM a role as a reference metrology. Scatterometry has greatly improved the measurement resolution of optical dimensional metrology, at least when applied to periodic arrays, allowing the inherent speed and absence of contamination of optical techniques to preserve a place for them in in-line metrology and inspection.

Without discounting the strengths of other methods, it seems appropriate at the beginning of a report on e-beam tools to dwell briefly on *their* particular strengths. The speed of e-beam tools is intermediate between AFM and optics, but the speed is much closer to the fast optical methods than to the slower AFM. On the other hand the spatial resolution of electron beams, even when used to generate secondary electrons (SE), which is not their highest spatial resolution mode, is on the order of a few nanometers, comparable to that of AFM (even if not three-dimensional) and much better than the native resolution of visible-light optical methods (given, e.g., by the classical Rayleigh criterion, as comparable to the wavelength).*

The combination of speed and spatial resolution is valuable in any of the many circumstances where one wants as rapid a measurement as possible on an individual feature, for example a defect, an individual transistor gate (how does it differ from the others?), or roughness. These tasks all become more challenging as feature sizes shrink and complexity grows. Looking to the future, the question should not be so much *which* of these measurement techniques can provide the solution as whether *all* of them, used in combination, will continue to be able to do so.

---

* By spatial resolution we refer to the characteristic broadening expected for a hypothetical 0-width feature, such as a sharp edge. This broadening is comparable to the characteristic interaction length of the probe (stylus, electron, or photon) with the sample. This sets the minimum spatial measurement uncertainty that can be *readily* obtained. These two concepts, spatial resolution and uncertainty in a spatial measurement, are often confused. In any of the techniques, measurement uncertainty may be improved beyond the spatial resolution through the use of averaging and modeling, a fact demonstrated most notably in scatterometry but also, perhaps less familiarly, by model-based measurement techniques in AFM and SEM.

Most of the imaging applications for electrons in semiconductor electronics use SE. In principle, SE are sample (i.e., not beam) electrons that gain energy in collisions with other high energy electrons. Some of these generated near the surface of the sample can escape and be detected. These electrons are mainly low energy, less than 50 eV. It is not possible to distinguish one electron from another, so for practical purposes SE are usually *defined* by their energy. Those with energy less than 50 eV are considered SE. Those with greater energies are called backscattered electrons (BSE). SE are produced in a cascade; beam electrons produce SE that may in turn produce other SE. SE are the preferred signal in most semiconductor electronics imaging applications because their greater number and easier collection (due to their low energy) provides a better signal level, which leads to higher measurement throughput. The characteristic size scale for this kind of interaction of electrons with solid matter is set by the size of the volume within which there is a significant number of energized electrons. This volume is larger for a higher energy beam and materials with a low density of lightly bound electrons than for a lower energy beam or higher density material. In silicon, for the beams below 1 keV that are now most common in semiconductor electronics imaging applications, this size is of the order of 10 nm. This is already uncomfortably close to the sizes of transistor gates, assist features, and most other objects of interest—and it is already much larger than the desired *uncertainties* with which we would like to measure the dimensions of such features.

The closeness of the size of the interaction volume to feature size results in imaging artifacts. "Image artifact" connotes some kind of imaging error. More precisely, an image artifact is a difference between the image and our expectations. The latter are usually based on a crude "what we see is what we get" model. Such a model may serve adequately when there is plenty of spatial resolution to spare, but it becomes increasingly problematic when we put more stringent demands on the measurement technique. Some examples will be given in the next section. It seems there are only two possible solutions to this problem: either a better measurement technique (i.e., with spatial resolution that is once again small compared to the desired uncertainty) or better interpretation of the existing technique, so that we are not so easily fooled by image artifacts. Motivation and a description of the latter approach applied to e-beam imaging tools are described in the next section.

2.0 Background of the project and JMONSEL

2.1 Project motivation

When features of our samples (or our required uncertainties) become comparable to the probe-sample interaction size, the data provided by the measuring tool is no longer enough to answer many of our questions. A model is also required. In the following paragraphs, we discuss three categories of questions that require models to answer: questions about image interpretation (especially quantitative interpretation), optimization, and limits.

Consider some common examples of image interpretation drawn from semiconductor electronics applications: (1) The CD (critical dimension) is a width, the distance between two edges. We associate the positions of features on the sample with the locations of corresponding intensity changes in the image; thus edges are associated with the bright edge bloom in an image (Fig. 1). But a sharp edge may nevertheless cause some brightening in the image over a distance of 10 nm or more. Where within this bright region is the actual edge? The height of the line, its composition, the angle (and detailed shape) of the edge, can affect the intensity profile, so affect the answer

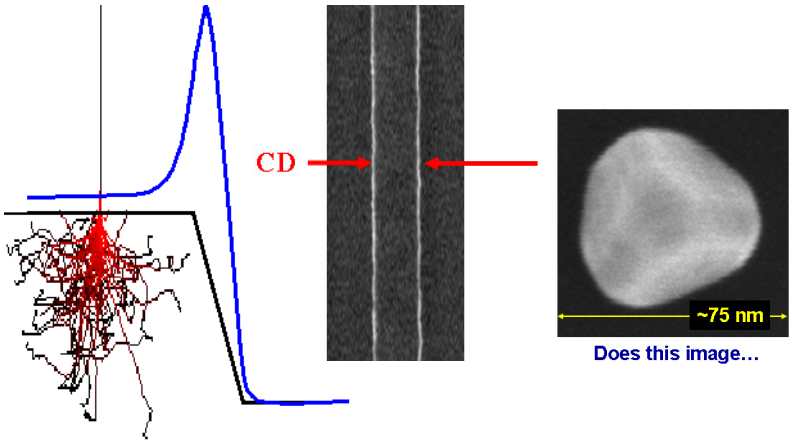

Fig. 1. Edge bloom: When the interaction volume intersects the edge, SE can escape from the side of a line in addition to the top. The increased brightness marks the edge locations, but the finite width makes their exact positions somewhat ambiguous.

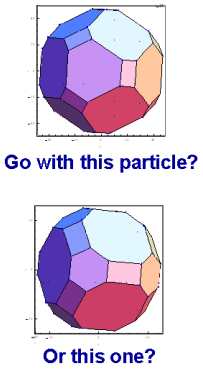

Fig. 2. Determining particle structure from an image requires comparing the measured image with the image that would be expected for likely structures.

to that question. (2) In contour metrology the goal is to measure the boundary of a complex feature (e.g., a transistor or assist feature on a mask) to make sure it is within the design specification. The neighborhood of such a boundary is rarely uniform. If the boundary is relatively isolated along part of its length, but has nearby neighbors in another part, the brightness of the image contrast depends in part on the differing likelihoods that electrons leaving the edge can escape recapture by the nearby (or not so nearby) obstacles. An apparent shift in the edge might be all or partly due to such proximity effects. (3) What structure(s) are consistent with a given observed image (Fig. 2)? The relevant question is in each case concerned with the consistency between observation and a particular state of the sample and instrument. The process of answering the question is some variation of this: If the edge were at x_1 (if the particle were faceted like A) what should the image look like? If it were at x_2 (faceted like B) what should it look like? Etc. Which of these possibilities are ruled out by the observation? Which are not? The observation half of these comparisons comes from a measurement. The answer to the "what would it look like" question must come from a model.

Optimization questions arise because we have choices in the settings at which we operate e-beam measurement tools. The choices include landing energy, spot size, scan speed, etc. Different applications dictate different choices. One application might call for optimizing an electron beam inspection (EBI) tool for topographic contrast (e.g., if the most likely defects manifest themselves as bumps or holes). Another might require optimizing for contrast between two particular materials (e.g., if the defect is composed of one and the designed structure the other). It may be possible to trade resolution for speed, but to decide the optimum point it is necessary to know details of the trade-off. In a quantitative metrology application, what landing energy minimizes measurement uncertainty? To answer optimization questions such as these, one needs to know how changing the instrument settings would change the result of a measurement.

Questions about limits include these: Is an open contact detectable by a particular method, or would a closed one look much the same? If the EBI tool detects no particle defects with size less than x nm, is that because there are no such defects or because such defects are not detectable? Are buried structures detectable? How close to the surface must they lie to be so? When we are finally able to make a gate with

CD < x nm (where x is not yet possible with today's technology but expected to be possible soon) will we be able to measure it? Would a particular redesign of the instrument improve the situation? Questions like these are often difficult to answer experimentally because they involve too many unknowns (e.g., unknown sample and unknown instrument response) relative to the amount of data or they involve hypothetical situations (e.g., future states of sample or instrument) that are not attainable today.

2.2 Background

2.2.1 The phenomena to be modeled

A number of different phenomena combine to produce the signals we measure in e-beam imaging tools. A simulator needs to model all of the important phenomena. Following is a list of the most important ones along with a description of their influence on the measured signals.

Elastic scattering: Elastic scattering (Fig. 3a) refers to scattering events in which the primary electron loses negligible energy. Scattering of electrons from atomic nuclei are events of this type. The nucleus of even a light atom is thousands of times heavier than the electron; even if the electron were to transfer the maximum momentum to the atom, the large mass difference insures that the kinetic energy imparted thereby would be small compared to the electron's energy. Elastic collisions are the most important source of large-angle scattering. Without such scattering, the beam electrons would brake in almost a straight line; the interaction volume would be a thin "pencil" within the sample along the beam axis. Instead, elastic scattering of the electrons in all directions results in a much broader interaction volume (Fig. 3b). A minority of electrons are backscattered in the direction from which they came. These can escape from the sample with high energy, and SE generated by them near the surface contribute to the SE signal. If the interaction volume intersects a boundary of the sample, some of those electrons can escape and be detected. Thus, this phenomenon helps determine the broadening of the signal at an edge.

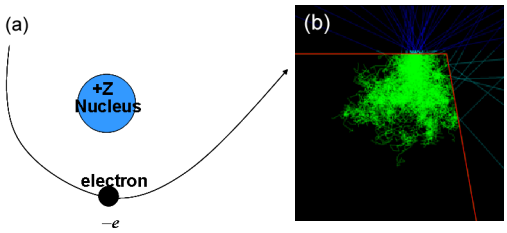

Fig. 3. (a) Elastic scattering schematic and (b) trajectories (green) inside a material, showing the interaction volume.

Secondary electron (SE) generation: Inelastic scattering events are those in which the incident electron loses a non-negligible amount of energy. Sometimes the energetic electron transfers energy to one of the sample's electrons (Fig. 4). These are usually the most important inelastic scattering events, events in which the primary electron loses significant energy. The SE may acquire enough energy to escape the sample, and they can produce more SE by scattering in their turn. This cascade is mainly a low energy phenomenon. The most likely energy transfer in this process is usually in the range of 20 eV to 50 eV, and the cross section is highest for primary electron energies below 100 eV. When the primary electron has energy large compared to the energy loss, these events do not result in a significant change in direction. Thus, these events do not so strongly affect the shape of the interaction volume as do the elastic events described in the previous paragraph. SE significantly outnumber the incident electrons. They are the reason electron yields greater than 1 (more electrons coming out of the sample than going in) are possible

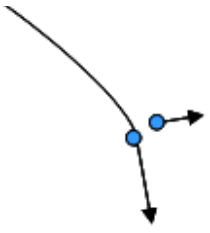
Fig. 4. SE generation schematic.

at favorable incident energies. Because of their larger numbers, SEs are the favored signal in many applications. The familiar images from a CD-SEM or EBI tool are SE images.

Slowing down of electrons: The beam electrons and their secondary offspring lose energy as they pass through matter. Eventually they have lost so much energy that they are indistinguishable from other electrons in the material. This loss of energy gives them a finite range. That range in turn limits the size of the interaction volume. In practice the interaction volume is somewhat smaller than the electron range in the material because elastic scattering means the electrons do not follow the shortest possible path to their eventual stopping point. SE generation is usually by far the most important energy loss process. For this reason we could have considered slowing down of electrons as an aspect of SE generation discussed in the previous paragraph. We give it its separate paragraph because in practice most e-beam interaction models have treated SE generation and slowing down separately. It is possible to treat them separately because there are well-known reasonably simple expressions for the stopping power (average energy loss per unit distance of travel) of materials. The most well known was first published by Hans Bethe in 1930.[1,2] At the time of its development, stopping powers were mainly of interest for nuclear radiation, cosmic rays, and other high energy (tens of keV at least, more often MeV) applications. While a good approximation for these cases, Bethe's approximation breaks down at the lower energies that are important for EBI and SEM. Later, Joy and Luo [3] and Rao-Sahib and Wittry [4] extended Bethe's formula to lower energies. Stopping powers for copper in these approximations and as computed by the dielectric function theory implementation used by JMONSEL are shown in Fig. 5. Some models use one of these or another approximation of the average energy loss to estimate the electron's energy after it has traversed a given distance. Because the energy in this approximation is lost continuously, it is known as the continuous slowing down (CSD) approximation. The advantage of the CSD approximation is simplicity. The electron's energy can be assigned even without a detailed model that describes the probability of losses of various sizes. The disadvantage is that the model is necessarily a description of average behavior; it cannot account for statistical deviations from the average, e.g., electrons with larger or smaller range than one would expect for their energy.

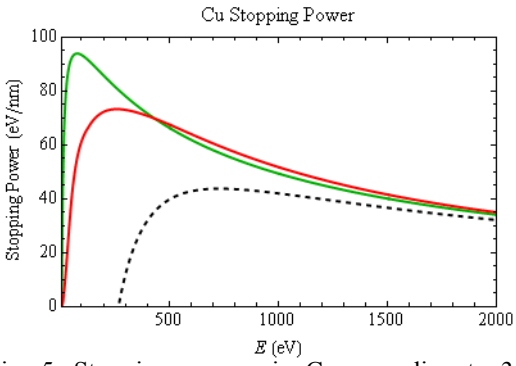

Fig. 5. Stopping powers in Cu according to 3 different approximations: Bethe (dashed, black), Joy & Luo (green) and JMONSEL's implementation of dielectric function theory (red).

Interface scattering: Electrons in different materials generally have different amounts of potential energy. Consequently, at an interface between materials, there is a potential energy step (Fig. 6). The potential energy in a conductor is always lower than in the surrounding air or vacuum. This is why the electrons remain trapped within the conductor. They must go "uphill" to escape. Electrons that move down the step (e.g., from vacuum into a metal) convert the potential energy difference into kinetic energy. Electrons going in the other direction

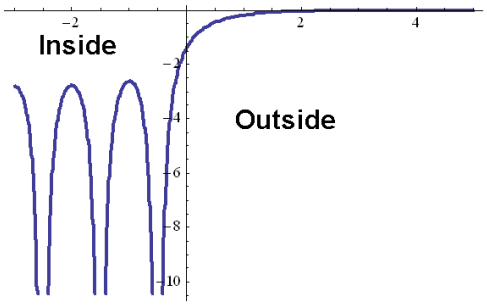

Fig. 6. Schematic of potential energy change at a material-vacuum interface.

lose an equivalent amount of kinetic energy. Either way, the electron's direction changes in a process closely analogous to refraction of light. Step heights are typically around 10 eV, and these phenomena become significant for electrons that have comparable energy. Consequently interface scattering is often ignored in models that concentrate on BSE or other high energy phenomena. However, as we have already discussed, SE have typical energies below 50 eV when they are generated, and they lose energy as they traverse the sample. For them, interface scattering significantly affects the probability of transmission and detection.

Charging and electric fields: In insulating materials, charge cannot flow to make up for a deficit of electrons (e.g., in regions where the initial yield was large) or an excess (e.g., in the interior of the sample, where beam electrons finally come to rest). Excess positive or negative charge generates electric fields that affect the paths and energies of nearby moving electrons. So, for example, an incoming electron can be deflected from the position at which it would otherwise have landed. It can accelerate or decelerate, so land with a different energy. Since yield varies with energy, this will in turn alter the yield. SE that escape through a surface need no longer escape the sample entirely. They may be turned around and recaptured if the local fields are large enough and of the right sign (Fig. 7). Even apart from electric fields induced by charging, external fields may be applied as part of the electron optical system, or to facilitate collection of SE.

