Yield Modeling and Enhancement for Optical Lithography

Edward W. Charrier and Chris A. Mack
FINLE Technologies, Inc., P.O. Box 162712, Austin, TX 78716

Abstract

A method is presented for predicting the CD limited yield of a photolithographic process using well established lithography modeling tools. The lithography simulator PROLITH/2 is used to generate several multi-variable process response spaces (for example, final resist critical dimension (CD) versus focus, exposure, resist thickness, etc.). Error distributions are determined for each input variable. By correlating the input error distribution with the process response space, a final CD distribution is generated. Analysis of the output distribution produces a predicted parametric CD yield. By analyzing predicted CD yield for multiple geometries, a yield curve (yield vs. feature size) is determined for an i-line process. By comparing CD yield for different nominal exposure doses, a yield-exposure curve is calculated for an i-line process.

Keywords: yield; microlithography; optical lithography simulation

I. INTRODUCTION

Predicting the yield of a lithographic process is difficult to the point that it is rarely attempted. For example, Cost of Ownership (COO) modeling and other cost methods for making decisions in lithography lack an accurate method for determining the parametric yield of a lithography step. As a result, most COO efforts simply assume a value for yield (e.g., 97%) and never change the value. Further, if yield could be predicted in a quantitative manner, the method for predicting yield could be used to optimize yield as well. Is a process which gives the nominal critical dimension (CD) at the nominal values for all inputs the same process which gives the maximum yield?

It is possible to accurately predict the parametric CD yield of a photolithographic process using well established lithography modeling tools [1]. In this paper we will introduce a four step process for predicting CD yield (Figure 1). First, error distributions are determined for each input variable. In this paper we use four input variables: focus, exposure, resist thickness and the development parameter R_{max}. Second, using the lithography simulator PROLITH/2, a multi-variable process response space is generated (for example, final resist critical dimension (CD) versus focus, exposure, resist thickness, etc.). Third, by correlating the input error distribution with the process response space, a final CD distribution is generated. Fourth, analysis of the output distribution produces a predicted CD-limited yield using some acceptance criterion for CD. This number can be used to help optimize the yield of a given process or can serve as a direct input to a COO modeling effort of the simulated process or equipment.

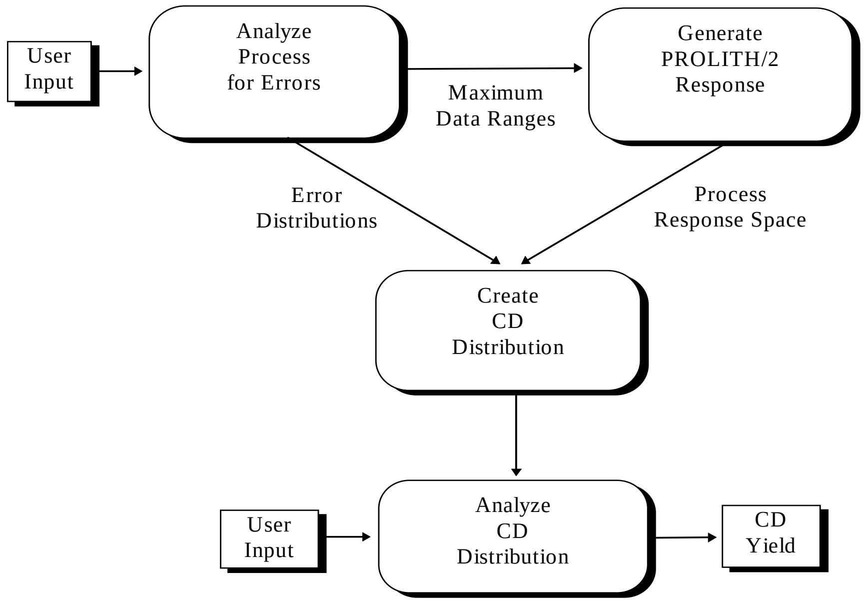


Figure 1. Graphical representation of four-step process for predicting CD yield. User input is required to produce the error distributions from analysis of the process and to analyze the CD distribution to determine acceptable values of CD. Calculation of process space, CD distribution and yield is automated using PROLITH/2 as the calculation engine.

