Focus Effects in Submicron Optical Lithography, Part 4: Metrics for Depth of Focus

Chris A. Mack
FINLE Technologies, Austin, TX 78716

Abstract

Common uses of the term "depth of focus" (DOF) are explored as it relates to semiconductor lithography. A definition of DOF is given which is most appropriate to photolithography for IC manufacturing and this definition is compared to other DOF metrics. In particular, simple methods for determining DOF (either experimentally or through simulation) lead to DOF-like metrics. These metrics are compared to the definition of DOF and their accuracy are evaluated. Examples of the use of the definition for DOF for studying trends in lithography are given.

Keywords: depth of focus, microltihography, optical lithography

I. Introduction

Focus effects continue to play a dominant role in extending optical lithography into the deep submicron regime [1-4]. When comparing technologies and processes for a given device generation, characterization of the depth of focus (DOF) is an essential part of the evaluation process. However, there is no universally accepted or standardized definition of DOF and the use of the term varies widely throughout the industry. This paper will begin by exploring the common uses of the term "depth of focus" and give a single definition which is most applicable to photolithography for IC manufacturing. Other "definitions," which typically include an experimental procedure for measuring DOF, will be described. In particular, pseudo-definitions for DOF involve the use of easier-to-measure metrics which are related to the actual DOF. Several of these simple metrics will be examined and their accuracy evaluated.

Finally, to demonstrate the value of using an accurate metric for DOF, several fundamental focus effects (such as DOF versus numerical aperture, wavelength, and feature size) will be explored using a simulation tool which can automatically evaluate the depth of focus.

II. Definition of Depth of Focus

Establishing a suitable definition for a commonly used concept such as depth of focus is not necessarily an easy task. Uses of the term vary widely, and some examples will be discussed in the next section. In general, DOF can be thought of as the range of focus errors that a process can

0-8194-1788-2/95/$6.00

tolerate and still give acceptable lithographic results. Of course, the key to a good definition of DOF is in defining what is meant by tolerable. A change in focus results in two major changes to the final lithographic result: the photoresist profile changes and the sensitivity of the process to other processing errors is changed. The first of these effects, the photoresist profile change, is the most obvious and the most easily observed consequence of defocus. Typically, photoresist profiles are described (in an oversimplified way) using three parameters: the linewidth (also called the critical dimension, CD), the sidewall angle, and the resist thickness of the feature (which is useful for lines or islands, but not spaces or contacts). In effect, the resist profile is modeled as a trapezoid, as shown in Figure 1. Usually it is more convenient to talk about resist loss (the difference between the original resist thickness and the final resist thickness), possibly as a percentage of the original resist thickness.

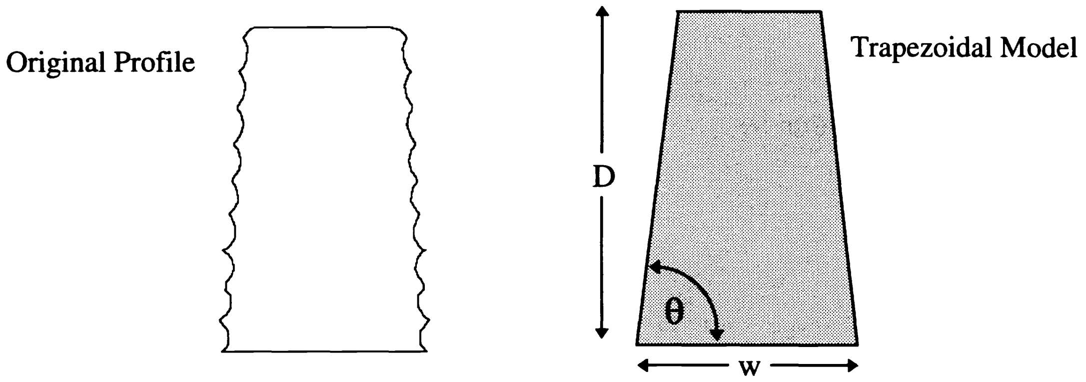
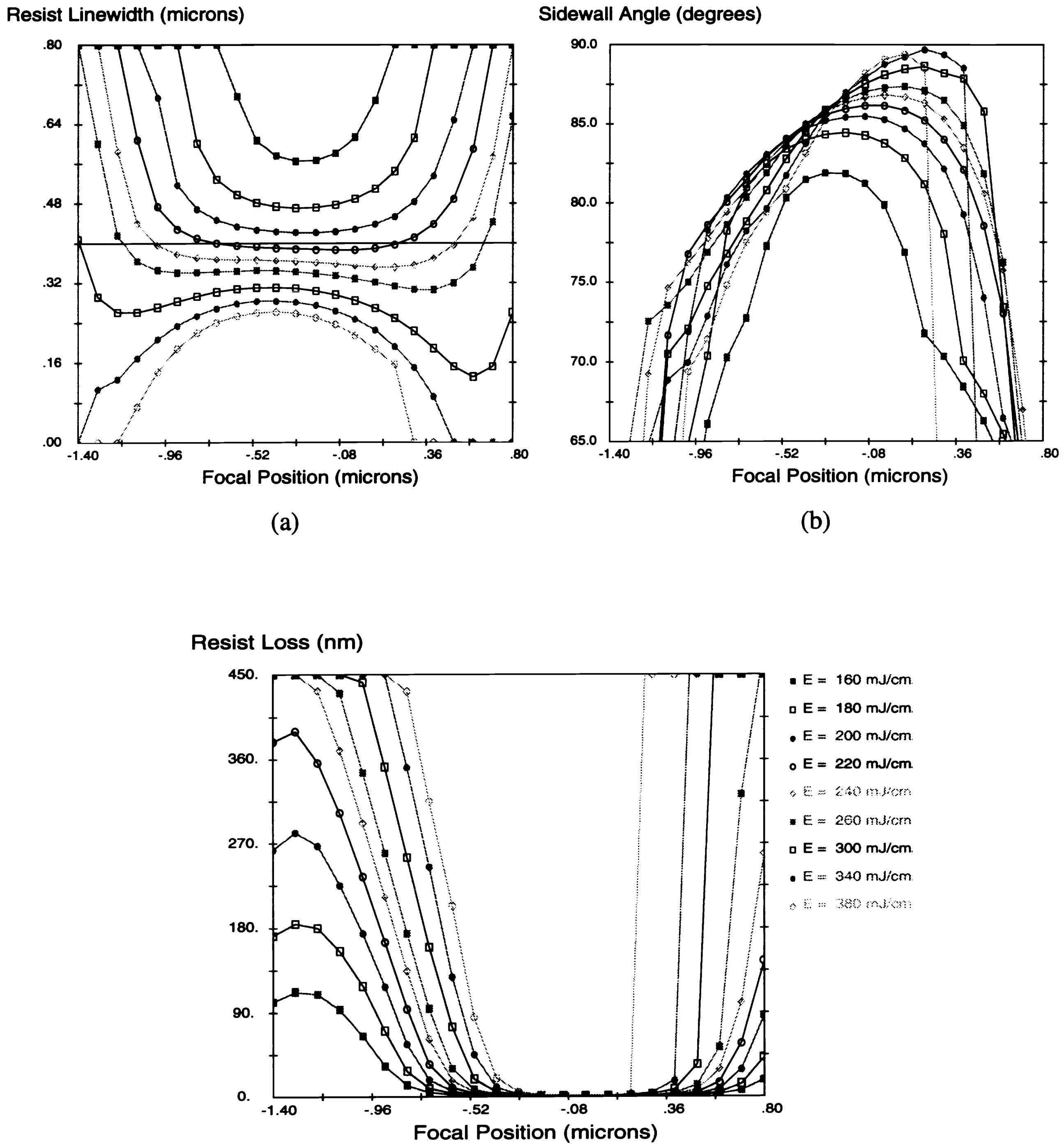