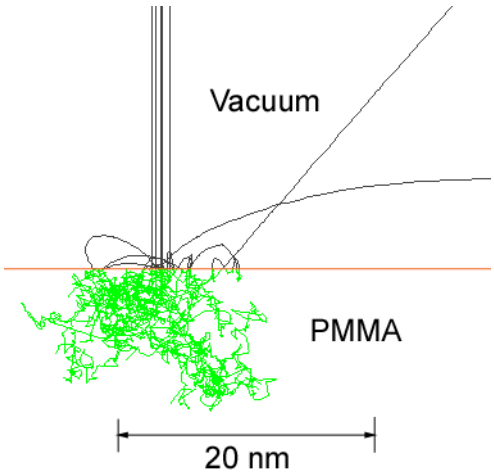

Fig. 7. Electron trajectories in PMMA, simulated by JMONSEL with charging. Note how some low energy SE return to the material.

Other insulator phenomena: We said above that SE generation is usually by far the most important inelastic scattering mechanism. Large-bandgap insulators provide an exception. If E_g is the bandgap, electrons in the sample cannot gain energy less than E_g. Therefore, energetic electrons cannot lose energy less than this via an SE generating mechanism. If SE generation were the only inelastic mechanism, primary electrons with energy less than E_g would never lose energy at all, and their range would be infinite. If such electrons have enough energy to escape the sample (i.e., more than the interface barrier), they will all eventually escape. When the SE generation rate goes to zero, it is unrealistic to continue to assume that competing processes remain relatively small. Otherwise small energy losses, e.g., losses to optical phonons, become important. Insulators may also have polarons or other states that can trap electrons and reduce yields.

2.2.2 Prior work

Prior e-beam imaging simulators usually employ one of three methods: a Monte Carlo method, a diffusion-like approximate solution of the Boltzmann transport equation, or an empirical method (i.e. a method based upon a function chosen to give a reasonable description of measurement results). Of these, the Monte Carlo methods represent the more relevant forebears of the simulator (JMONSEL) under development in the present project. It seems to us that Monte Carlo methods are often slower than the alternatives but have the advantages of simplicity, accuracy, and generalizability.

Simplicity: Contrast in an e-beam image is determined by yield variations across the scanned area. The probability that a given incident electron will backscatter and/or produce one or more SE is an average over many alternative particle histories. It can scatter first here or there, produce an SE with this energy or that one, heading in direction A or direction B. These possibilities describe the first event. For each outcome of the first event, there is a set of possible outcomes of the second event, then the third, and so forth. In a Monte Carlo simulation, the process is broken down into individual steps, much as just described. Each step is characterized by its own probability distribution, which is simulated by using the computer's random number generator. In this way, each step is rendered relatively simple; the inherently complex features of the problem, which arise from the infinity of possibilities at each step and the probability density convolutions required by multiple scattering are dealt with by a statistical sampling instead of by attempting an explicit averaging.

Accuracy: The sampling can be made to approach the correct value (i.e., the averages implied by particular choices of physics for the individual events) as closely as desired by making the number of samples large enough. That is, the Monte Carlo methods require only minor approximation beyond the unavoidable (and therefore common to all physics-based models) approximation entailed in choosing among the often uncertain physics for the individual events.

Generalizability: Consider a method based upon an empirical rule, such as that the escape probability for an electron generated a distance d beneath a planar surface is $0.5\exp(-d/\lambda)$ where λ is a material-dependent constant. How is this to be generalized if the surface is not planar? In that case, presumably, we must take into account the different lengths of the various escape paths. Perhaps the boundary in question is the surface of a substrate, and it is not strictly planar because it has a line on it, but the line's position, shape, and size could be anything. We cannot even be sure, for a general sample, that the paths to any of these surfaces represent *escape* routes for the electrons; maybe they are only routes into an interior void with more material on the other side. If, as programmer, we knew that all samples would be lines of a particular shape or spherical particle defects, etc., on a substrate, simplifying assumptions and rules might be made. However, the range of possible interesting samples is such that sample definition is best left to the user, at run-time. Since, as far as the programmer is concerned, the sample could be absolutely any shape or composition, it is necessary to have a completely general solution to the average-over-escape-routes problem.* Monte Carlo methods are well suited to this; it is difficult to conceive of an alternative with the required generality.

Even with restriction to Monte Carlo modeling, however, the relevant literature is vast. Each of the phenomena mentioned in Sec. 2.2.1 has its own often long history of investigation. An exhaustive review is outside the scope of the present report. A review of a subset of this literature relevant to dimensional metrology has recently been published.[5] For the present report, we confine ourselves to the program (JMONSEL) under development in this project, those prior developments that are particularly relevant to its present form and its immediate forebears.

Modeling of e-beam imaging systems at NBS/NIST appears to have begun in the 1970s [6] with a simulator written by R. Myklebust in what was then NIST's Institute

* Even framing the problem in terms of escape routes represents an over-specialization for some purposes. Escape routes (routes to the detector) are relevant for imaging, but there are other problems for which routes that do not reach the detector are important. For example, paths that stop inside the sample are nevertheless important for charging or for e-beam resist exposure.

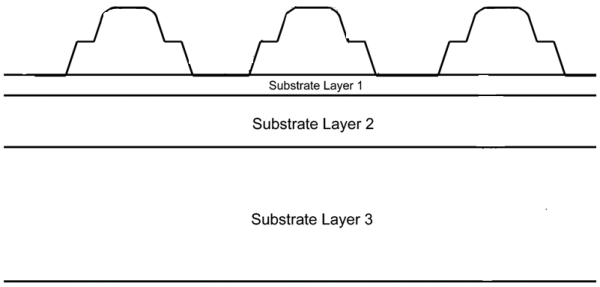

Fig. 8. MONSELCFN samples were 3 lines on a 3-layer substrate. Parameters (linewidth, pitch, wall angle, corner radius, layer thicknesses, etc.) provided variability within this class of shapes.

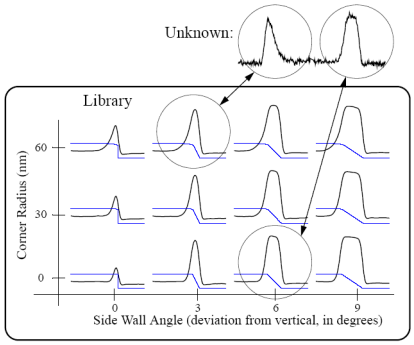

Fig. 9. Schematic use of an edge library. Each library entry is an edge-shape/simulated-image pair. Shape (or other) parameters vary from entry to entry. Measured images are matched to the library. (From Ref. 27.)

for Materials Research.* This program was intended mainly for microanalysis, i.e., the chemical analysis of samples by means of the characteristic x-rays emitted under electron bombardment. There was no need to model SE, since these have energies too low to generate such x-rays. SE images were being used, however, for critical dimension metrology in semiconductor industry applications. This interest was the impetus for a second branch of development, a program named MONSEL.[8,9,10] MONSEL started from the Myklebust program but with the x-ray generation parts shorn and other parts relevant to dimensional metrology with SE added. The added parts included, of course, a model for SE generation, but also new sample shapes that were important for CD metrology, e.g., lines on a layered substrate (Fig. 8).

MONSEL modeled elastic scattering via the Mott cross sections.[11] The alternative screened Rutherford model results in a simple expression for the cross sections that is easy to implement in a simulator, and for this reason it had been popular for earlier simulators. However, the Mott cross sections are more accurate, particularly for low energies (below, say, 1 keV) or heavier elements. MONSEL's implementation used the empirical interpolation formula developed by Browning.[12,13,14] Slowing of electrons was treated in the CSD approximation with Joy and Luo's [3] stopping power to which was added a small residual energy loss that could be adjusted to improve agreement with measured yields.

One of the oldest methods for estimating SE yields has its roots in ideas of Salow [15], subsequently expanded upon by others.[e.g., 16,17,18,19] In this model the inverse mean free path is

$$\lambda_{\mathrm{mfp}}^{-1} = -\frac{1}{E_{\mathrm{SE}}}\frac{dE}{ds}. \qquad \text{(EQ 1)}$$

This is the average number of SE produced per unit path length. Given E_{SE}, which is an effective energy required to produce an SE, one could use one of the stopping power models discussed above to provide $-dE/ds$ and thereby permit determination of the number of SE produced in any given trajectory segment. The probability that these electrons will then be detected after escaping the sample through a planar sur-

* There was even earlier work at NBS/NIST, for example by Berger in 1963 [7], but at that time the work was mainly concerned with basic transport phenomena and computational procedures. It does not appear to have yet been applied to imaging.

face at a distance d is taken to be of the form $0.5\exp(-d/\lambda)$. The factor of 0.5 comes from the assumption that SE are produced isotropically (so half have paths directed toward the surface), and λ is an escape depth. Neither E_{SE} nor λ is known a priori. They are determined for each material by fitting the simulation to the measured yield vs. energy. Lin and Joy provide tabulated values of λ and E_{SE} for many of the elements.[20] In this method the model expression for the detection probability replaces simulation of the SE cascade, which is therefore not performed explicitly. There may be geometrical complications for nonplanar exit surfaces, as mentioned in our discussion on generalizability above.

In MONSEL, inelastic scattering was modeled using the Möller cross section [21] for binary scattering from a free electron, in the non relativistic limit, scaled for the density of free electrons in each material. Valence electrons in the target material were treated as though they were free, while dislodging of more tightly bound electrons was ignored. Generated SE were treated like other electrons; their trajectories, subsequent scattering events, creation of additional SE, etc. were tracked until the energy fell below a cutoff value (usually the vacuum level, below which electrons cannot escape from the sample). MONSEL was, as far as we know, the earliest simulator to do both samples with topography and the full SE cascade. Such a treatment has advantages for samples with non-planar geometry or when details of the cascade are relevant to the problem, as for example for determining charge or energy distributions. The number of fitting parameters was reduced to one (the residual energy loss) from the previous model's two (λ and E_{SE}). There are, however, also shortcomings and difficulties associated with this approach. Explicit treatment of the SE cascade requires a model of the stopping power at energies below 50 eV. The existing expressions were in this energy range little more than uncertain extrapolations from higher energy. Also, the treatment of a dense material by a binary collision model is an approximation valid only for large energy transfers, where many-body effects may safely be ignored. This approximation overestimates the production of low energy SE, an overestimate that was in practice compensated (at least insofar as yields were concerned) through choice of the fitting parameter.

At the time of MONSEL's initial development, the approximations in the inelastic scattering model could only be lifted with difficulty. There did exist a theory for SE generation that included many-body effects. This dielectric function theory was based on the following expression [22] for the many-body differential inverse mean free path:

$$\frac{\mathrm{d}^2\lambda_{\mathrm{in}}^{-1}}{\mathrm{d}(\hbar\omega)dq} = \frac{1}{\pi a_0 E}\mathrm{Im}\left[-\frac{1}{\varepsilon(q,\omega)}\right]\frac{1}{q} \qquad \text{(EQ 2)}$$

with $\hbar\omega$ and q the energy loss and momentum change of the primary electron, a_0 the Bohr radius (≈0.053 nm), E the energy of the incident electron, and $\varepsilon(q,\omega)$ the dielectric function including finite momentum transfer. $\varepsilon(0,\omega)$ is the familiar optical dielectric function. Initially, practical application of the theory was limited because $\varepsilon(q,\omega)$ was not known for many materials. Lindhard [23] had derived an expression for it in a free electron gas. This expression was applicable to a subset of metals (the free-electron ones, mainly alkali metals plus aluminum). Monte Carlo simulators were written based upon Lindhard's expression. These were mainly intended to explore the implications of the theory for stopping power, SE generation, etc., not for topographic imaging, so the simulations were in bulk or planar samples without topography. Although the physics in the model was attractive, its applicability to only a limited set of materials made it unsuitable for a general-purpose simulator, so it was not implemented in MONSEL.

Instead, for some time research was mainly focused on the question of how best to use model results to improve metrology. The essential problem was that simulators provide a means to calculate the expected image for given sample and instrument conditions. In practice, however, the image and some of the beam conditions are known, given by a measurement. It is the sample parameters that are desired but unknown. That is, the practical problem is the inverse of the problem simulators are designed to solve. If the model were some simple invertible function, this would pose no difficulty. However, a complicated algorithmic calculation does not have a simply expressible inverse, and the forward calculation is too time-consuming for a naive iterative inversion to be practical. Accordingly, a model-based library method was implemented.[24,25,26] In this method, a library of simulation results (Fig. 9) is pre-computed for a discrete set of parameter values in the neighborhood of the expected ones. Measured images can then be matched directly to the library, as shown in Fig. 9. Alternatively, interpolation of the library provides a good approximation of the simulator's results within this neighborhood. With fast library interpolation now acting as a surrogate for the slower simulation, the parameters associated with a particular image are assigned by an iterative nonlinear least squares fit.[25] On Si this procedure agreed to better than 2 nm (Fig. 10) with results obtained by cross-sectioning the lines.[27] Agreement on resist lines was also good, apart from a small systematic difference consistent with resist shrinkage during measurement.[28]

Because it is common to encounter sub-nanometer stated uncertainties for SEM CD measurements, we should remark that such sub-nanometer values, when they are valid, are invariably *repeatabilities* (e.g., same-site measurement precision). They are for this reason only a part (usually not the largest part) of the total uncertainty in a CD measurement. Agreements mentioned in the last paragraph should be compared to the 10 nm or more offsets often measured (e.g., Ref. 29 and 30) when comparing CD-SEM measurements to other techniques. MONSEL simulations had predicted such offsets [31], including variation in the size of the offset with sidewall angle—a variation with which independent measurement [29] was in good agreement.

As a result of the above developments, both an SE image simulator and a method to apply its results to the improvement of CD measurements were available. They had been demonstrated to improve measurements. However, some shortcomings still remained. Initially the most evident of these were restrictions by MONSEL to lines on a substrate. Such samples were sufficient for its initial intended CD metrology application, but there has been increased interest in modeling other sample configurations, a few of which are shown in Fig. 11, that are not available given MONSEL's restrictions. Such configurations include contacts and vias, particles, samples with regions containing different materials within a single layer (like filled Damascene trenches), line edge roughness, sidewall roughness, FinFETs (a type of three-dimensional field effect transistor), and complex patterns for contour metrology. Extension of MONSEL to fully three-dimensional samples with general shapes was the initial impetus for a new version of the code, called JMONSEL.

While MONSEL was being developed in NIST's Semiconductor Electronics and Precision Engineering Divisions, back in NIST's Surface and Microanalysis division, the original Myklebust program had been succeeded by NISTMonte [32], a Java program that added the capability of modeling more complex three-dimensional samples. With small modifications, the geometrical implementation was compatible with requirements for SE models. To take advantage of this existing capability, JMONSEL was written as a Java package that interfaced with NISTMonte.[33] The JMONSEL package provides the necessary models and modifications required for SE generation and interface scattering.

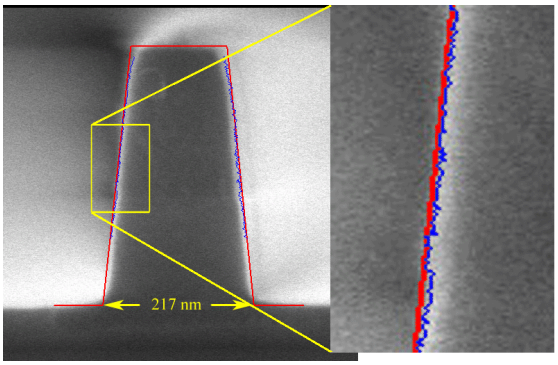

Fig. 10. Comparison of edge assignment from the cross section image (blue) with the assignment from a top-down image using the model-based library technique (red).