II. THEORY

There is, of course, an extensive literature and experience base on error analysis that can be applied to the prediction of CD limited yield. A very typical approach to predicting an error in critical dimension (ΔCD) resulting from a number of input errors (Δx_i) is the use of the total derivative.

$$\Delta CD = \frac{\partial CD}{\partial x_1}\Delta x_1 + \frac{\partial CD}{\partial x_2}\Delta x_2 + \frac{\partial CD}{\partial x_3}\Delta x_3 + \ldots \tag{1}$$

where each partial derivative represents the process response of CD to the input variable x_i. Equation (1) is exact in the limit of infinitely small input errors. As the errors become larger (i.e., realistic) equation (1) remains accurate so long as the process response remains linear and uncoupled over the range of the error (i.e., the partial derivatives remain constant). If, however, the process response is non-linear or coupled to other variables the use of the total derivative error equation can be both misleading and inaccurate.

If both the input error distribution and the process response are known, the assumption of linearity is not needed -- the resulting CD output distribution can be calculated directly. While both the input errors and the process response can be determined experimentally, process simulation offers the capability to accurately predict the process response, thus significantly reducing the effort required to perform the study.

Consider a simple example to illustrate the method -- the effect of exposure errors on linewidth. The process response in this case is the well known exposure latitude curve. If the input error distribution is known, correlation of the input error probability with the process response function gives the output error distribution. For this example let us assume that the exposure errors are normally distributed about the mean with a 3σ of 10% (Figure 2). The error distribution is plotted as the frequency of occurrence (or probability of occurrence) versus exposure energy with arbitrary units for frequency. The process response is linewidth versus exposure energy and in this case was predicted using the lithography simulation program PROLITH/2. The parameters for the simulation are given in Table I(a). For any given exposure energy, there is a probability that this energy will occur. From the process response curve, an exposure energy corresponds to a specific CD (for example, 0.513 μm for an energy of 200 mJ/cm^2) and thus must have the probability of occurrence corresponding to the probability of the exposure energy. Correlation of the input error distribution with the process response results in a list of linewidth values with corresponding frequencies of occurrence. The linewidth can then be divided up into equal size bins (for Figure 2, the bin size is 0.004 μm) and all of the probabilities with CDs within a given bin are summed. The result is plotted as a histogram of frequency versus CD and represents the resulting output CD error distribution.

The procedure described above and illustrated in Figure 2 is not limited to one input error variable. Multiple simultaneous variations in input parameters can be modeled to produce a full process response space. To produce this response space, the multi-dimensional probability function needs to be determined. For example, if you assume the input errors are independent, then individual input error functions can be combined by multiplying the individual probabilities to create a single multi-dimensional probability function. The multi-dimensional error function and the multi-dimensional process response are then mapped to the standard one-dimensional output distribution.