Figure 1. Comparison of an actual, complex photoresist profile with its trapezoidal model used to determine linewidth, sidewall angle, and resist loss.

The variation of linewidth, sidewall angle, or resist loss with focus can be readily determined for any given set of conditions. If these were the only responses of importance, specifications on these responses would lead to a simple definition of the depth of focus: the range of focus which keeps the linewidth, sidewall angle, and resist loss within their stated specifications. There is, however, a second affect of focus which is significantly harder to quantify and of great importance. As an image goes out of focus, the process becomes more sensitive to other processing errors such as exposure dose or develop time. Of these secondary process errors, the most important by far is exposure. To state the issue in another way, focus and exposure are coupled in their effect on the process.

Since the effect of focus is dependent on exposure, the only way to judge the response of the process to focus is to simultaneously vary both focus and exposure in what is known as a focus-exposure matrix. Figure 2 shows typical examples of the output of a focus-exposure matrix using linewidth, sidewall angle, and resist loss as the responses. The most common of these curves, Figure 2a, is called the Bossung plot [5] and shows linewidth versus focus for different exposures.



(a)

(b)

(c)

Figure 2. Examples of the effect of focus and exposure on the resulting resist profile: (a) linewidth, (b) sidewall angle, and (c) resist loss. Focal position is defined as zero at the top of the resist with a negative focal position indicating that the plane of focus is inside the resist.