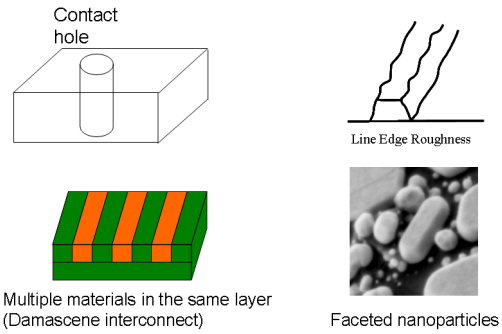

Fig. 11. Examples of relevant samples that do not fit the uniform lines-on-a-substrate model.

2.2.3 JMONSEL

JMONSEL is the starting point for the current project. In this section we describe its initial capabilities, particularly as they relate to the goals of the project.

The NISTMonte/JMONSEL combination is a library of Monte Carlo simulation-related routines written in Java. JMONSEL was written in Java to take advantage of the existing NISTMonte code. The original choice of Java for NISTMonte was made for cross-platform compatibility. The code has a modular construction (Fig. 12) in which various components (electron guns, materials, sample shapes, detectors, and physics models) are implemented as separate modules that connect via defined interfaces with the Monte Carlo executive that actually runs a simulation. Samples are divided into regions, each of which is associated with a particular MaterialScatterModel. The MaterialScatterModel defines the material and associated scattering mechanisms that characterize the region. The boundary of each region is defined by a Shape. Shapes are defined using constructive solid geometry (CSG), in which any number of three-dimensional shape primitives (Fig. 13) may be freely combined using standard operations (translation, rotation, union, intersection, difference, complement,…). Regions may contain subregions (e.g., a void, buried defect, or intentional buried structure) nested to any depth. Together, these capabilities can be used to define and simulate images from complicated samples (e.g., the FinFET shown in Fig. 14 and its simulated image in Fig. 15).

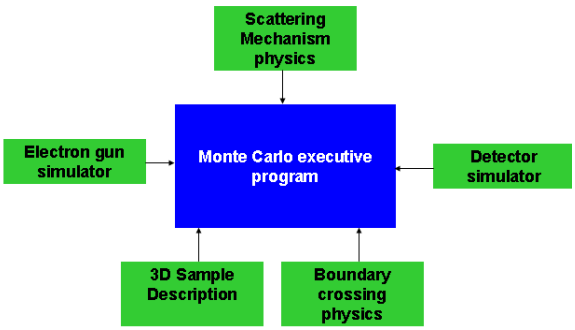

Fig. 12. Modular construction of NISTMonte/JMONSEL. Electron guns, detectors, and scattering physics are written as separate interchangeable modules.

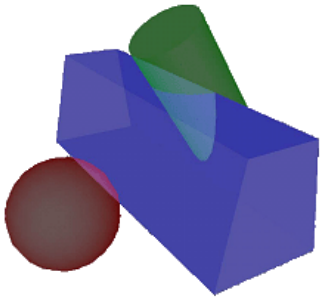

Fig. 13. Examples of constructive solid geometry primitives. Complex samples are built up from spherical, cylindrical, and polyhedral building blocks.

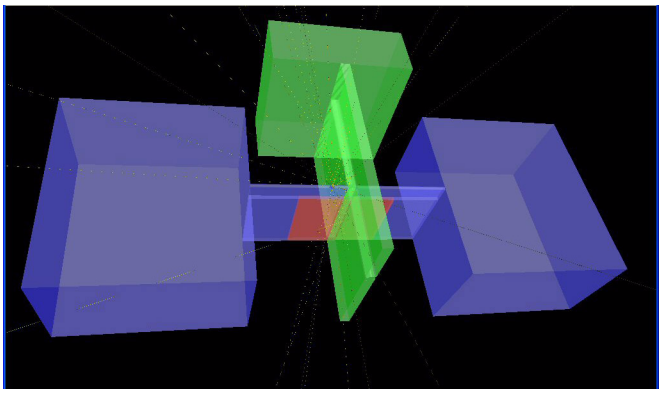

Fig. 14. VRML rendering of FinFET model in JMONSEL (Reproduced from Ref. 33).

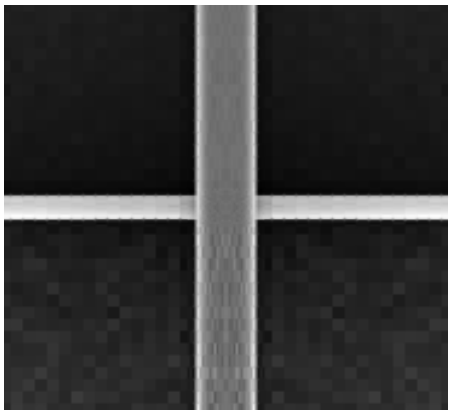

Fig. 15. JMONSEL simulated image from the center of the FinFET sample, where the fins cross. (Reproduced from Ref. 33).

The code is accessed via a Jython scripting utility, i.e., from a window into which Jython commands may be typed and executed interactively, or from which Jython scripts may be run. The user in effect writes a short program to define the simulation. The scripting language has conditional statements, loops, and other standard programming language constructs. This gives the interface considerable flexibility. A linescan or series of them is implemented as a repeat loop in which the landing position of the electrons is changed. This allows SE yield vs. position (a standard SE linescan or image) to be simulated. Depending upon how one configures the detectors, one might also record BSE yields, energy spectra, or other properties vs. position. One could as easily vary angle of incidence, beam energy, beam size, etc., to simulate dependence upon any of these quantities. Loops can be nested to, e.g., simulate an image (yield vs. position inner loop) as a function of varying sample shape or beam energy (outer loop). This has obvious application to building a model-based library for doing subsequent CD measurements or for deciding the best landing energy to use for defect detection. Conditional statements allow a series of simulations in which the outcome of early simulations determines which simulations will be run next. Such modes might be used to minimize user-defined cost functions (e.g., defect detection failures or CD measurement uncertainty) for solving optimization problems. Parameters to minimize such a function are more efficiently found via gradient search, Levenberg-Marquardt, or any of a number of other standard algorithms in place of exhaustive mapping.

The initial implementations of the modules shown in Fig. 12 were meant to duplicate MONSEL in all respects apart from the sample description. Elastic scattering and SE generation were as described above. The electron gun produced electrons moving down the column with starting positions normally distributed with a specified standard deviation about a specified center. The detector by default recorded energy and angle histograms of electrons that escaped the sample. Counts in various portions of the histogram could be summed as desired. For example, to obtain the SE yield, counts with $E < 50$ eV (the usual operational definition of SE) were summed and divided by the number of incident electrons. Similarly, for the BSE yield the numerator in this ratio was replaced by total counts with $E > 50$ eV. The user can optionally save VRML (Virtual Reality Modeling Language) files (like Fig. 14) that can be viewed with widely available viewers. The default trajectory image file generator produces a projection of electron trajectories (Fig. 16) into the x-z plane.

Upon completion of the initial implementation, JMONSEL and MONSEL were compared as a check against coding errors.[33] Results were indistinguishable apart from small variations (expected due to the stochastic nature of Monte Carlo simulations) for samples within the repertoire of both simulators.

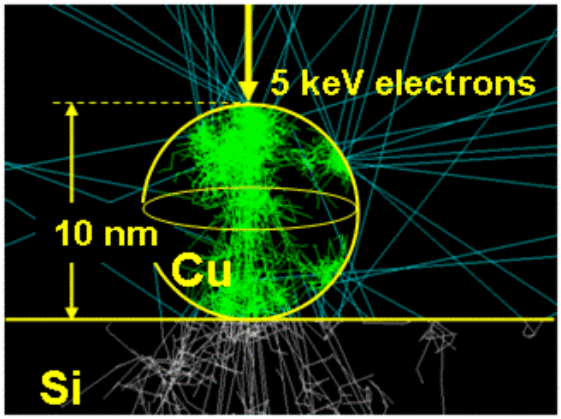

Fig. 16. JMONSEL trajectory plot for 5 keV electrons incident on a gold sphere on a silicon substrate. Yellow annotations were superimposed afterwards.

3.0 Project overview

As discussed in Sec. 2.1, the project goal is to develop and deliver e-beam imaging simulation software useful for image interpretation, measurement optimization, and assessment of measurement tool limits in the context of dimensional metrology and defect analysis. JMONSEL as it existed at project inception fell short of this goal in several respects. We discuss these first to motivate the project tasks that follow.

The version of JMONSEL described in Sec. 2.2.3 had removed the *geometrical* but not the *modeling* limitations, of the preceding MONSEL simulator. The modeling limitations included these: (1) There was no facility to model charge build-up in insulating parts of the sample, the electric fields derived from such charges, or the effect of those fields on incident or emitted electrons. (2) Although the approximately 2 nm offsets achieved by MBL (see Fig. 10 and the accompanying discussion in Sec. 2.2) with the MONSEL model were a considerable improvement over offsets in the absence of modeling, they nevertheless remained higher than the industry's sub-nanometer [34] goals. For the source of the remaining error, suspicion was naturally directed at the approximations (e.g., neglect of screening) employed in the SE generation and transport model. Without alternative models for comparison, it was difficult to estimate how significantly these approximations might affect metrology. The approximations could reasonably be expected to affect electron inelastic mean free paths, which determine the electrons' range of motion before they come to rest or escape. The final position of non-escaping electrons was not important for MONSEL's modeling of conductors (since charges can readily flow as required to equilibrate), but the new goal of modeling charging in insulators renders such internal charge distributions more important. (3) JMONSEL initially, like MONSEL before it, had only a single model for each of the main phenomena: elastic scattering, SE generation, slowing down, and interface scattering. A single model would be enough if we were sure it were correct. However, the physics is not certain, particularly at low energies where the phenomena are complex and little explored. In such circumstances, it is desirable to be able to examine alternatives; in this way one obtains not only a result, but also a measure of the sensitivity of that result to the assumptions one is forced to make. The current project aims to address these limitations. As we describe next, we are in some respects in an advantageous position to do so.

Regarding the previous paragraph's item #2 (approximations in the SE generation model), as mentioned in Sec. 2.2.2 there does exist an SE generation model that includes a more comprehensive treatment of the physics. The obstacle at MONSEL's inception was the limited number of materials to which this treatment was applicable. The problem was that the Lindhard theory only permitted the calculation of $\varepsilon(q,\omega)$ in Eq. 2 for free-electron metals. However, in December 2000, NIST published a stan-

dard reference database of inelastic mean free paths (SRD 71).[35] A large number of elements and a few compounds were included. The methods employed to determine the mean free paths were an improvement on the generality of earlier methods. In 1974 at NIST, Powell had introduced the idea of using optical data in mean free path calculations.[36] Penn (also at NIST) later carried out and extended this idea, applying the theory to calculate mean free paths in Cu and Ag, neither of which is particularly free-electron-like.[37] The optical data provide a measurement of $\varepsilon(0,\omega)$. This is extended to $q \neq 0$ via a suitable dispersion relation. The connection to the scattering theory that had previously been developed in terms of the Lindhard dielectric function was made by writing this extended optical dielectric function as an integral expansion of Lindhard functions with an appropriate spectral weighting. In a series of approximately 20 papers between 1987 and 2008, Tanuma (Nippon Mining), Powell, and Penn (henceforth TPP) applied these methods to determine inelastic mean free paths and stopping powers for a large number of materials, including mean free paths for 41 elemental solids, 15 inorganic compounds, and 14 organic compounds. These, together with results from other workers and measurement results by other techniques eventually became the foundation for SRD 71. The optical input data required for determining mean free paths and stopping powers are the same input data required for implementation of a Monte Carlo simulator based on Eq. 2, as demonstrated by Ding and Shimizu.[38] These developments have ameliorated to a significant extent the restrictions on applying a more accurate model.

Regarding item #3 (restriction to a single model), JMONSEL's design lends itself to implementation of alternatives. Unlike MONSEL, which was an implementation of a particular model, NISTMonte/JMONSEL was designed as a "plug and play" test bed for multiple models. To implement a new physics model, one need only write those few routines dictated by the appropriate (e.g., ScatterMechanism) interface. These new models are then simply added to the library of routines that comprise NISTMonte and JMONSEL. Users may select from this library whichever routines are most appropriate for their particular application.

The above goals, obstacles, and opportunities prompted the current project. The fundamentals of the project are as follows:

- Improve the scattering physics models. This task includes a number of sub-tasks. Most significantly, implementation of a dielectric function theory (DFT) model for SE generation requires (1) collecting the energy loss function ($\mathrm{Im}[-1/\varepsilon(\omega)]$) and certain other material properties data required for simulation and (2) using those inputs to compute scattering tables for materials of particular interest in semiconductor electronics applications. At project inception we knew how to calculate the tables for metals, but the method had to be generalized for insulators. The DFT does not completely specify the final energy and momentum of SE,* so additional (and uncertain) assumptions are required. These uncertainties motivate the next task.
- Study sensitivity of model results to choice of model. In practice this means implementing more than one model, e.g., for different a priori reasonable assumptions, performing simulations with the alternatives, and evaluating the different outcomes. If measurement data exist to which the outcomes can be compared, results of this exercise can determine which of the assumptions are the most reasonable.

* DFT is a theory for energy and momentum *transfer*. These amounts must be added to the pre-collision energy and momentum of the target electron. These are not specified by the DFT. Matters can be further complicated by the effects of binding on the target electron.

TABLE 1. Project milestone schedule from scope-of-work document.

Description	Due Date
Project start	6/15/2009
Supplier will improve the SE generation model	7/31/2009
Supplier will deliver a presentation describing results to date	11/16/2009
Supplier will conduct a sensitivity study on e-beam tool parameters	12/31/2009
Supplier will lay groundwork for a charging model	1/28/2010
Supplier will develop Java classes to track charges	2/28/2010
Supplier will deliver a presentation describing results to date	3/31/2010
Supplier will prepare the interim report	4/30/2010
Supplier will develop a solver	6/30/2010
Supplier will modify transport algorithms	8/31/2010
Supplier will deliver a presentation describing results-to-date	10/31/2010
Supplier will debug the code	10/31/2010
Supplier will simulate cases of interest	11/30/2010
Supplier will deliver a presentation describing results to date	12/31/2010
Supplier will provide the Final Report for publication and distribution to Sematech and FOUNDATION	12/31/2010
Supplier will provide an executable file and source code for e-beam simulation	12/31/2010

If measurement data do not exist, the exercise helps us decide whether the assumptions are important (make a significant difference in results) or not.

- Add charge modeling capability to the simulator based on the best available method. This model depends critically upon completing the steps mentioned above because the charge distribution within the sample is determined by the SE generation and transport models. The charging distribution is of course an essential input for computing the electric fields. Starting with the charge distribution together with other defined boundary conditions such as the presence of arbitrarily shaped dielectric materials and conductors held at ground or other potentials, Poisson's equation must be solved to determine electric fields. Finally, the effect of these fields on the trajectories of primary and secondary electrons must be taken into account.
- Final model is being made available to ISMI Member Companies beginning December 31, 2010.

The project milestone schedule is given in Table 1.

4.0 Accomplishments and results

4.1 Improvements to scattering physics

4.1.1 SE generation model development

At project inception we had already established procedures for a DFT-based SE generation model of metals. Copper was used as a test case during development because of

the availability of reliable measurement data for that element. Consequently, we had scattering tables for Cu. The procedures derived for Cu could be used for other conductors. Our present procedure treats the Lindhard dielectric function in the single-pole approximation. Inaccuracy introduced by this approximation is considered small for Cu and other non-free-electron metals, but is important for aluminum, silicon, and the alkali metals. This DFT-based model provides enough detail in its description of SE production that we can (and do) dispense with the CSD approximation when we use it. That is, instead of using a separate formula for average energy loss with distance, each time we create an SE we subtract its initial energy from the primary electron that produced it. These random and discrete losses are more realistic than the CSD approximation.