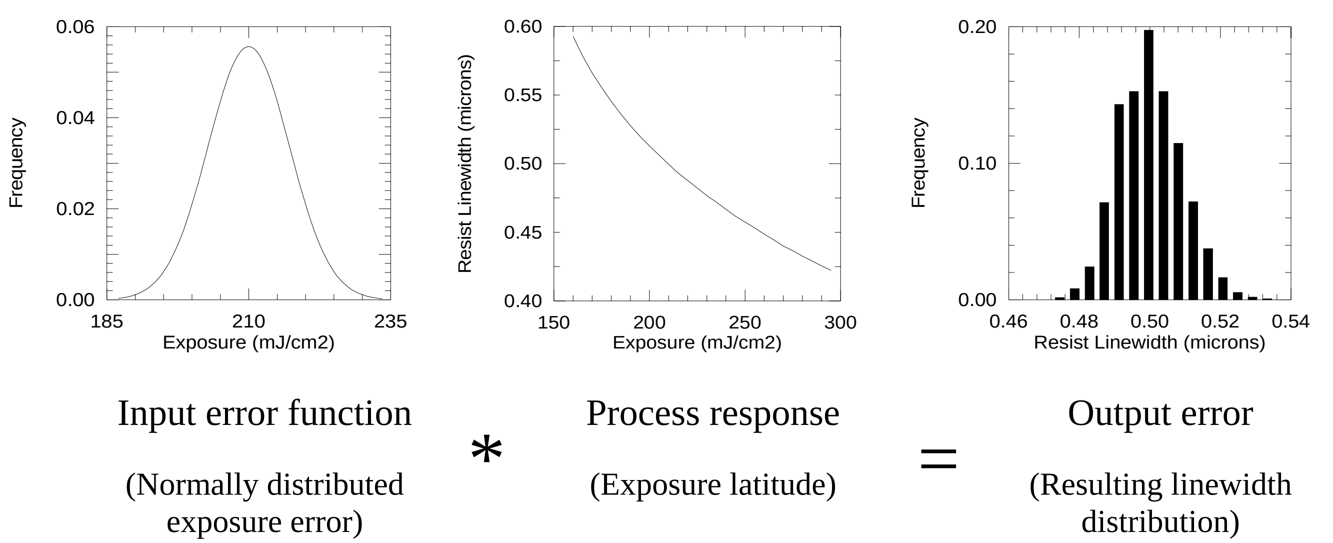


Figure 2. Example of the calculation of an output error distribution from a normal input error function and a simulated process response for the one-dimensional case.

Also, the procedure described above and illustrated in Figure 2 is not limited by the type of input error distribution. In this paper, for convenience, we assume a normal error distribution for each variable, but this assumption is not necessary. Any observed or estimated error distribution can be used. Therefore, this procedure can be used with data which has been compiled from actual processes.

Once a CD distribution has been predicted, the calculation of CD limited yield used here is straightforward. Given some CD specification (for example, the mean ±10%), the frequencies of all CDs within spec are summed and divided by the sum of the frequencies for all CDs to give the yield. There are several subtle assumptions built into this approach, depending on the interpretation of the input error distribution and the impact of an out-of-spec linewidth. In this paper we assume that each CD data point is the mean CD of a die. This assumption is valid if the input error distribution is a die-to-die variation in the variable. In this case, a CD out of spec means that CDs over the entire die are out of spec, resulting in a failed die.

However, if the input error distribution is an error across one die, then the output is the CD distribution across one die. Yield loss would then depend on the specifics of the die. For example, yield loss may result if even one CD is beyond some specification, or it could result if some number of CDs are beyond some other specification, or both. In addition, a specific CD variation could result not in a failure, but in a predetermined bin sort of the die. Given an understanding of how CD errors cause device failures and bin sort, a suitable prediction of yield from a known CD distribution can be accomplished. Of course, such an analysis is device specific.

If both across the die and die-to-die errors must be considered, two separate (though coupled) yields may be necessary. It should also be noted that a feature may "fail" because of other attributes besides CD. For example, the photoresist sidewall angle may have a specification which could result in a failure. The modeling approach presented here could easily be extended to include sidewall angle, resist loss and other metrics of lithographic quality.

III. RESULTS

Several examples of the use of the methodology described in this paper are presented here. First, two examples of using one-dimensional responses are shown for exposure dose and resist thickness inputs. Second, a two dimensional case in which both exposure and resist thickness vary simultaneously is presented. Finally, four inputs are allowed to vary to illustrate the extension to more realistic process conditions. Several examples of the four-variable case are illustrated to show effects of feature type and feature size on yield.

1-D Inputs

Figure 2 illustrates the simplest example of one error input. Although the calculations are straightforward, the results are nonetheless revealing. Notice that the normal (Gaussian) input error distribution does not result in a normal output distribution. Because the process is non-linear, the output distribution is slightly skewed (the probability of getting a +0.02 μm CD error is higher than the probability of getting a -0.02 μm error). Although the exposure response in this case is only slightly non-linear over the ±10% exposure range, it is sufficient to show the effect. Also notice that since the process response in monotonic, the CD with the highest probability of occurrence corresponds to the exposure dose with the highest probability of occurrence.