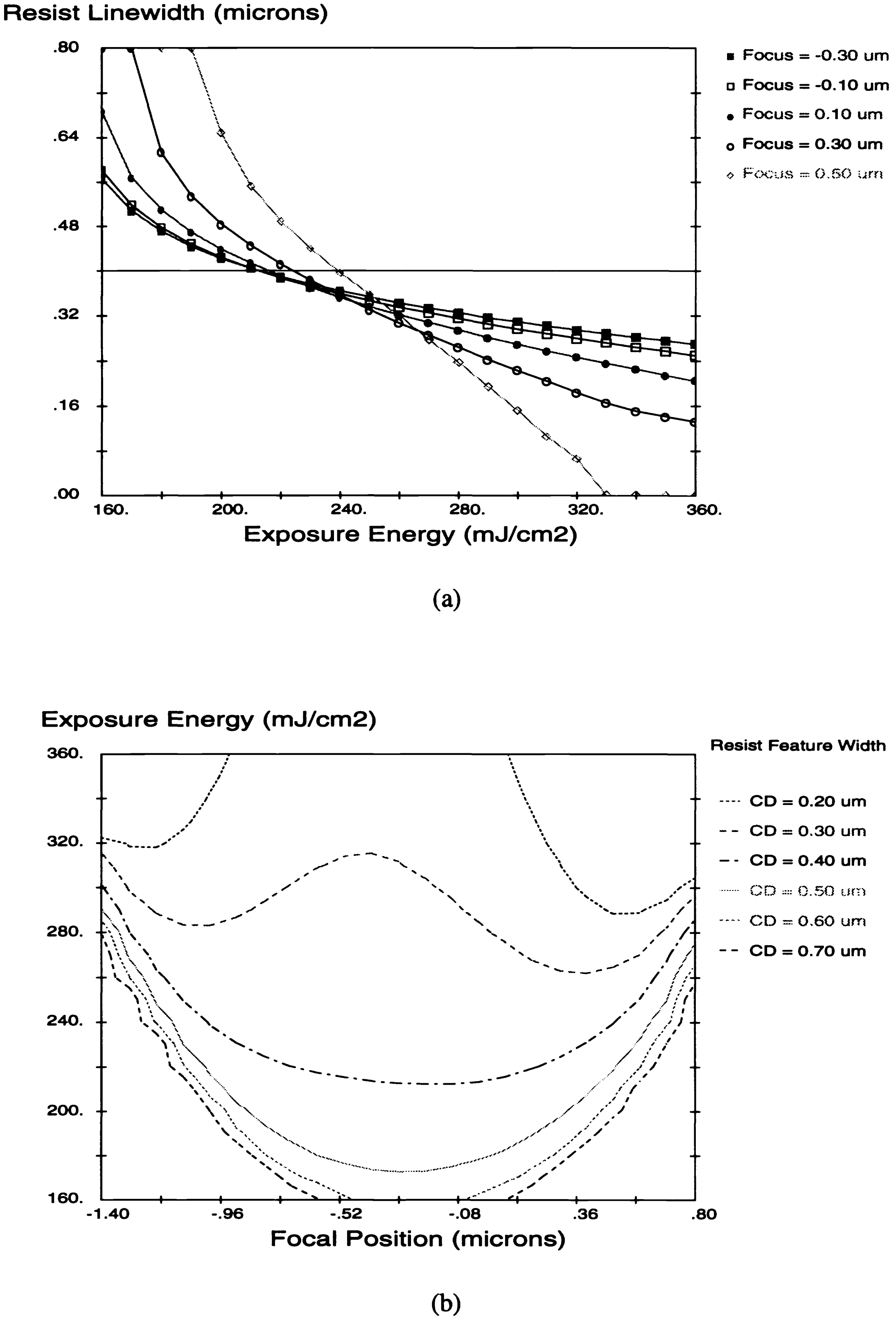


Figure 3. Displaying the data from a focus-exposure matrix in alternate forms: (a) decrease in exposure latitude resulting from defocus, and (b) contours of constant CD versus focus and exposure.

Each plot in Figure 2 contains a large amount of data and interpretation can become a problem. Of course, one output as a function of two inputs can be plotted in several different ways. For example, the Bossung curves of Figure 2a could also be plotted as exposure latitude curves (linewidth versus exposure) for different focus settings (Figure 3a). This is very useful in showing how defocus causes a reduction in exposure latitude. Probably the most useful way to plot the two-dimensional data set of CD versus focus and exposure is a contour plot -- contours of constant linewidth versus focus and exposure (Figure 3b). Obviously, sidewall angle and resist loss could also be plotted in these alternate forms if desired.

The contour plot form of data visualization is especially useful for establishing the limits of exposure and focus which allow the final image to meet certain specifications. Rather than plotting all of the contours of constant CD for example, as was done in Figure 3b, one could plot only the two CDs corresponding to the outer limits of acceptability -- the CD specifications. Because of the nature of a contour plot, other variables can also be plotted on the same graph. Figure 4 shows an example of plotting contours of CD (nominal ±10%), 80° sidewall angle, and 10% resist loss all on the same graph. The result is a *process window* -- a region of focus and exposure which keeps the final resist profile within all three specs (shown as the shaded area of Figure 4).