The material data required as input for a scattering table calculation include many relatively easy-to-obtain material properties (density, elemental constituents, their stoichiometry and core-level binding energies, …), others that are sometimes somewhat more difficult (Fermi energy, band gap, work function, possibly phonon energies), and the measured energy loss function (ELF). The last of these is the least widely available. An example ELF is shown in Fig. 17. As is evident, the ELF is a curve, not a single number. It can be expressed as a function of frequency, ω, or energy, $\hbar\omega$. In electron scattering calculations, these values correspond to energy losses by the primary electron. In an inelastic collision the primary electron with energy E can lose energy approximately in the range from E_g to $E - E_F - E_g$, where E_F is the Fermi energy and E_g is the bandgap. Hence it is necessary either to measure or otherwise estimate the ELF for all energies in this range. Typical values of E in a CD-SEM are close to 1 keV. Higher energies may be used in analytical SEM or e-beam defect detection tools. The high end of the energy range is thus in the UV or soft x-ray part of the spectrum. Much of the available data derive from synchrotron optical measurements, collected for example by Palik.[39] Data derived from electron scattering measurements would be potentially even more useful, because they could in principle provide $\varepsilon(q, \omega)$ for $q \neq 0$. However, inelastic mean free paths for electrons are typically much smaller than even the thinnest achievable films. Consequently, interpretation of electron scattering data is complicated by multiple scattering and the mixing of bulk with strictly surface phenomena. There have been recent proof-of-concept demonstrations [40,41] but data derived from this source are not yet available for many materials.

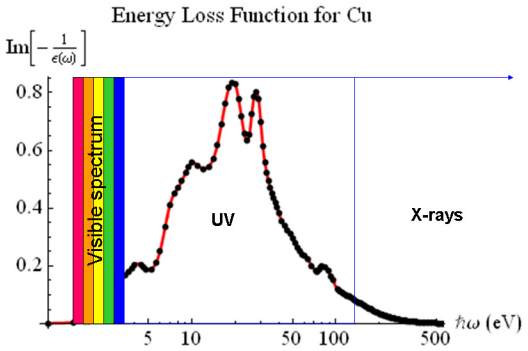

Fig. 17. Energy loss function for copper.

From Cedric Powell (and indirectly from Shigeo Tanuma) we collected ELF data for 40 elements (including 3 forms of C) and 4 compounds. These data had been used by the TPP collaboration for their contributions to SRD 71. TPP had vetted the data by checking agreement with known sum rules. We also contacted Eric Bosch at FEI, who had collected ELF data for 34 materials. The majority of these were for materials already in the TPP collection, but by pooling our resources we obtained ELF curves for 13 new compounds. The pooled collection contains ELF data for these elements: Ag, Al, Au, Be, Bi, Co, Cr, Cs, Cu, C (diamond), C (glassy), C (graphite), Dy, Fe, Gd, Ge, Hf, In, Ir, K, Li, Mg, Mo, Na, Nb, Ni, Os, Pd, Pt, Re, Rh, Ru, Sc, Si, Sn, Ta, Tb, Ti, V, W, Y, Zr and these compounds: Al2O3, AlAs, AlN, GaAs, GaN, H20, InAs, MgO, NiSi, PMMA, polyimide, Si3N4, SiC, SiO2, TiN, TaN, TiO2.

As received, these curves are not yet quite in the form required for simulations. Via Eq. 2 they provide the double differential free path (or cross sections). Together with the other data mentioned above (work functions, etc.) these specify probability distribution functions that, with a random number generator can be used to select the outcome of a given scattering event. However, those outcomes are related to the cross sections via integrals, and numerical integration is prohibitively slow to be repeated millions of times in the course of a simulation. Consequently, the data are used to precompute tables that can be rapidly interpolated by a simulator. Four tables are required for each material: (1) inelastic mean free path vs. energy, (2) energy loss vs. energy and a random number, (3) scattering angle vs. energy, energy loss, and a random number, and (4) pre-collision energy of the target electron as a function of energy loss and a random number. In the project so far, we have completed such tables for these materials: Ag, Al, Au, Cu, Fe, graphite, glassy C, Mo, Pt, Si, SiO_2, Ru, Ta, Ti, Sn, W. Tables for poly(methyl methacrylate), abbreviated PMMA, are in process. Our original procedures for computing the tables were developed for Cu. They had to be generalized for two of the materials in this group, the insulators SiO_2 and PMMA.

As mentioned in the overview (Sec. 3.0), one of the uncertainties in the SE generation model is associated with assigning initial energy and momentum to the target electron. We have implemented 3 different ways of doing this, one for each of the models suggested in references 38, 42, and 43.

What do we do if we need to simulate a material for which ELF data are not available? In the long run, Sematech or other interested parties might consider commissioning additional ELF measurements for important materials. In the shorter run, we have two back-up models, presumably less accurate than the DFT-based one, but with less stringent requirements. One of these back-ups is the original MONSEL model. The other is a newly implemented variant of the model represented by Eq. 1. Previous implementations, as described in Sec. 2.2.2, computed an escape probability for the SE at the moment they were generated. Our implementation differs in that we actually instantiate the SE (with initial energy equal to E_{SE}) and track them through subsequent scattering events. This procedure is somewhat more time-consuming, but it is more general, inasmuch as this method can deal with non-planar shapes. There are still two fitting parameters in our variant—a stopping power parameter replaces the escape depth parameter. These parameters are chosen by fitting the simulation results to measurements (e.g., yield vs. E) as is typically done.

4.1.2 Elastic scattering models

As mentioned above, JMONSEL originally used the same Browning empirical approximation of the Mott elastic cross sections as had previously been used by MONSEL. This approximation was originally chosen over 10 years ago (for MONSEL) because it is a simple, quickly calculable formula that captures the main trends of energy and atomic number (Z) dependence. It agrees with the observed energy dependence and yield vs. Z goes smoothly through the data, with measurements roughly equally likely to be higher as lower. However, the approximation misses variations in cross section with small changes in Z as the periodic table is traversed. These variations are real, and *not* solely measurement error. Accordingly, we have implemented two methods that obtain cross sections by interpolating tables of Mott cross sections. One of these uses the NIST SRD 64, which contains cross sections determined from the Dirac-Hartree-Fock potential.[44,45] The other uses the tables of Czyzewski et al.[46] A fourth option uses a screened Rutherford cross section. Tabulated methods can only be used for energies in the range of existing tables, 50 eV to 20 keV for the NIST tables and 20 eV to 30 keV for the Czyzewski tables. Our implementations switch to a scaled

Browning for energies lower or screened Rutherford cross section for those higher than those in the tables.

4.1.3 Interface scattering models

JMONSEL has improved upon the initial interface scattering model that it inherited from MONSEL. The new model describes the potential energy step with the simple exponential form, $U(x) = \Delta U/[1 + \exp(-2x/w)]$. This expression contains parameters descriptive of the height, ΔU, of the step and its width, w. The probability that an electron incident on such barrier will be transmitted can be solved exactly.[47] This solution has been implemented in JMONSEL.[48]* An electron that does not transmit reflects specularly from the interface. One that does transmit has its momentum component normal to the barrier altered to change its kinetic energy by an amount equal to the step height. This produces an angle change closely analogous to refraction of light at an interface.

4.1.4 Detectors

JMONSEL's initial detector module operated by triggering when electron trajectories intersect the walls of the vacuum chamber, i.e., when electrons were well clear of the sample. This kind of detection permits electrons that escape from low-lying areas of a sample with topography to re-enter the sample if their straight-line escape path is obstructed by neighboring sample features. This is one cause of the proximity effect commonly observed in SEMs. (Feature appearance depends upon other features in the neighborhood.) This is a realistic mode for some kinds of SEMs, but industry CD-SEMs frequently employ large electric fields in the neighborhood of the sample. These fields are intended to fully extract all secondary electrons directly to the detector as soon as they leave the sample. A new "RegionDetector" module for JMONSEL was written. This is a detector that triggers whenever electrons enter any of a user-specified list of regions. By setting the trigger region(s) to the space surrounding the sample, the full-extraction mode can be simulated. The RegionDetector is, however, more general than this. It can be used in other modes, for example, to define a detector with any desired acceptance angle (e.g., the commonly used split annular backscatter detector). One can even place a region detector inside the sample, where it can provide a nondestructive monitor of a sort not experimentally realizable but that nevertheless returns useful statistics about a simulation. When triggered, the detector tabulates electron energy and angular distributions that can later be post-processed to account for the characteristics of a real detector (e.g., detection efficiencies or acceptance angles).

4.1.5 Models for insulator phenomena

Scattering tables for insulators, as we mentioned above, required some generalization of the method of calculation. There is, however, a more serious issue with insulators, illustrated by Fig. 18. In conductors, slowing down of fast electrons is mainly due to electronic excitation. In insulators, however, excitations below a minimum energy

* This is the only such implementation of which we are aware. Most Monte Carlo models intended for high-energy electrons can afford to ignore the surface barrier. Some SE models do not explicitly track trajectories of SE, so whatever happens at the barrier is presumably incorporated, along with losses in transit to the nearest surface, into the overall escape probability. Models that do have an explicit model for the barrier have previously used either the limit of an abrupt barrier, $w = 0$, or a classical barrier, with w much greater than the electron wavelength.

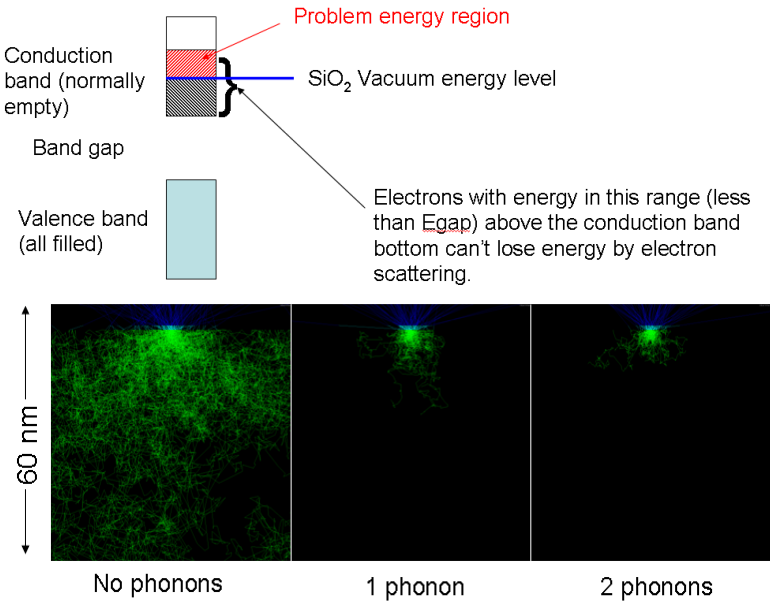

Fig. 18. An issue with modeling insulators and its solution. Upper: Schematic energy diagram for a wide-bandgap insulator with occupied valence band and mostly unoccupied conduction band. Primary electrons within the bracketed (shaded) region have too little energy to promote a valence electron. Those with energies greater than the vacuum level (shown, blue line) can escape the sample. Overlap of these two ranges is shown in red. Lower: trajectory plots in SiO_2 with 0, 1, and 2 phonon modes.

(the band gap) are forbidden. Primary electrons with energy less than the band gap, that is, those in the energy range marked by the curly bracket in Fig. 18, cannot lose energy via electronic excitations. The modeled range of these electrons becomes infinite, which is unphysical. (See the trajectory plot at lower left.) If the bandgap were narrow, as in a semiconductor, this would not pose a problem for simulating images because these poorly modeled electrons would not contribute to images—even the most energetic of them would have too little energy to surmount the surface potential barrier and escape. However, in wider bandgap materials, like SiO_2 shown in the figure, part of this energy range exceeds the vacuum energy. Therefore, it becomes necessary to consider other energy loss mechanisms, for example phonons. We therefore made a model for electron-optical phonon scattering. It is a Monte Carlo implementation as described by Llacer and Garwin.[49] Reduction of the predicted electron ranges when this model is included are shown along the bottom of Fig. 18. Required inputs are the sample temperature, the dielectric constant at high and zero frequency, the phonon mode's energy, and a rate multiplier that is ordinarily an integer equal to the number of modes at the stated energy. In this model the stopping power contributed by phonons at different energies is roughly proportional to E_{ph}^{2}, apart from a slowly varying logarithmic factor. Although the code allows inclusion of as many phonon modes as desired, this quadratic weighting with energy means it is usually possible to simplify the simulation by including only the highest energy phonon modes.

Insulators may also trap charges. We have included in JMONSEL a polaron trapping mechanism according to the model of Ganachaud and Mokrani.[50] This model assumes a trapping probability of $S_{\mathrm{trap}}\exp(-\gamma_{\mathrm{trap}}E)$, with E the electron's energy and the other symbols representing fitting parameters. This is a rather crude model, and it is less than desirable to have two free parameters, but this appears to be the best available model at the moment.

As a test of JMONSEL's new procedures for scattering tables for insulators and phonons, we chose data that had been published by Kadowaki et al.[51] The main measurements consist of SEM intensity profiles across sloped Si edges in a trench. The edges were measured with and without an SiO_2 layer. Ancillary measurements established the shape of the edge and SiO_2 layer thickness. These careful measurements were undertaken in an effort to explain the failure of another model (without phonons) to reproduce the observed brightening of the edge with the addition of the SiO_2 layer.

This seems a reasonable system on which to test JMONSEL's newly implemented insulator models. Testing on a bulk insulator would be problematic, because charging would be expected to skew the results. Kadowaki et al. argue that in their thin films charging is, however, very slight because the incident beam induces conductivity (EBIC, electron beam induced conductivity) and it goes all the way through the film, thereby providing a path to ground through the Si substrate. Furthermore, if it were necessary to include the effect of trapping, interpretation would become more doubtful: would agreement mean the physics was right, or only that fitting parameters can hide errors? As it is, choices of trapping parameters by Ganachaud and Mokrani give mean trapping distances large compared to the film thickness, suggesting that trapping can to a good approximation be ignored in these films. Absence of the trapping parameters makes a more stringent test of the other components of the model.

Because brightness and contrast are adjustable in SEM images, the image intensities do not provide a direct measure of absolute yields. Accordingly, to compare profiles with different SiO_2 layer thicknesses, Kadowaki et al. applied a linear adjustment (scale and offset) to their measured intensities to make them agree at mid-line and mid-trench. We applied a similar scaling to JMONSEL simulation results. Once this is done, differences in peak height and shape can be compared. JMONSEL simulations were performed with and without oxide. The with-oxide simulation used SiO_2 film thickness fixed at 0.9 nm, the value measured by Kadowaki et al. after 16 days growth. The raw model results were convolved with a Gaussian shape to account for the finite beam size. Since this size was not known a priori, it was varied to find the best fit. At this point, the width, shape, and height of the simulated and measured peaks agreed well, as shown in Fig. 19.

4.2 Addition of a Charging Model

4.2.1 Overview of the approach

A schematic of our approach to modeling charging in the SEM is shown in Fig. 20 The plan is based on using finite element analysis (FEA) to determine the electrostatic potential as a function of position in and near the sample. FEA has the advantage of being a very general technique. It is applicable to arbitrarily-shaped samples, and it can accommodate a wide variety of boundary conditions.

1. The sample and space through which electrons travel are meshed, i.e., subdivided and tiled by smaller volumes as shown in Fig. 20, step 1. The program maintains a charge counter associated with each mesh element.
2. In step 2, we perform Monte Carlo simulations of trajectories for N primary electrons, with N an adjustable parameter. During these simulations we must keep track of charge accumulation or depletion. Thus, whenever an electron's trajectory terminates in an insulating mesh element of the sample (because, e.g., its energy falls below a critical value or because a trapping model determines it is trapped) one electron charge is added to the counter for that mesh element. When a SE is

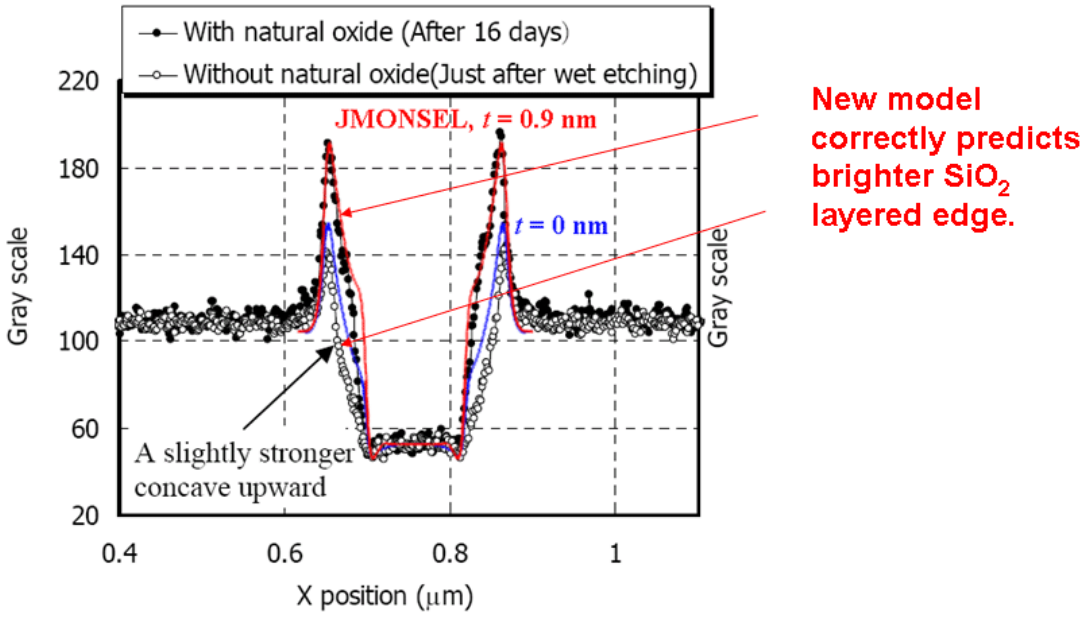

Fig. 19. SEM profiles across a trench feature comparing experimental and modeling results. Parts in black are from Ref. 51, used with permission. Superimposed parts in color are present work.

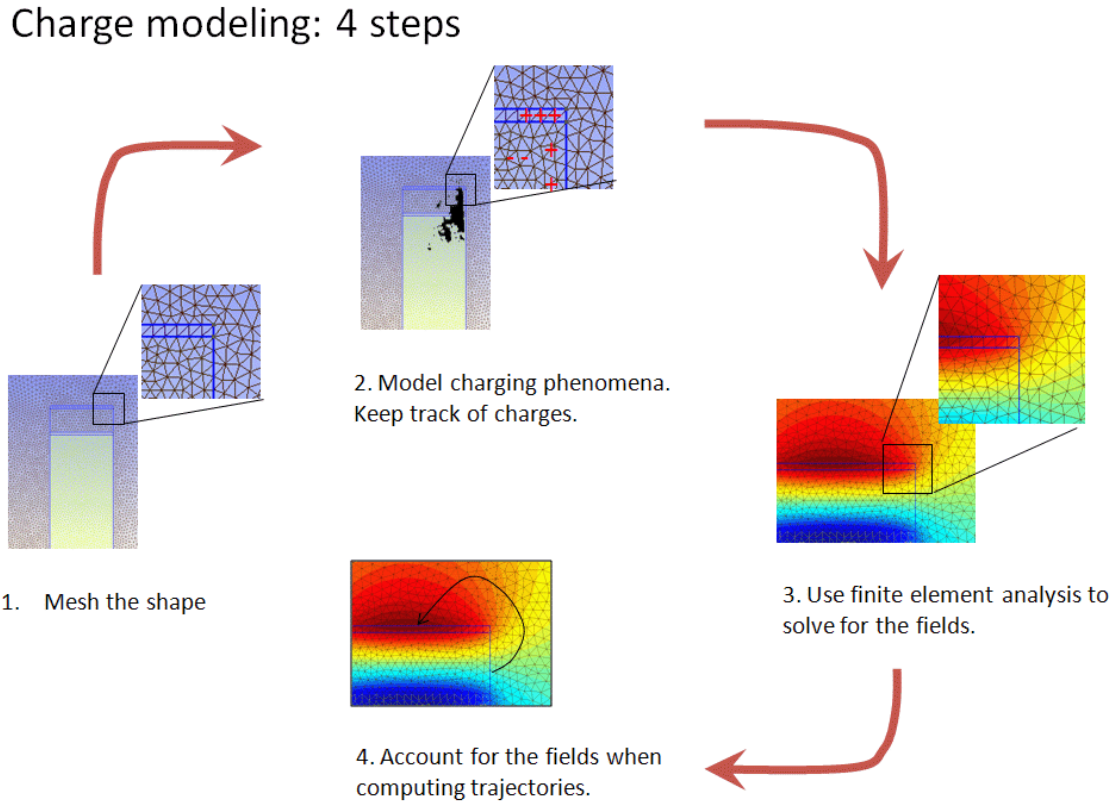

Fig. 20. Schematic of the 4-step plan for modeling charging in the SEM.

generated, one electron charge is subtracted from the counter for the mesh element in which it was generated.

3. Upon completion of the N^{th} primary electron, Monte Carlo simulation is paused while a finite element analysis is performed. This analysis determines the electrostatic potentials at the nodes of all mesh elements. The potentials within the elements can then be determined by interpolation, and the electric fields can be determined from $E = -\nabla\varphi$.
4. The Monte Carlo simulation is then resumed for the next N electrons as in step 2, but now with the electron trajectories modified by the accelerations implied by the most recently computed electric fields. Execution alternates in this way between Monte Carlo simulation of trajectories and finite element solutions of accumulated charge until all the required electron trajectories have been run.

The choice of a value for N is a compromise between speed (larger N means less frequent FEA pauses, so is faster) and accuracy (smaller N is more accurate).

4.2.2 Implementation

Meshing of three-dimensional spaces and performance of FEA on such spaces with appropriate boundary conditions are themselves non-trivial problems with a long history of development. To reproduce a significant part of that development within the time allotted for this project would have been difficult. Accordingly, we sought to perform these tasks with independently developed specialized software. There are, of course, commercial packages to perform these tasks. However, since JMONSEL is public domain, we considered it undesirable to require users to purchase licenses for third party software if this could be avoided. As part of this project, a number of open-source candidates were identified. The initial implementation uses Gmsh [52,53] to produce the mesh and GetDP [54,55] to perform the FEA. Both of these packages are open source and freely distributable under license terms described on their respective web sites.[53,55] (Generally speaking, these are GNU General Public Licenses but with some exceptions allowing additional freedom to combine the software with certain other software.) On computers running Windows, both of these programs may be run via command line from a DOS prompt. In this case the programs obtain their necessary input data from files previously prepared by the user, and results of their calculations are likewise stored in data files.

The operation of these programs and the accuracy of results were first tested by application to problems simple enough to permit exact solution of Maxwell's equations. In such cases, the FEA solution can be compared to the exact one. One such problem is illustrated in Fig. 21. In the schematic on the left, the two 100 nm tall rectangles represent PMMA lines standing on a Si substrate (the plane at $y = 0$). When an electron beam scans such a sample, SE escaping from the top of the line should leave a net positive charge behind while the more energetic primary electrons should come to a stop some distance below that. Motivated by this, we have imposed a uniform layer of positive charge at the top and a corresponding negative layer 10 nm farther down. In this way the scenario to be solved is motivated by a practical measurement. However, some simplifying assumptions were made to render its solution exactly solvable. These were: (1) the dielectric constant of the lines was set to 1, the same as the surrounding vacuum, and (2) the charges were equal and opposite in well-separated layers of uniform charge density equal to 1 electron per cubic nanometer. With these simplifications, the problem can be exactly solved to determine the potential as a function of position. The method of image charges is used to account for the Si ground plane. A contour plot with labelled contours (the potential in volts) superim-

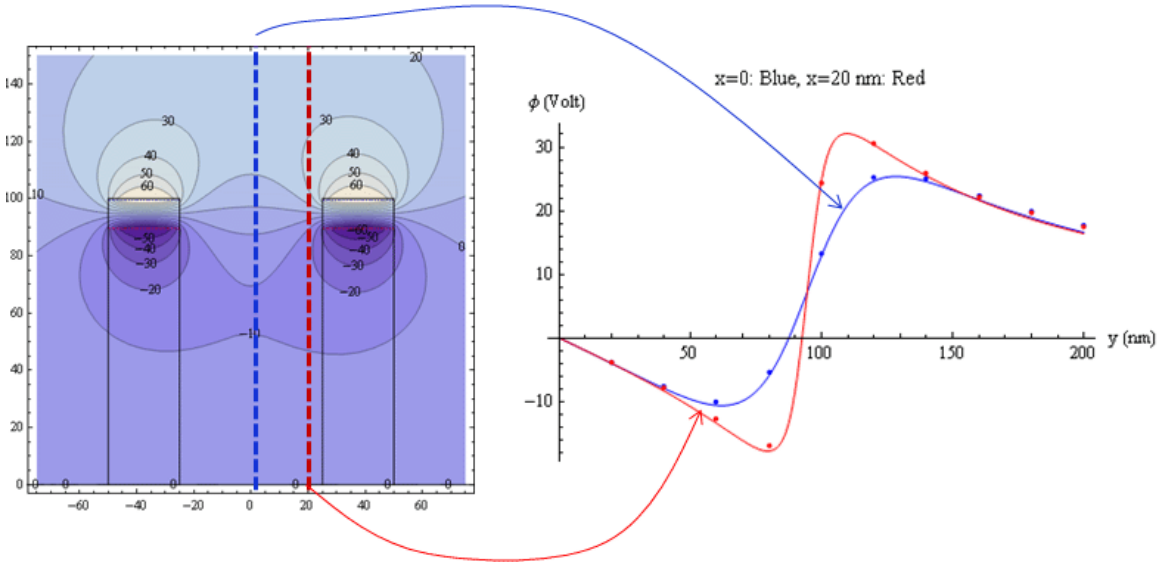

Fig. 21. A test of the FEA solution. The geometry on the left, with a dipole charge separation approximately 100 nm above a grounded plane, was solved exactly using the method of images. The resulting potentials along the blue and red vertical lines are displayed by the blue and red lines in the plot on the right, where the horizontal axis represents height above the plane. The FEA solution is displayed as discrete points in the plot on the right.

poses this exact solution on the schematic. The potentials vs. height above the Si plane along the blue and red dashed lines are shown in the graph on the right by the blue and red continuous curves respectively. The values from the FEA solution are shown as blue and red dots. Errors are deemed to be acceptably small.

The problem in Fig. 21 is translationally invariant in the direction perpendicular to the page. Because of this it was possible to solve its FEA problem in seconds using a two-dimensional mesh. As soon as we require solution of more realistic charge distributions and more general sample shapes, the translational invariance is lost and a 3-D solution is essential. This places a much greater demand upon computer resources, as can be appreciated by the following simple argument. The range of 500 eV electrons in most solids is on the order of 10 nm to 30 nm. Thus, we expect charge separations of about that size. Consequently, the potential will vary significantly over distances of 10 nm or so. For accuracy, it is desirable to keep mesh elements small compared to this. A target size of 1 nm seems reasonable. At the same time, we need the size of the meshed space to be relatively large. Charges are deposited over a surface area equal to that scanned by the SEM. This can easily be several hundred nanometers square. It is desirable to mesh volumes out to some multiple of this distance, at least to a distance such that further changes in the potential become negligible. In some cases, if external fields are imposed by grids and we wish to include those in the simulation, it might be desirable to mesh volumes of many cubic centimeters. If we were to naively mesh a 1 cubic *micro*meter volume with 1 nm^3 cubes, we would need 10^9 of them. Similarly meshing 1 cm^3 would require 10^{21} mesh elements. The latter is clearly out of the question, and even the former would likely overwhelm most desktop computers. (It would require 12 GB just to store single-precision coordinates for 10^9 nodes.)

The solution to this problem is to use a variable mesh size. With variable size the mesh can be made small in a region of interest (near where we are depositing charges and where we will be using image intensities to draw quantitative conclusions) and then progressively larger with distance from that region. Gmsh meshes the space with tetrahedra. It has several different options for specifying how the size is to

be varied. In the one we have found most useful the user specifies the target mesh size as a function of position. Fig. 22 illustrates. In this case the neighborhood of the line on the right has been designated the region of interest. In this way it was possible to mesh a 1 μm × 1 μm × 5 μm volume with approximately 2.5 million tetrahedra joining 377 000 nodes. This number of tetrahedra should be storable in less than 1 GB. Gmsh performed the meshing in approximately 2 min.*

The FEA problem associated with a mesh of this size requires solution of almost 377 000 equations in as many unknowns. (The number of equations is equal to the number of nodes reduced by the in this case relatively small number of constrained nodes.) We had some initial difficulties solving problems of this size with GetDP. Despite that GetDP is a 32-bit application, and hence should permit memory usage up to 2 GB, GetDP would fail with an out-of-memory error when usage approached 1 GB. This cost some delays for trouble-shooting, but was eventually resolved by recompiling GetDP with a substitute linear algebra library. The solution requires approximately 15 min including the time for reading the mesh and writing the output.

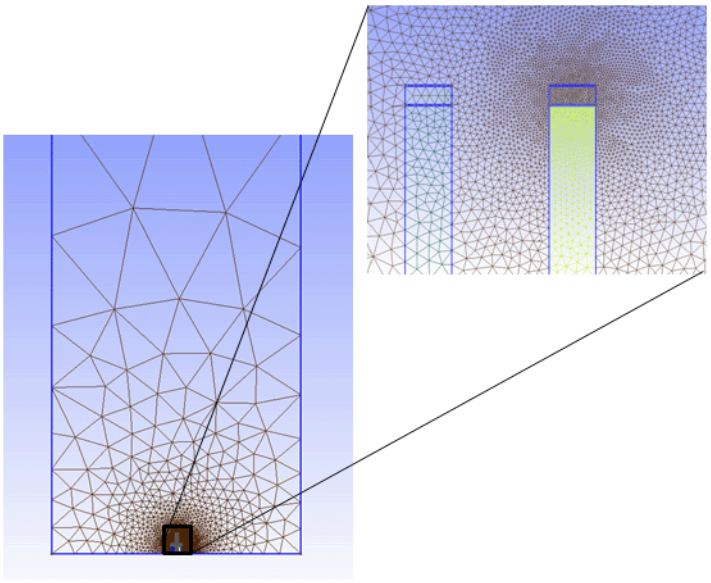

Fig. 22. Variable mesh size. The number of mesh elements is limited by computer memory and available computation time. The best accuracy for a fixed number of mesh elements is obtained by making the size small only where necessary.

With the problem thus in-principle solved, it remained to integrate the solution with JMONSEL's Monte Carlo simulations. The new JMONSEL program classes implemented for this integration are called Mesh, MeshElementRegion, MeshedRegion, and ChargingListener. Brief descriptions of the functions of each of these follow.

The Mesh class is JMONSEL's internal representation of the tetrahedral mesh produced by Gmsh. At over 2000 lines of code and comments, it is by far the largest of the new classes. The Mesh class obtains its data by reading a mesh file previously produced by Gmsh. It can also read problem resolution files produced by GetDP. These contain the solution (potentials at all nodes) of the FEA problem. The Mesh class's duties include storing the data associated with the mesh and providing access to that data to other program modules. The data include, for example, the coordinates of all nodes in the mesh, a list of all the tetrahedral mesh elements including for each one the indices of its nodes, information on the mesh topology (which tetrahedron if any is adjacent to which other tetrahedron through a given face), the potentials at each of the nodes, the charge contained in each of the tetrahedral elements, etc.