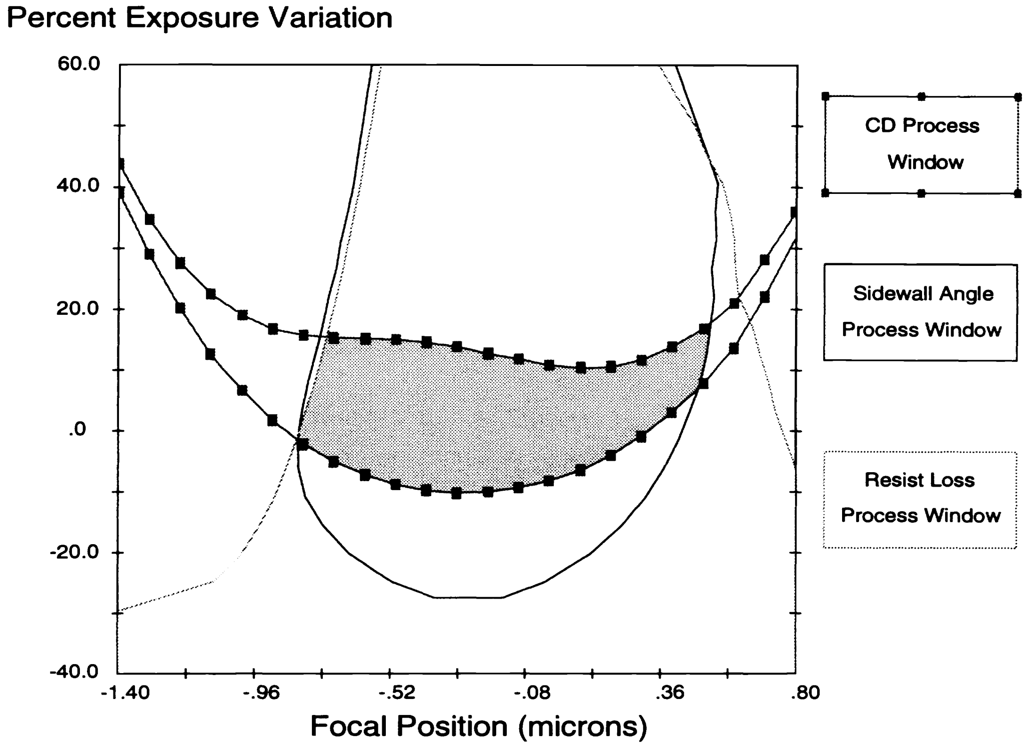
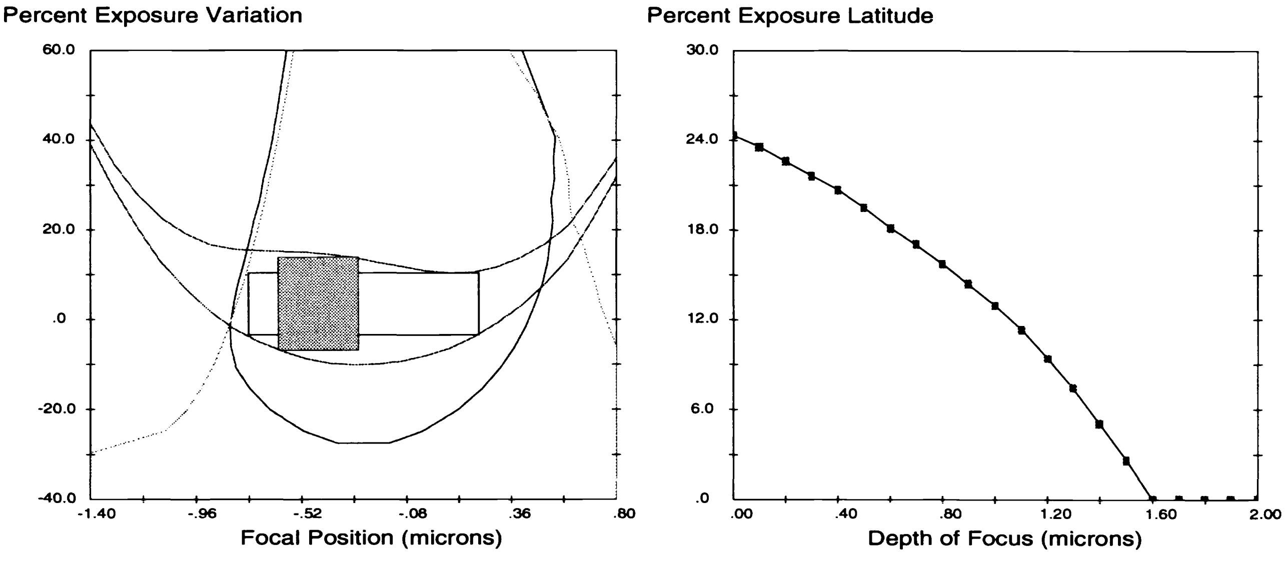


Figure 4. The focus-exposure process window constructed from contours of the specifications for linewidth, sidewall angle and resist loss. Shaded area shows overall process window.

The focus-exposure process window is one of the most important plots in lithography since it shows how exposure and focus work together to affect linewidth, sidewall angle and resist loss. The process window can be thought of as a *process capability* -- how the process responds to changes in focus and exposure. How can we determine if a given process capability is good enough? An analysis of the error sources for focus and exposure in a given process will give a *process requirement* [3]. If the process capability exceeds the process requirements, yield will be high. If, however, the process requirement is too large to fit inside the process capability, yield will suffer. A thorough analysis of the effects of exposure and focus on yield can be accomplished with yield modeling [6,7], but a simpler analysis can be used to derive a number for depth of focus.

What is the maximum range of focus and exposure (that is, the maximum process requirement) that can fit inside the process window? A simple way to investigate this question is to graphically represent errors in focus and exposure as a rectangle on the same plot as the process window. The width of the rectangle represents the built-in focus errors of the processes, and the height represents the built-in exposure errors. The problem then becomes one of finding the maximum rectangle which fits inside the process window. However, there is no one answer to this question. As Figure 5a shows, there are many possible rectangles of different widths and heights which are "maximum", i.e., cannot be made larger in either direction without extending beyond the process window. (Note that the concept of a "maximum area" is meaningless here.) Each maximum rectangle represents one possible trade-off between tolerance to focus errors and tolerance to exposure errors. Larger depth of focus can be obtained if exposure errors can be minimized. Likewise, exposure latitude can be improved if focus errors are small. The result is a very important trade-off between exposure latitude and DOF. Figure 5b shows an analysis of the process window where every maximum rectangle is determined and their height (the exposure latitude) is plotted versus their width (depth of focus).



(a) (b)

Figure 5. The process window (a) is analyzed by fitting all of the maximum rectangles, then plotting their height (exposure latitude) versus their width (depth of focus) as in (b).