* Times reported here and in the next paragraph were from a PC equipped with Intel Xeon E5520 2.26 GHz processor, 12 GB of RAM, and Windows XP-64 operating system.

The Mesh class contains a number of subclasses—Element, Line, Triangle, and Tetrahedron—that are internal representations of the components of a mesh. E.g., a Tetrahedron is a volume element that represents tetrahedral mesh elements, a Triangle is a two-dimensional component that represents one of the 4 faces of a tetrahedron, and a Line is a one-dimensional component that represents one of the edges of a triangle. It also contains a subclass, MeshedShape, that represents the boundary of the meshed volume (i.e. the boundary between it and the surrounding unmeshed space). Most notable among these are the Tetrahedron and the MeshedShape. The Tetrahedron class does double duty. JMONSEL needs to be able to determine whether a given point lies within a particular shape. I.e., the shape must be able to report whether a point lies within it or not. Also, if an electron moves from $\hat{x}_0$ to $\hat{x}_1$ before it scatters, we need to know whether it intersects a boundary of the shape somewhere in the intervening interval. The Tetrahedron class fulfills these functions for mesh elements. Additionally, the FEA requires that we can determine the charge density within a particular mesh element, compute the electrostatic potential and associated electric field at a point, and determine the mesh element that an electron enters when it leaves its present one through a given face. The Tetrahedron class also fulfills these functions. The MeshedShape performs similar functions for the mesh as a whole. This is mainly required when electrons are outside of the meshed region. In that case, it is sometimes necessary to determine which facet of the mesh is first struck by an electron on a given path. When the electron is inside the mesh, this is not necessary because it is then also inside a known tetrahedral mesh element—it must intersect one of the boundaries of that element before it can intersect any others.

The MeshElementRegion and MeshedRegion classes are new types of Regions. To explain their function, some background is necessary. JMONSEL describes samples by dividing them into regions. Each region has a shape (the shape of its boundary) and a MaterialScatterModel. The latter is a representation of the material that occupies the region and the scattering properties of that material. These regions were represented by a Region class. Among the properties of this class: Regions could have wholly contained subregions nested to any depth, they could be translated and/or rotated, they could find and return the innermost region that contains a given point. JMONSEL accesses these properties of a Region to determine whether and how electrons scatter. If we are to use meshed samples, analogous structures are required to perform these functions. The existing Region class could not be used because the mesh differs in some important respects from JMONSEL's previous regions. For example, by associating a MaterialScatterModel with a Tetrahedron, we make a MeshElementRegion. This has many of the properties of JMONSEL's previously existing Regions, but MeshElements cannot be independently rotated or translated, since this would destroy their relationship to the rest of the mesh. Neither can MeshElementRegions have subregions. On the other hand, there is no reason not to translate or rotate the mesh as a whole. The new MeshedRegion class represents the entire mesh. Like JMONSEL's prior Regions, a MeshedRegion is rotatable, translatable, and can have subregions. Unlike them, its subregions can only be MeshElements, and therefore they can only be nested one level deep. The new classes also make use of some advantageous properties of meshes. In a mesh, interior elements share each of their faces with a single neighboring element. That is, there is topological information that is missing in the case of JMONSEL's original Regions. This information is advantageously employed to render certain operations (boundary crossings, locating the innermost containing region for a point) more efficient. The MeshedRegion class also knows how to write the input files required by GetDP to perform an FEA solution of its current state, and it knows how to read the results of such an analysis to update the potentials on all its nodes.

The ChargingListener class performs the iteration through steps 2, 3, and 4 shown in Fig. 20. During Monte Carlo simulations of trajectories, the simulator takes note of certain events (scattering event, boundary crossing, trajectory start, trajectory end, SE generation, …). Whenever such an event is generated, all "Listeners" that have been registered with the simulator are informed. This is NISTMonte/JMONSEL's standard mechanism for implementing detectors. With the new ChargingListener class, it is also used for charge modeling. The ChargingListener takes these actions:

1. On trajectory start events (which happen when an electron is generated in the model's electron gun) it increments a counter. If the value in the counter is not equal to a user-specified N, it simply returns. Otherwise it resets N to 0 and runs an FEA. Running an FEA means it causes GetDP's required input files to be written into a temporary folder. These input files contain the problem specification: the mesh, values of dielectric constants, constraints, charge densities, etc. It then launches an instance of GetDP, waits for GetDP to finish, then reads the result files.
2. On trajectory end or secondary end events (which happen when an electron is dropped from the simulation because it has become trapped, either through its energy falling below a user-specified threshold or through the operation of an explicit trapping model) the region containing the electron has its total charge decreased by e (where $-e$ is the charge on the electron).
3. On a secondary generation event, the containing region's total charge is increased by e.
4. A scattering event or boundary crossing event terminates an electron's "step". On such events the ChargingListener obtains the electric field inside the Tetrahedron where the electron resides, and it corrects the trajectory for the effect of this field.

Most of the above operations are, of course, internal to the program. The only ones visible to a user are those that change the procedure for running a simulation that includes charging. The new elements of the procedure are these:

1. The sample, at least those parts of it that need to be meshed, and surrounding space must be generated and meshed using Gmsh. Gmsh provides a graphical user interface to perform this operation in a manner similar to computer aided design programs. The user must take care to "tag" distinct regions. Tagging consists in assigning a number to a particular geometrical object. When the object is meshed, all mesh elements (e.g., tetrahedra) that are part of the object inherit that tag. The tags are later used to associate MaterialScatterModels and/or constraints with volumes or surfaces in the geometry. Therefore each volume or surface that has a distinct material or that will be subject to a boundary condition must be assigned a unique tag. Generally, it is advisable to use Gmsh's variable mesh size feature to assign a small mesh size to the critical region (typically the defect location for defect metrology or the edge location for CD metrology). The resulting mesh is then saved to a file.
2. In JMONSEL, the Mesh is generated by calling the Mesh constructor with the name of the mesh file produced in Step 1 as its argument. If necessary, the mesh can then be translated or rotated.
3. A MeshedRegion is then generated by calling the MeshedRegion constructor. One of its arguments is the mesh that was created in Step 2. Another argument is a user-provided map that associates the tag numbers of volume elements with MaterialScatterModels. The MeshedRegion constructor uses these to assign materials to the various parts of the mesh. Another argument is a similar map that associates tag numbers with constraints. Any regions or surfaces that are to be held at fixed voltages are specified in this way. The MeshedRegion is made a subregion of a

user-specified region, usually the chamber. Any number of MeshedRegions may be generated in this way. It is anticipated that most simulations will require only one. MeshedRegions should be large enough for fields to be close to 0 outside of them. The space outside of MeshedRegions is assumed to be field-free, since there is no mechanism for representation of fields in ordinary NISTMonte/JMONSEL regions.

4. The rest of the simulation description is the same as simulations that do not model charging, except that the user must create a ChargingListener and add it to the model's list of listeners. The ChargingListener is created by calling its constructor. Among the constructor's arguments are the value of the FEA interval, N, and the location of the temporary folder that will be used to write and read GetDP's input and output files.

4.3 Sensitivity Study

4.3.1 What is a sensitivity study?

Consider this analogy from uncertainty analysis. Suppose we have a function that depends upon two parameters, $f(a, b)$. If the values of the parameters are uncertain, the uncertainty in the function value is given by

$$u_f^2 = \left(\frac{\partial f}{\partial a}\right)^2 u_a^2 + \left(\frac{\partial f}{\partial b}\right)^2 u_b^2 \qquad \text{(EQ 3)}$$

where u_a and u_b are the standard uncertainties (e.g., the standard deviations, measured or estimated) in a and b. The derivatives of f with respect to the parameters are called the sensitivity coefficients. A large sensitivity for a particular parameter is an indication that small errors in that parameter produce large errors in f. Clearly, if one wants to measure f, it is good to know which parameters are sensitive ones and which insensitive. When we are uncertain about many things, this helps to prioritize those parts of the calculation of f for which we would like to do better. This example is not a perfect analogy to our modeling problem. For one thing, the difference between model #1 and model #2 usually cannot be expressed as the difference in a continuous parameter in some kind of overarching super-model. We are often faced with a choice between discrete alternatives. Since the difference is not a continuous variable, we cannot literally take a derivative. However, we can still examine the amount by which such choices affect outcomes. Such an examination is not a full uncertainty analysis. The difference, in analogy with Eq. 3 again, is that we are not at this stage necessarily committing to particular values for the u_x terms. In loose analogy to the example just given, we refer to this examination as a sensitivity study.

4.3.2 Why do we need a sensitivity study?

Many components of the theory that underlies an e-beam simulator remain uncertain. We already referred to some of these in Sec. 3.0. As a consequence of these uncertainties JMONSEL often has implemented alternative ways to model the same phenomenon. If we are to do better than make a random choice among the options, we must know what differences these choices make and compare results with different choices to available data. Much of the data in the literature are in the form of yields and stopping powers at varying landing energies and target materials. A useful catalog of these was assembled by David Joy under a previous Sematech contract.[56,57] There are uncertainties in these data as well. SE yields measured by different groups often disagree by a factor of 2 or more. Some of these differences may be due to dif-

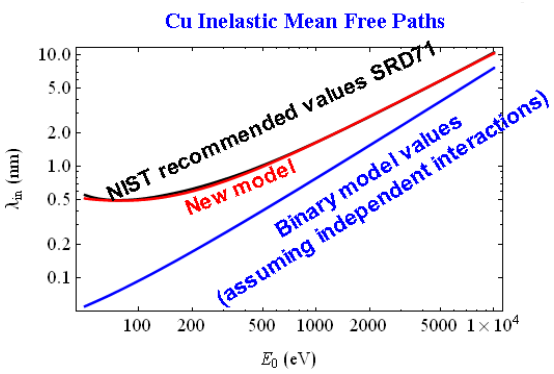

Fig. 23. Comparison of inelastic mean free paths in Cu from a binary collision model (blue), the present JMONSEL's DFT model (red), and NIST SRD71 (black).

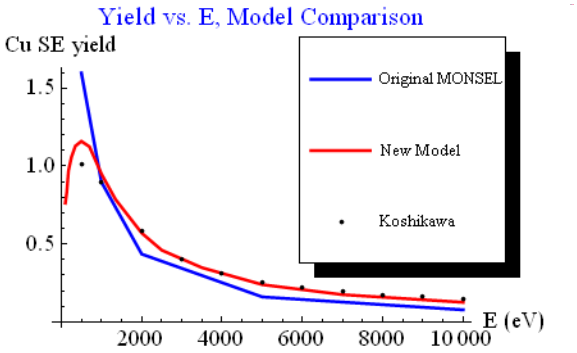

Fig. 24. Yield from a flat Cu sample vs. landing energy as measured [59], from the original MONSEL model (blue), and from JMONSEL (red).

ferent sample preparation or cleanliness; this suggests that not only model variations should be tried but also some sample variations.

The most obvious result of these differences is often a difference in predicted yield, which in turn implies a difference in image brightness. However, brightness is only indirectly relevant to industry use of an e-beam imaging tool. Measurement values are determined by algorithms applied to images, and the information in an image comes from *contrast*. A statement of the form "Model #1 produces 20% more yield than model #2" is not directly relevant to a metrologist's bottom line. We are looking instead for statements of the form "Using model #1 results in an X nm difference in measured CD value compared to using model #2." To make such statements we must perform alternative simulations of images with contrast, and we must apply appropriate measurement algorithms to them. We must see how differences in inputs propagate to what we care about. I.e., we must do a sensitivity study.

4.3.3 The study and its results

The DFT SE generation model in JMONSEL was developed using Cu as a test case. The development was motivated in part by the results shown in Fig. 23. The black curve shows the inelastic mean free path in Cu from NIST SRD 71 for electrons with energy E_0. The blue curve shows the result from a binary (omits screening) scattering model. JMONSEL's DFT model result (red curve), which includes screening, agrees with the reference values. The inelastic mean free path is the average distance between SE generation events. Short mean free path means a large number of SE are generated per unit path length. Other things being equal, this would increase the predicted yield of the model. The effect on predicted yields is shown in Fig. 24, where it is compared to measured values. MONSEL's fitting parameter allowed the excess SE generation to be compensated by higher stopping power, allowing overall yield to be matched to experiment at one energy. The decrease of yield with energy had approximately the right functional dependence at high energies, so it was possible to match the resulting curve to measured yields towards the right side of the graph. However, the model could not reproduce the position and magnitude of the observed yield maximum at lower energy. The revised model in JMONSEL does a better job of this.

As mentioned in Sec. 4.1.2, as part of the model improvements for this project we increased the number of choices for elastic scattering cross sections from one (Browning) to four (Browning, NIST SRD 64 table, Czyzewski table, and screened Rutherford). Elastic scattering plays an important role in the backscattered electron yield.

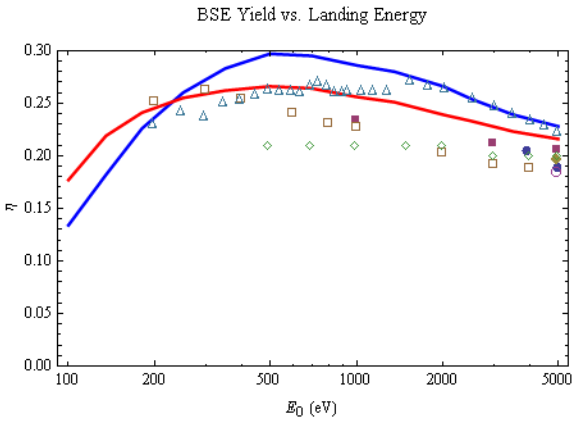

Fig. 25. Si backscattered electron yields vs. landing energy from JMONSEL with elastic cross sections determined by the Browning empirical form (blue) or by interpolation of NIST SRD 64 tabulated Mott cross sections (red). Data for comparison (symbols) are from a number of sources as described in and digitized from Fig. 4c of Ref. 60.

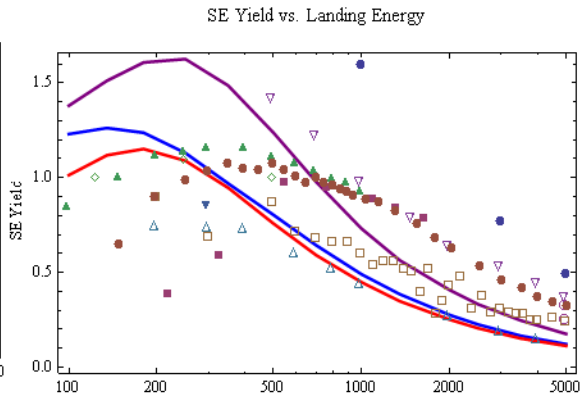

Fig. 26. Secondary electron yields vs. landing energy from JMONSEL for clean Si (middle curve, blue), Si with a thin SiO_2 layer (upper curve, purple), and Si with SiO_2 and carbon layers (lower curve, red). Data for comparison are from Joy's database and Walker et al.[56,57,61]

With more than one option, we were able observe the sensitivity of predicted yield to the different choices (Fig. 25), The backscattered electron yield itself plays a role in SE yield, because backscattered electrons, like any higher energy electron that passes within the SE escape depth of a surface, can generate SE that can escape (the so-called SE2s).

Si SE yield vs. energy has proven difficult to match. Partly this is because of the very large scatter in measured yields (symbols in Fig. 26) in the database. We simulated yield vs. energy for a large number of different model choices and assumptions about the sample's surface state. The model variables included:

1. SE generation mode (assumptions about the initial energy and momentum of the target electron)
2. Assumed density of states. The default free-electron density of states was suspected of being a poor approximation for Si, so a more realistic treatment was implemented, but the results turned out to be quite insensitive to this difference.
3. Interface barrier width. If we assume an abrupt change (compared to the electron's wavelength) at the surface barrier, transmission is reduced relative to a gradual barrier. The ratio of yields does not change either much or rapidly with energy, however, so the effect is mostly an overall scaling of the yield curve without much change in its shape.
4. Presence of absence of SiO_2 and C (modeled by graphite) contamination layers. These had strong effects on simulated yields.