An improvement can be made to the analysis of the process window in order to account for the statistical nature of focus and exposure errors. If all focus and exposure errors were systematic, then the proper graphical representation of those errors would indeed be a rectangle. The width and height would represent the total ranges of the respective errors. If, however, the errors were randomly distributed, then a probability distribution function would be needed to describe them. For the completely random case, a gaussian distribution with standard deviations in exposure and focus, σ_E and σ_F respectively, is used to describe the probability of a given focus error.

$$p(\Delta E, \Delta F) = \frac{1}{2\pi\sigma_E\sigma_F}\exp\left(-\Delta E^2 / 2\sigma_E^2\right)\exp\left(-\Delta F^2 / 2\sigma_F^2\right) \tag{1}$$

where focus errors and exposure errors are assumed to be independent. In order to graphically represent the errors of focus and exposure, one should describe a surface of constant probability of occurrence. All errors in focus and exposure inside the surface would have a probability of occurring which is greater than the established cut-off. What is the shape of such a surface? For fixed systematic errors, the shape is a rectangle. For a gaussian distribution, the surface can be derived by setting the probability of equation (1) to a constant, p^*.

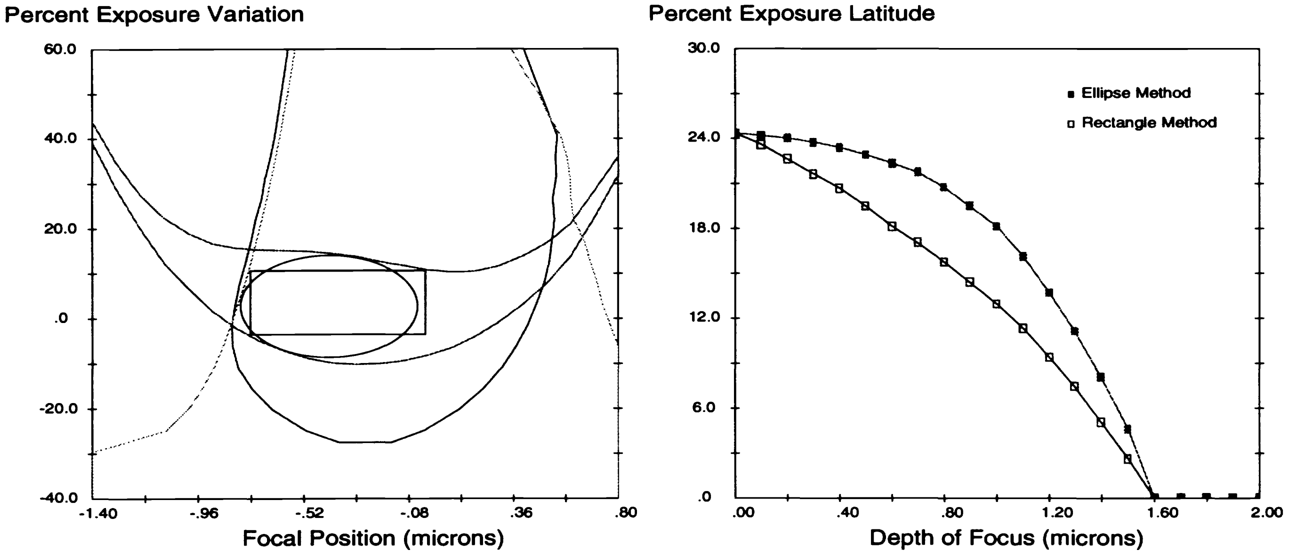
$$p^* = \frac{1}{2\pi\sigma_E\sigma_F}\exp\left(-\Delta E^2 / 2\sigma_E^2\right)\exp\left(-\Delta F^2 / 2\sigma_F^2\right)$$

$$-\ln\left(2\pi\sigma_E\sigma_F p^*\right) = \frac{\Delta E^2}{2\sigma_E^2} + \frac{\Delta F^2}{2\sigma_F^2} \tag{2}$$

Equation (2) is that of an ellipse. Suppose, for example, that one wishes to describe a "3-sigma" surface, where p^* corresponds to the probability of having an error equal to 3σ in one variable. The resulting surface would be an ellipse with major and minor axes equal to $3\sigma_E$ and $3\sigma_F$.

$$1 = \frac{\Delta E^2}{(3\sigma_E)^2} + \frac{\Delta F^2}{(3\sigma_F)^2} \tag{3}$$

Finding all of the maximum ellipses which fit inside the process window will also give the trade-off between exposure latitude and depth of focus. However, since it is the corners of the rectangles which limit their size in general, the ellipses which can fit inside the process window are larger, as seen in Figure 6. In reality, focus and exposure errors have both systematic and random components [3]. Thus, the two methods of analyzing the process window (rectangles corresponding to systematic errors and ellipses corresponding to random errors) will bracket the actual response of a real system. The rectangle method can be thought of as pessimistic, whereas the ellipse method is somewhat optimistic. An average of the two can also be used as a simple, more realistic metric.



(a) (b)

Figure 6. The ellipse, corresponding to a surface of constant probability for two random variables, results in a larger estimate of the depth of focus than the rectangle, which assumes only systematic errors.

Once a process window has been generated and analyzed to give the exposure latitude - defocus curve, a definitive value for the depth of focus can be obtained. The depth of focus can be defined as the range of focus which keeps the resist profile within all specifications (linewidth, sidewall angle, and resist loss) over a specified exposure range. For the example given in Figure 6, a minimum acceptable exposure latitude of 15%, in addition to the other profile specifications, would lead to the following depth of focus results:

DOF (rectangle) = 0.85 μm

DOF (ellipse) = 1.14 μm

DOF (average) = 1.00 μm

(Note: the days of quoting DOF as ± some distance are over. Focus behavior for small geometries is quite asymmetric so that only the total range has a useful meaning.)

Comparison to Other Metrics of DOF

The definition of depth of focus given above is physically significant and very flexible. It is independent of the means of obtaining the data (experimental or simulated data are analyzed in the same way), and can be applied in virtually any situation. In fact, many of the pseudo-DOF metrics in use in the industry are a subset of the definition given here. Consider the following DOF-like metrics.

A. Constant exposure DOF

This metric is used to reduce the number of data points that must be collected to determine the DOF. It has the added feature (some may say benefit) of exaggerating the reported depth of focus. In this metric the exposure is held constant (usually at the dose which gives the nominal CD, but it is sometimes varied in order to maximize the resulting DOF). Focus is varied and the range of focus required to keep the resist profile within spec is determined. No effort is made to evaluate the effect of focus on exposure latitude. The constant exposure DOF is a special case of the actual DOF, corresponding to an exposure latitude specification of 0. As can be seen from Figure 6b, this results in a maximum evaluation of the DOF.

Similarly, a constant focus exposure latitude can be defined as the range of exposure that keeps the linewidth within spec at best focus. This corresponds to the exposure latitude of Figure 6b when the DOF is 0. The combination of the constant focus exposure latitude with the constant exposure DOF is equivalent to drawing the largest *cross* inside the process window. One would expect that the size of the cross would be proportional to a more realistic measure of DOF. The problem with using only the constant exposure DOF is illustrated in Figure 7. Clearly process A is superior to process B in terms of DOF, but the constant exposure DOF metric would claim both processes give the same depth of focus. Simultaneous specification of both the constant exposure DOF and the constant focus exposure latitude (i.e., specifying the cross) would show the differences between these processes and result in a reasonably useful metric. *The constant exposure DOF by itself, however, can be very misleading.*

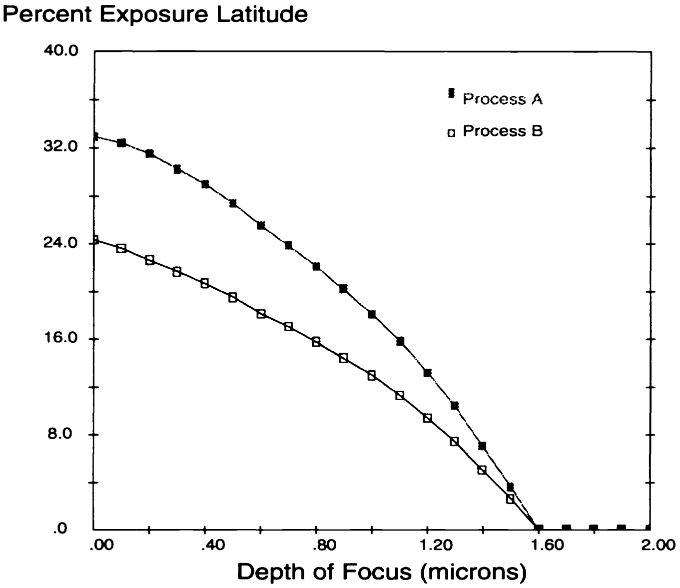


Figure 7. Comparison of two exposure latitude-depth of focus curves for different processes. Use of only the constant exposure DOF would conclude that these two processes had the same DOF.

B. Log-slope defocus curve

The log-slope defocus curve has become a very useful way to evaluate aerial images [1-4]. Because of the speed with which these results can be calculated, it has been explored as a way to estimate the depth of focus. In essence, the log-slope of the aerial image, evaluated at the nominal line edge, is proportional to exposure latitude. Thus, a plot of the degradation of the log-slope with defocus gives an indication of the nature of the full exposure latitude-depth of focus curve presented earlier. A minimum acceptable image log-slope corresponds directly to a minimum acceptable exposure latitude. There is one problem, however. The log-slope is proportional to the exposure range between the high and low CD specifications, but does not describe any curvature in the CD process window. Curvature in the process window is also expressed as an isofocal bias -- an out-of-focus image requires a different dose to size than an in-focus image. Often the process window curvature can have a large effect on the resulting depth of focus, causing a discrepancy with the log-slope defocus curve method. However, the simplicity and speed of the log-slope makes this method attractive for examining basic trends, especially when investigating image enhancement techniques.

C. SEMI specification for DOF

SEMI has published a standard definition for DOF entitled "SEMI P25-94, Specification for Measuring Depth of Focus and Best Focus." The language of the standard is almost vague enough that it could include a useful DOF measurement scheme. It is certainly vague enough that it could include many useless DOF measurement schemes. The specification describes depth of focus as

> "... the total distance of defocus where over the whole of the processed image field, the processed image is sufficiently resolved for practical use."