Various combinations of the variables were tried. A sample of three of the resulting curves is shown in Fig. 26 in comparison to measured yield data.

The position of left and right edges of a feature defines its width. The effect of changing the instrument model upon the assigned edge position was studied by performing simulations on Cu and Si trapezoidal edges (as in Fig. 27). Simulations were performed for nine values of edge angle in the range 0° to 10°, four beam landing energies from 500 eV to approximately 4.95 keV (Si) or 5 keV (Cu), three extraction values (i.e., degrees of interpolation between full extraction and no extraction, as

described in Sec. 4.1.4) and 5 beam width values (the standard deviation of an assumed Gaussian beam, ranging from 0 nm to 5 nm). The SE generation models used were MONSEL's original binary model, the phenomenological "backup" model described in Sec. 4.1.1, and the DFT model. Within each of these main model categories there were further variants for choice of interface scattering (abrupt vs. gradual barrier), stopping power model, or (for the DFT model) the method of determining the target electron's initial energy and momentum. All together, eight different models or model variants were tried.

An example of the different shapes of intensity profiles predicted by the different models is shown in Fig. 27. Differences in edge position from one model to another were judged by shifting and scaling the test model to best agree (in the least squares sense) with a reference model. Each of the eight model variants was given a turn as the reference model for these comparisons. Histograms from a subset of the results (landing energy chosen close to a typical one for CD-SEM) are shown in Fig. 28. Differences between the models resulted in differences of approximately 1 nm in edge position assignment. This translates to 2 nm in CD assignment, where two edges are involved.

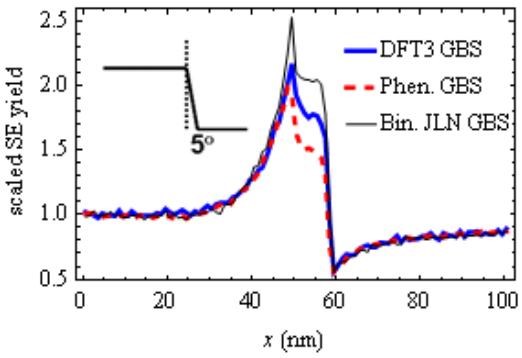

Fig. 27. Effect of SE generation model for 1 keV incident electrons on a 5° Si edge.

We have called this a sensitivity study because what it does in the first instance is quantify the effect of different model assumptions upon a CD result. Perhaps some further discussion is in order concerning the relationship between these values and CD uncertainties.

One difference is that the total uncertainty includes components from sources other than modeling errors. Noise in the measurement, vibration,… Some of these things are random errors, and their magnitudes can be estimated by the repeatability of measurements. Measurement "precision," as this is often called, in a good industrial CD-SEM is nowadays a fraction of a nanometer. Ease of measurement makes these precisions well-known numbers. They are sometimes mistaken for total uncertainties

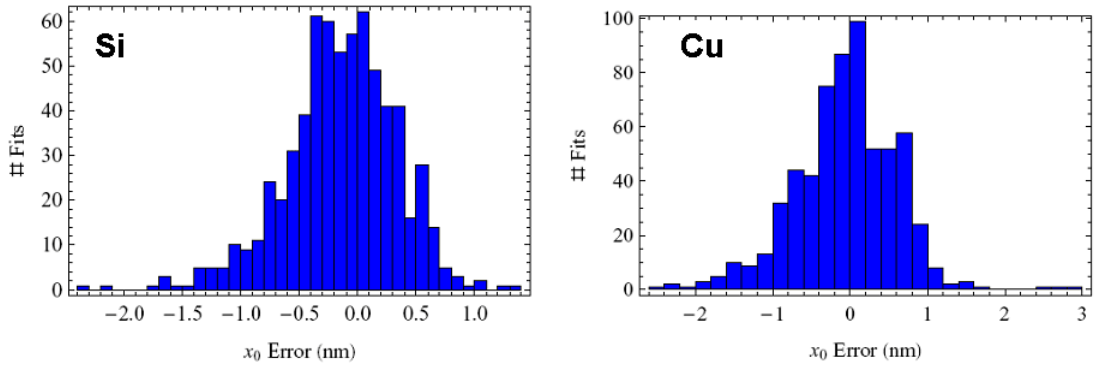

Fig. 28. Histograms of edge position differences in 662 fits at 1 keV landing energy and varying choice of model on Si (left) and Cu (right). For Si, 95% of single-edge errors fall within ±1 nm. On Cu 95% were within ±1.3 nm. From Ref. 48, used by permission.

even though they leave out, as our discussion above should have made obvious, other significant sources of error. The approximately 2 nm differences in results between the various model-based methods used in the present study are small compared to the differences between the typical CD algorithms (e.g., based upon intensity thresholds or slopes) employed for industrial measurements [31] and differences between different measurement techniques.[29,30]

An uncertainty is an estimate of error. Error is the difference between our result and the truth. In a manner of thinking, we can consider the true result to be the one we would get if we could use "nature's true model" instead of only our best estimate of it. To equate the magnitude of our observed 2 nm differences with the modeling uncertainty requires an additional assumption, the assumption that the errors in our models are random, uncorrelated, and similar in magnitude to the differences between them and the true model. If that were so, then an estimate of 2 nm model uncertainty would be appropriate. Is it so? We would like to think that these assumptions are pessimistic, that we are converging on the truth, therefore that differences between our best model and models that have recognizable shortcomings *over*estimate our current errors. However, in the nature of things, this is not something easily proven. For many materials, absence of the right kind or good enough quality of measurements hinders our ability to distinguish between models.

4.4 Demonstrations of model functionality

On two occasions, simulations were performed for sample shapes requested by Sematech. Additionally, upon initial completion of the charge modeling modules, some initial simulations were performed to exercise this part of the simulator. This section describes those simulations and shows the results. These were not intended to be detailed studies. Rather, they were meant to demonstrate software functionality for relevant samples.

4.4.1 A contour metrology demonstration

The first set of simulations was on a series of array patterns (dense array, comb, etc.) as shown in Fig. 29. In this figure the light blue design patterns were submitted by Sematech. A sample description was then created in JMONSEL. In this description, the line features in these arrays were 30 nm high and 30 nm wide. The closest spacings were also 30 nm. The samples were Si with a 1.3 nm surface oxide. The beam landing energy was 600 eV. The indicated portions of the samples were simulated, and the resulting images are shown.

As an illustration of a potential use for this type of simulation, consider Fig. 30. This figure reproduces the simulated image of the interleaved comb structure on the left. On the right the same intensity values are rendered as contour plots of successively increasing magnification, such that the plot at lower right is narrowly focused on the edge of one of the lines as it transitions from an isolated region (i.e., where there are no nearby neighbors to that side) to a dense region (with a nearby neighbor). The simulated sample had perfectly straight edges. I.e., there is no *actual* shift in edge position. The fact that the simulated intensity contours do have a shift of a bit over 1 nm means contour metrology could have measurement error of about this size. The red and blue contours were results from two simulations with different pixel sizes, to demonstrate that pixel size was not significantly affecting the interpretation.

After the preliminary results described here, work on contour metrology was continued under separate Sematech funding of a small project for that purpose. The additional work included careful comparison of AFM and SEM measurements of the same

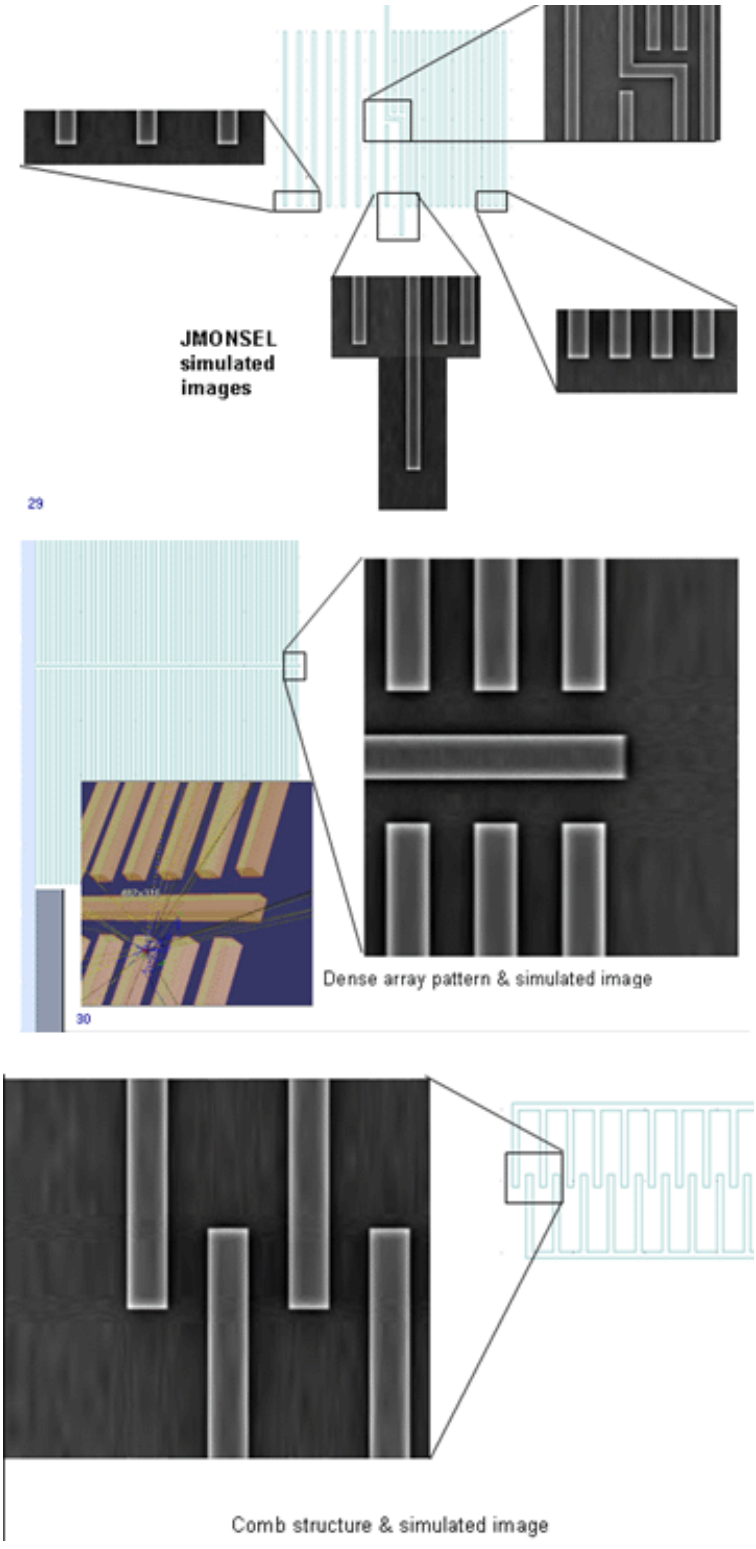

Fig. 29. Simulated images of selected portions as shown) of various array structures submitted by Sematech (light blue). The three-dimensional rendering in the middle group is a screen shot of the view of the simulator's virtual reality modeling language (VRML) output.

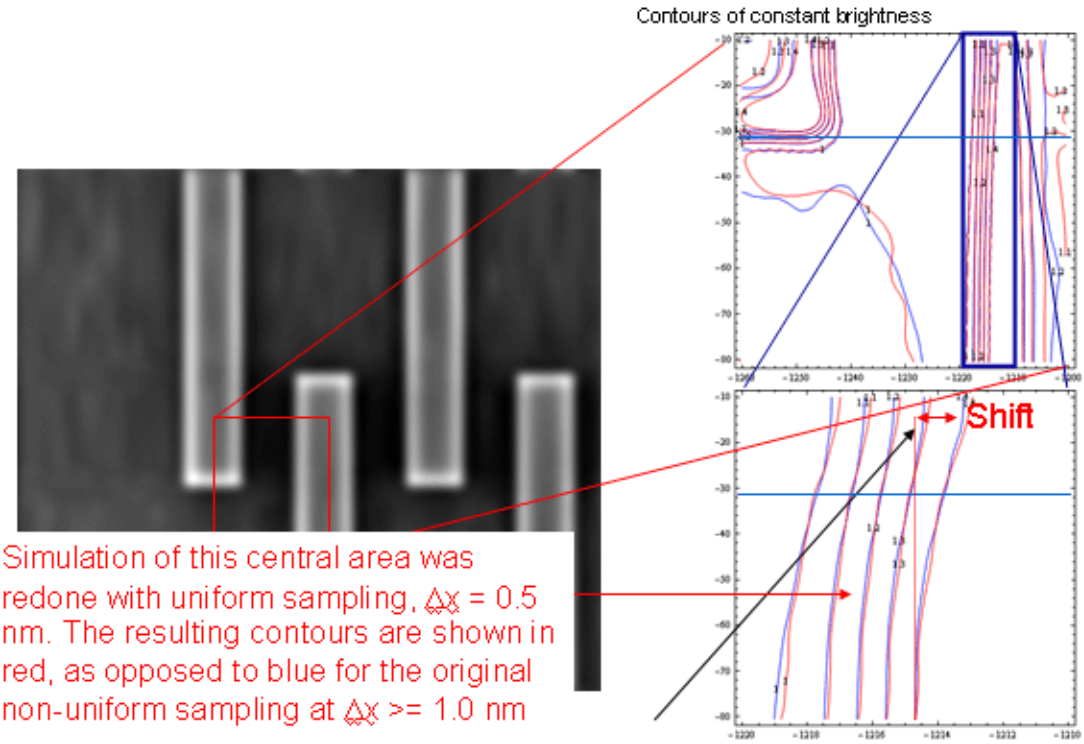

Fig. 30. Detail from simulated SEM image showing how contours of constant brightness shift between isolated and dense regions of a sample.

region of a sample that contained lines with a transition between isolated and dense areas. This comparison confirmed the model's prediction of an iso/dense metrology offset if the SEM contours are assigned based on a fixed brightness threshold, as in the above case. The bias is significantly reduced if relative brightness levels are used. These results have been published separately.[58]

4.4.2 A non-charging defect demonstration

The second set of simulations (Fig. 31), are from IDAs. An IDA is an intentional defect array. It may be used, for example, to qualify a defect detection or inspection tool. The left member of each pair is an IDA image supplied by Sematech. The right image is simulated for a similar structure. In JMONSEL's sample description, no attempt was

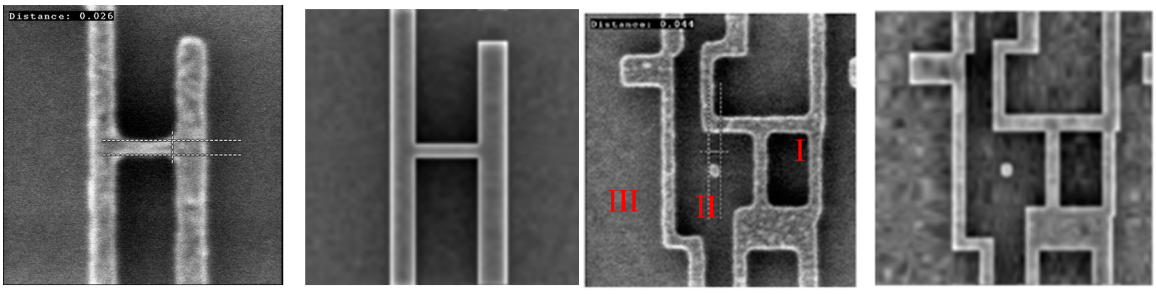

Fig. 31. Simulations of intentional defect arrays. In each pair the left image is a measurement, and the right image is a simulation of a similar structure (with some details, like roughness, granularity, and corner rounding omitted).

made to capture all of the details (edge roughness, granularity, corner rounding, etc.) of the actual samples. Nevertheless, the model structure is close enough that the image can demonstrate qualitative agreement with many of the features of the actual images. Note, for example, the similarity of brightening along feature edges and corners. Note also the variation of average background brightness levels depending upon the neighborhood. In particular, in the roughly H-shaped IDA on the left, the real and simulated images show similar darkening in the neighborhood of the horizontal element that connects the vertical lines. In the more complex pattern shown by the right image pair, there are distinct neighborhoods (labeled I, II, and III). The background brightness in these neighborhoods goes from darkest to brightest as the proximity of enclosing walls goes from closest to farthest. The background in the corresponding regions of the simulation exhibit similar behavior.