It goes on to explain that an acceptable photoresist image is "...where the variation of the pattern is not outside the permissible deviation..." In the strictest sense, the definition seems to apply only photoresist profile considerations to the evaluation criterion when determining DOF. The wording could be interpreted more loosely, however, to allow for a process latitude specification in the image criterion such as a minimum acceptable exposure latitude.

Interestingly, although the SEMI specification is purposely vague on what image criterion to use it is explicit in the need to evaluate DOF over the entire image field of the exposure tool. I believe this to be a mistake since it confuses the difference between a process response and a process requirement. Theoretically, the DOF is a process response -- how does the process respond to changes in focus? The variation of focus over the field due to the field curvature and astigmatism of an imaging system is not a process response, it is a process error. Measuring DOF is always difficult because the experiment itself induces some focus errors (wafer non-flatness, autofocus uncertainties, etc.) which influence the results. Careful consideration of these experimental errors should always be a part of any DOF measurement technique. But to purposely add focus errors by measuring the photoresist response over the image field is to purposely *not* measure the depth of focus. What is measured is something like a focus budget remaining (the DOF minus the systematic errors of field curvature and astigmatism). This is a useful quantity, especially when evaluating exposure tools, and is sometimes given a distinctive name such as practical DOF or usable DOF. The best course, however, is to measure and specify the DOF correctly, then separately specify a tool's field curvature and astigmatism as part of an overall specification of built-in focus errors [3].

The SEMI specification correctly espouses the need for a complete description of the experimental conditions for the DOF measurement. Depth of focus is not just an optical phenomenon, and it is not just a photoresist phenomenon: it is a lithography process phenomenon. All process conditions should be given when quoting a DOF number, as well as the specific DOF criterion used.

D. Extensions of the DOF definition

It is possible to extend the definition of DOF given previously to include various process conditions or situations which may occur simultaneously. For example, DOF is a strong function of feature size and feature type. If a given mask layer includes several feature sizes and feature types which must be imaged simultaneously, one could evaluate the process window for each feature and find the overlap process window (see Figure 8). Evaluation of this overlapped window would then lead to an overall depth of focus number.

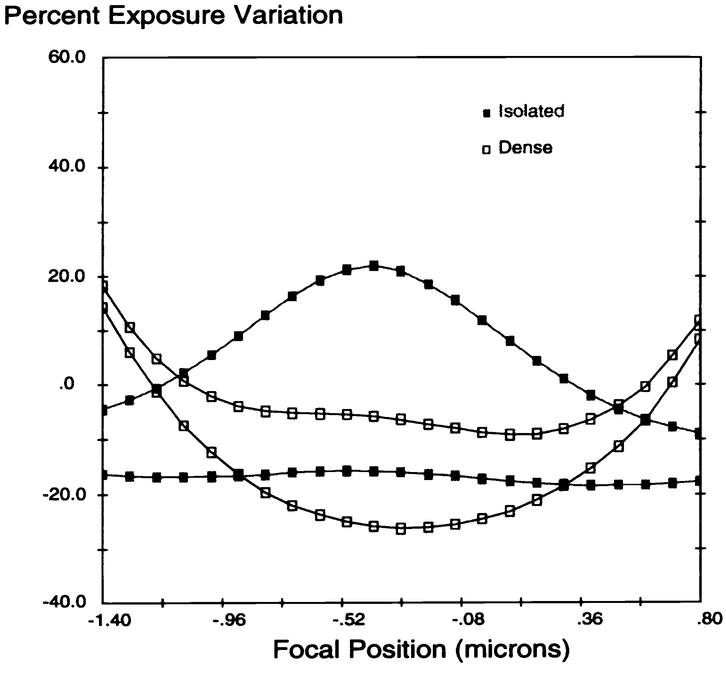


Figure 8. One example of overlapping CD process windows for the simultaneous imaging of dense and isolated lines. Obviously, the print bias between dense and isolated lines reduces the overlap in their process windows.

Using the Definition of DOF

Now that a suitable definition of depth of focus has been found and a systematic procedure for obtaining consistent results determined, DOF can be used as a metric to judge various processes. A few examples of how this metric can be applied will be given here using the lithography simulator PROLITH/2 v4.0 as the source of the data to be analyzed. In addition, PROLITH/2 has automated the data analysis described above using the rectangle method to evaluate DOF.

It is now well known that for any given situation there is one numerical aperture which maximizes the depth of focus. Figure 9 shows how this optimum varies with wavelength for a given feature size and equivalent resists. Interestingly, a relatively low 0.35NA deep-UV system is quite effective in printing 0.5 μm dense features, as one major U.S. DRAM manufacturer has already discovered.

Figure 10 shows how DOF falls off with feature size for a given lithography process and several different types of features.

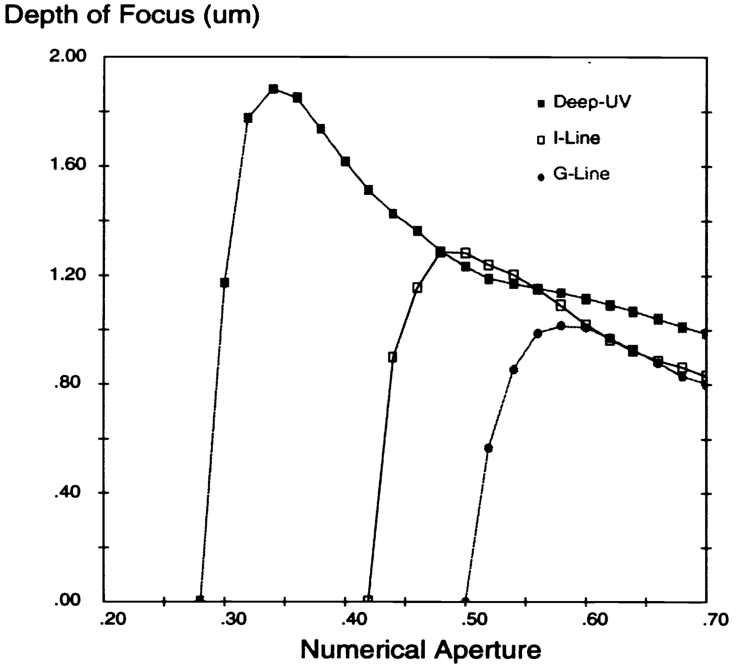


Figure 9. Depth of focus calculated for 0.5 μm lines and spaces as a function of numerical aperture for three different wavelengths. Equivalent resist parameters were used for all wavelengths and the DOF criterion were ±10% CD, sidewall angle > 80°, resist loss < 10% and a minimum exposure latitude of 20%.

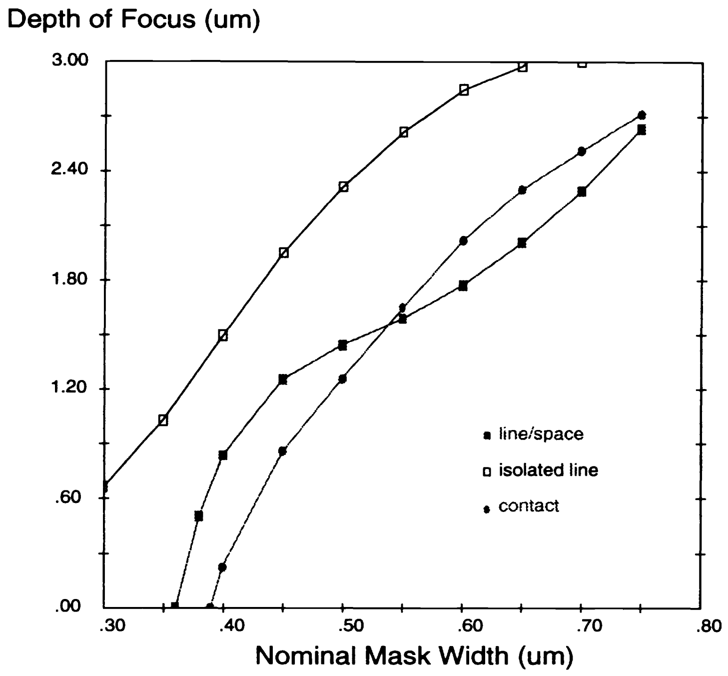


Figure 10. DOF is a strong function of feature size. Here, the fall-off of the depth of focus with feature size is shown for three different features using a simulated 0.55 NA i-line stepper with a typical photoresist.

Conclusions

A definition for depth of focus has been given which adequately reflects the manufacturability of a process in the presence of focus errors. The DOF is defined as the range of focus which keeps the photoresist profile within spec (CD, sidewall angle, and resist loss) over a specified exposure range. The value of the DOF can be obtained by a straightforward analysis of the data from a standard focus-exposure matrix. Once obtained, the measured DOF can be compared to a statistical analysis of the built-in focus errors of a process to determine its manufacturability.

Other means of "measuring" or approximating the depth of focus using a more limited set of data can lead to fairly significant errors if an actual DOF value is needed. These pseudo-DOF metrics can be used for relative DOF comparisons, however, if properly applied.

Finally, a simulator equipped with the analysis capabilities required to determine the DOF can be used to study many important and fundamental questions about optimum processes and future directions of lithography.

References

1. C. A. Mack, "Understanding Focus Effects in Submicron Optical Lithography," *Optical/Laser Microlithography, Proc.*, SPIE Vol. 922 (1988) pp. 135-148, and *Optical Engineering*, Vol. 27, No. 12 (Dec. 1988) pp. 1093-1100.

2. C. A. Mack and P. M. Kaufman, "Understanding Focus Effects in Submicron Optical Lithography, part 2: Photoresist effects," *Optical/Laser Microlithography II, Proc.*, SPIE Vol. 1088 (1989) pp. 304-323.

3. C. A. Mack, "Understanding Focus Effects in Submicron Optical Lithography, part 3: Methods for Depth-of-Focus Improvement," *Optical/Laser Microlithography V, Proc.*, SPIE Vol. 1674 (1992) pp. 272-284.

4. C. A. Mack, "Understanding Focus Effects in Submicron Optical Lithography: a Review," *Optical Engineering*, Vol. 32, No. 10 (Oct. 1993) pp. 2350-2362.

5. J. W. Bossung, "Projection Printing Characterization," *Developments in Semiconductor Microlithography II, Proc.*, SPIE Vol. 100 (1977) pp. 80-84.

6. C. A. Mack and E. W. Charrier, "Yield Modeling for Photolithography," *OCG Microlithography Seminar, Proc.*, (1994) pp. 171-182.

7. E. W. Charrier and C. A. Mack, "Yield Modeling and Enhancement for Optical Lithography," *Optical/Laser Microlithography VIII, Proc.*, SPIE Vol. 2440 (1995)