4.4.3 Electron trajectories with charging model

The charging model was exercised on a sample similar to those already shown in Fig. 21 and Fig. 22. The meshed volume was 1 μm × 1 μm × 5 μm. The lines were 100 nm tall (along the z direction) PMMA, 25 nm wide (x direction) and 500 nm long (from $y = -250$ nm to $y = 250$ nm). The landing position was in the center of the line's length (i.e., at $y = 0$ nm) and close to the right edge of the left line, as shown in Fig. 32. To charge the PMMA, 100 electron trajectories were run before beginning recording of trajectories for the trajectory image. Then the next 50 electron trajectories were recorded. These are shown in black (vacuum), green (PMMA), and blue (Si) in the figure. The Si was unmeshed and grounded. Curved trajectories in vacuum like the ones visible here are not seen in simulations without charging, since in that case there are no electric fields to cause the paths to deviate from straight lines.

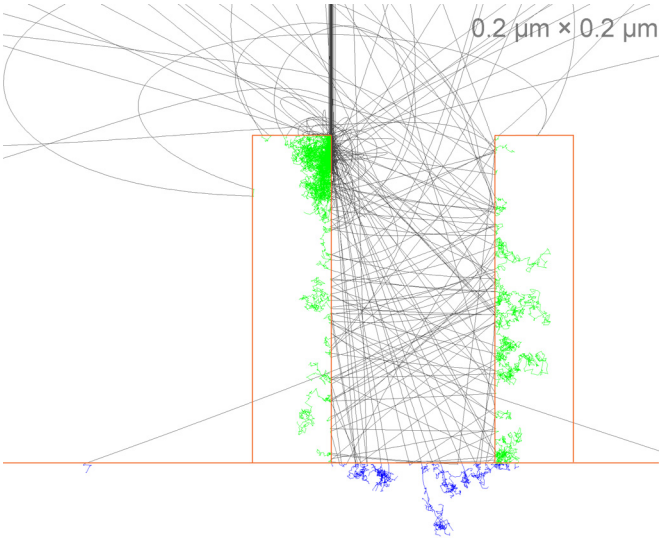

Fig. 32. Trajectories of electrons incident near the edge of PMMA line with charge model included.

5.0 Conclusions and future work

The JMONSEL simulation models have been improved in a number of respects:

1. Three new options for elastic scattering have been added. One of these determines the relevant cross sections by interpolating "NIST Electron Elastic-Scattering Cross-section Database" (SRD 64).
2. The method of calculating scattering tables for the dielectric function theory (DFT) secondary electron (SE) generation model has been generalized to include semiconductors and insulators.
3. The required energy loss function (ELF) input data for scattering table calculations have been collected for 40 elemental solids and 17 compounds. The source of many of these was the same input data used for contributions to the "NIST Electron Inelastic-Mean-Free-Path Database" (SRD 71).

4. Scattering tables have been computed for 17 materials.
5. A backup SE generation model has been written. The backup model has different and generally less stringent prerequisites than the DFT model. Many materials for which the requisite ELF data are unavailable can be simulated using the backup model instead.
6. An interface scattering model allows the potential energy barrier between materials to be approximated with both a barrier height and width.
7. Electron energy loss through electron-phonon processes and electron trapping can be simulated.
8. A charging model has been implemented. The charge distributions are determined by the scattering and SE generation model. Fields are computed using finite element analysis with the sample and neighboring space tetrahedrally meshed. Electron trajectories are adjusted for the effect of the electric field.

Simulations have been performed for a number of different model choices to observe the sensitivity of results to these choices. Such tests included (1) the effect of elastic scattering model on backscattered electron yield, (2) the effect SE generation model and presence or absence of SiO_2 and carbon films on Si yields, and (3) the effect of a large number of different model choices on edge position assignments. Results were compared to measurement data when such data were available.

Preliminary modeling results using the new insulator-modeling capabilities (Fig. 19) are consistent with experimental data from Toshiba[51] showing how native oxide growth affects SEM imaging of silicon trench features.

Image simulations of electrical test structures at the poly gate level show an iso/dense shift in contour lines representing constant brightness.

The charging model has been exercised on a sample consisting of 25 nm wide PMMA lines on a grounded Si substrate. Trajectory plots indicate that all of the parts of the model are working together.

Future work will add new reporting capabilities to the charging model. For example, it is desirable to obtain data showing the charge and potential distributions in user-specified planes. There is a need for comparisons between simulation results and measurements to validate the models. There is the possibility to improve upon some approximations (e.g., the single-pole approximation) that were employed to determine the initial set of scattering tables. Finite element solutions for finely meshed 3-D samples inherently involve solution of hundreds of thousands of simultaneous equations. This computationally demanding job is time consuming. It is desirable to investigate the possibility of improving the speed.

References

1. H. Bethe, “Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie,” *Materie Ann. Phys*. **5**, 325-400 (1930).
2. H. A. Bethe and J. Ashkin, “Passage of Radiations Through Matter,” in Emilio Segrè (ed.), *Exp. Nucl. Phys.* **1**, 166–357 (1953).
3. D. C. Joy and S. Luo, “An Empirical Stopping Power Relationship for Low-Energy Electrons,” *Scanning* **11**, 176-180 (1989).
4. T. S. Rao-Sahib and D. B. Wittry, “X-Ray continuum from thick elemental targets for 10–50 keV electrons,” *J. Appl. Phys.* **45**, 5060–8 (1974).
5. C. G. Frase, D. Gnieser, and H. Bosse, “Model-based SEM for dimensional metrology tasks in semiconductor and mask industry,” J. Phys. D: Appl. Phys. **42**, 183001 (2009).
6. R. L. Myklebust, D. E. Newbury, and H. Yakowitz, “NBS Monte Carlo Electron Trajectory Calculation Program,” in *NBS Spec. Pub.* **460**, 105-128 (1976).

7. M. J. Berger, “Monte Carlo calculation of the penetration and diffusion of fast charged particles,” in Methods in Computational Physics, ed. by B. Alder, S. Fernbach, and M. Rotenberg (Adademic, NY, 1963) Vol. 1 pp. 135-215.
8. J. R. Lowney and E. Marx, “User’s Manual for the Program MONSEL-1: Monte Carlo Simulation of SEM Signals for Linewidth Metrology,” *NIST Spec. Pub.* 400-95, 1994.
9. J. R. Lowney, “Application of Monte Carlo simulations to critical dimension metrology in a scanning electron microscope,” *Scanning Microscopy* **10**, 667-678 (1996).
10. J. R. Lowney, “Monte Carlo simulation of scanning electron microscope signals for lithographic metrology,” *Scanning* **18**, 301-306 (1996).
11. N. F. Mott, Proc. Royal Society London A**124**, 425 (1929).
12. 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,” Journal of Vacuum Science & Technology B **9**, 3578-3581 (1991).
13. R. Browning, T. Z. Li, B. Chui, J. Ye, R. F. W. Pease, Z. Czyzewski, and D. C. Joy, “Empirical Forms for the Electron-Atom Elastic-Scattering Cross-Sections from 0.1 to 30 Kev,” Journal of Applied Physics **76**, 2016-2022, (1994).
14. R. Browning, T. Z. Li, B. Chui, J. Ye, R. F. W. Pease, Z. Czyzewski, and D. C. Joy, “Low-Energy-Electron Atom Elastic-Scattering Cross-Sections from 0.1-30 Kev,” Scanning **17**, 250-253 (1995).
15. H. Salow, "Sekundarelektronen-Emission," Phys. Z. **41**, 434-442 (1940).
16. J.L.H. Jonker, “On the theory of secondary electron emission,” Philips Res. Rep. **7**, 1 (1952).
17. A.J. Dekker, “Secondary electron emission,” Solid State Phys. **6**, 251 (1958).
18. H. Seiler, “Secondary electron emission in SEM,” J. Appl. Phys. **54**, RI (1983).
19. D. C. Joy, “A model for calculating secondary and backscattered electron yields,” J. Microsc. **147**, 51-64 (1987).
20. Y. Lin and D.C. Joy, “A new examination of secondary electron yield data,” Surf. Interface Anal. **37**, 895-900 (2005).
21. C. Möller, “Zur Theorie des Durchgangs schneller Elektronen durch Materie,” Ann. Phys. **14**, 531 (1932)
22. D. Pines and P. Nozieres, *The Theory of Quantum Liquids* (Benjamin, New York, 1966).
23. J. Lindhard, “On the properties of a gas of charged particles,” *Mat. Fys. Medd. Dan. Vide Selsk.* **28**, 1-57 (1954).
24. M. P. Davidson and A. E. Vladár, “An inverse scattering approach to SEM line width measurements,” *Proc. SPIE* **3677**, 640-649 (1999).
25. 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).
26. J. S. Villarrubia, A. E. Vladár, and M. T. Postek, “Scanning electron microscope dimensional metrology using a model-based library,” *Surf. Interface Anal.* **37**, 951-958 (2005).
27. 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).
28. 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).
29. V. A. Ukraintsev, “Effect of bias variation on total uncertainty of CD measurements,” *Proc. SPIE* **5038**, 644-650 (2003).
30. N. Rana, C. Archie, W. Lu, and B. Banke, “The measurement uncertainty challenge of advanced patterning development,” *Proc. SPIE* **7272**, 727203 (2009).
31. J. S. Villarrubia, A. E. Vladár, and M. T. Postek, “A Simulation Study of Repeatability and Bias in the CD-SEM,” *J. Microlith., Microfab., Microsyst.* **4**, 033002, (2005). This was a slightly revised update of a paper that first appeared in a conference proceedings, *Proc. SPIE* **5038** (2003).
32. N. W. M. Ritchie, “A new Monte Carlo application for complex sample geometries,” *Surf. Interface Anal.* **37**, 1006-1011 (2005).
33. J. S. Villarrubia, N. W. M. Ritchie, and J. R. Lowney, “Monte Carlo modeling of secondary electron imaging in three dimensions,” Proc. SPIE **6518**, 65180K (2007).
34. *International Technology Roadmap for Semiconductors,* 2007, Metrology, Table 3a.
35. C. J. Powell and A. Jablonski, NIST Electron Inelastic-Mean-Free-Path Database - Version 1.1, National Institute of Standards and Technology, Gaithersburg, MD (2000), http://www.nist.gov/srd/nist71.cfm
36. 2C. J. Powell, “Attenuation lengths of low-energy electrons in solids,” *Surf. Sci.* **44**, 29 (1974).
37. D.R. Penn, “Electron mean-free-path calculations using a model dielectric function,” *Phys. Rev. B* **35**, 482-486 (1987).
38. Z.-J. Ding and R. Shimizu, “A Monte Carlo Modeling of Electron Interaction with Solids Including Cascade Secondary Electron Production,” *SCANNING* **18**, 92-113 (1996).
39. *Handbook of Optical Constants of Solids* and *Handbook of Optical Constants of Solids II*, edited by E.D. Palik, Academic Press, 1985 and 1991 respectively.
40. W. S. M. Werner, M. R. Went, M. Vos, K. Glantschnig, and C. Ambrosch-Draxl, “Measurement and density functional calculations of optical constants of Ag and Au from infrared to vacuum ultraviolet wavelengths,” *Phys. Rev. B* **77**, 161404(R) (2008).
41. M.R. Went, M. Vos, and W.S.M. Werner, “Extracting the Ag surface and volume loss functions from reflection electron energy loss spectra,” Surf. Sci. **602**, p. 2069 (2008).

42. Z. J. Ding, X. D. Tang, and R. Shimizu, “Monte Carlo study of secondary electron emission,” *J.Appl.Phys.* **89**, 718 (2001).
43. S.F. Mao, Y. G. Li, R. G. Zeng, and Z. J. Ding, “Electron inelastic scattering and secondary emission calculated without the single pole approximation,” *J.Appl.Phys.* **104**, 114907 (2008).
44. A. Jablonski, F. Salvat, and C. J. Powell, NIST Electron Elastic-Scattering Cross-Section Database - Version 3.1, National Institute of Standards and Technology, Gaithersburg, MD (2002), http://www.nist.gov/srd/nist64.cfm
45. A. Jablonski, F. Salvat, and C.J. Powell, “Comparison of electron elastic-scattering cross sections calculated from two commonly used atomic potentials,” J. Phys. Chem. Ref. Data **33**, p. 409 (2004).
46. Z. Czyzewski, D. O. MacCallum, A. Romig, and D.C. Joy, “Calculations of Mott scattering cross sections,” J. Appl. Phys. **68**, p. 3066 (1990).
47. L.D. Landau and E.M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory*, (Pergamon Press, 1958) p. 75.
48. J. S. Villarrubia and Z. J. Ding, “Sensitivity of SEM width measurements to model assumptions,” *J. Micro/Nanolith. MEMS MOEMS* **8**, 033003 (2009).
49. J. Llacer and E.L. Garwin, “Electron-Phonon Interaction in Alkali Halides. I. The Transport of Secondary Electrons with Energies between 0.25 and 7.5 eV,” *J. Appl. Phys.* **40**, 2766-2775 (1969).
50. J.P. Ganachaud and A. Mokrani, “Theoretical study of the secondary electron emission of insulating targets,” Surf. Sci. **334**, 329-341 (1995).
51. M. Kadowaki, A. Hamaguchi, H. Abe, Y. Yamazaki, S. Borisov, A. Ivanchikov, and S. Babin, “Investigation of factors causing difference between simulation and real SEM image,” *Metrology, Inspection, and Process Control for Microlithography XXIII,* Proc. SPIE **7272,** 727231 (2009).
52. C. Geuzaine and J.-F. Remacle, “Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities,” Int. J. Numer. Methods. Eng. **79**, 1309-1331 (2009).
53. http://www.geuz.org/gmsh/
54. P. Dular, C. Geuzaine, F. Henrotte, W. Legros, “A General Environment for the Treatment of Discrete Problems and its Application to the Finite Element Method,” IEEE Trans. Magn. **34**, 3395 (1998).
55. http://www.geuz.org/getdp/
56. D. C. Joy, “Database of Electron-solid interactions and measurements of electron yields,” Sematech Technology Transfer 96063130A-TR, 1996.
57. D. C. Joy, “A database on electron-solid interactions,” SCANNING 17, 270-275 (1995). Also <web.utk.edu/~srcutk/database.doc>.
58. J. S. Villarrubia, R. D. Dixson, and A. E. Vladár, “Proximity-associated errors in contour metrology,” Proc. SPIE **7638,** 76380S (2010).
59. T. Koshikawa and R. Shimizu, “Secondary electron and backscattering measurements for polycrystalline copper with a spherical retarding-field analyzer,” *J. Phys. D* **6**, 1369–1380 (1973).
60. M.M. El Gomati, C. G. H. Walker, A. M. D. Assa’d, and M. Zadrazil, “Theory experiment comparison of the electron backscattering factor from solids at low electron energy (250–5000 eV),” SCANNING **30**, 2-15 (2008).
61. G.G.H. Wallker, M.M. El-Gomati, A.M.D. Assa’d, and M. Zadrazil, “The secondary electron emission yield for 24 solid elements excited by primary electrons in the range 250–5000 eV: A theory/experiment comparison,” SCANNING **30**, 365-380 (2008).