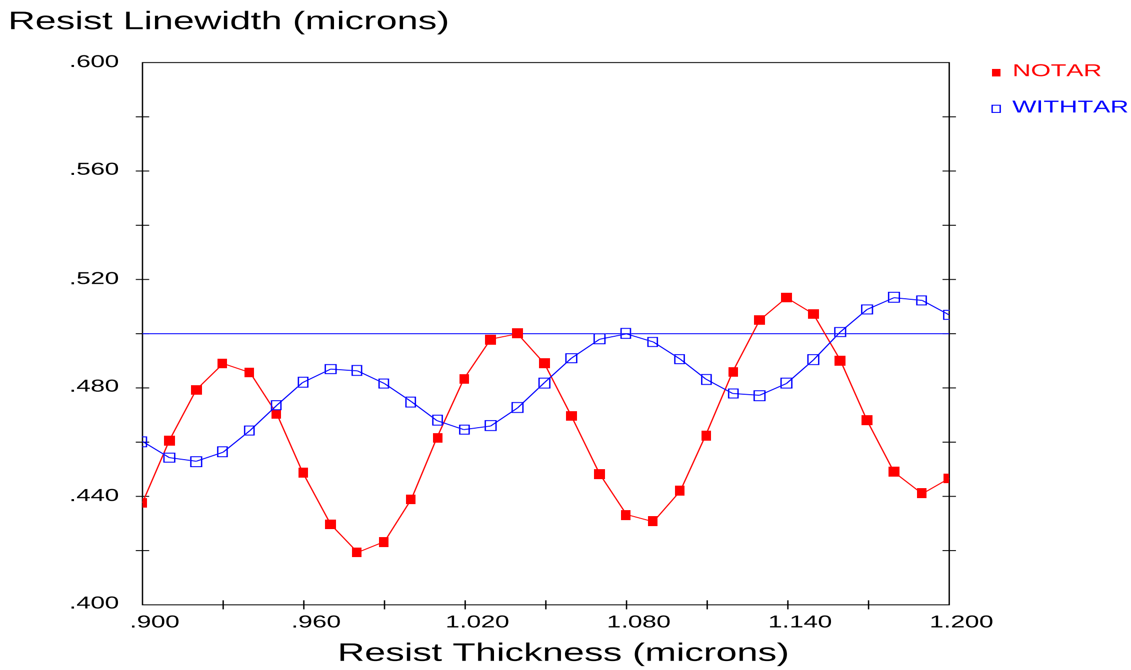


Figure 3. Swing curves from a PROLITH/2 simulation for cases with and without top antireflective coating.

A second more interesting example of a single input error is resist thickness. The response of the linewidth to changes in resist thickness (called a swing curve) is not only highly non-linear, it is not monotonic. Figure 3 shows swing curves simulated with PROLITH/2 for resist exposure on polysilicon both with and without a top layer antireflection coating (TAR). Because of the sinusoidal nature of the response, the resulting CD distribution is very sensitive to both the mean and the 3σ of the input error. Figure 4 shows various output CD distributions for various input errors using the swing curve without TAR as the process response. Figure 4a shows the result where the nominal resist thickness is set at a maximum of the swing curve and the error is assumed to be normal with a 3σ width of 0.06μm, corresponding to half of one swing period. As one would expect, the CD distribution is highly skewed. All resist thickness errors result in smaller than nominal CDs. Operation at a minimum of the swing curve results in a similar, though reversed, skewed distribution (Figure 4b). In both cases, the *mean* of the CD distribution is not the CD at the minimum or maximum of the swing curve.

Figure 4c shows the case where the nominal resist thickness is set halfway between a minimum and a maximum of the swing curve. Even though the input error is normally distributed, the non-monotonic nature of the process response results in a bimodal output distribution. In fact, the CD which corresponds to the most probable resist thickness is one of the least probable CD results. A similar result occurs whenever the resist thickness errors include an appreciable amount of both a

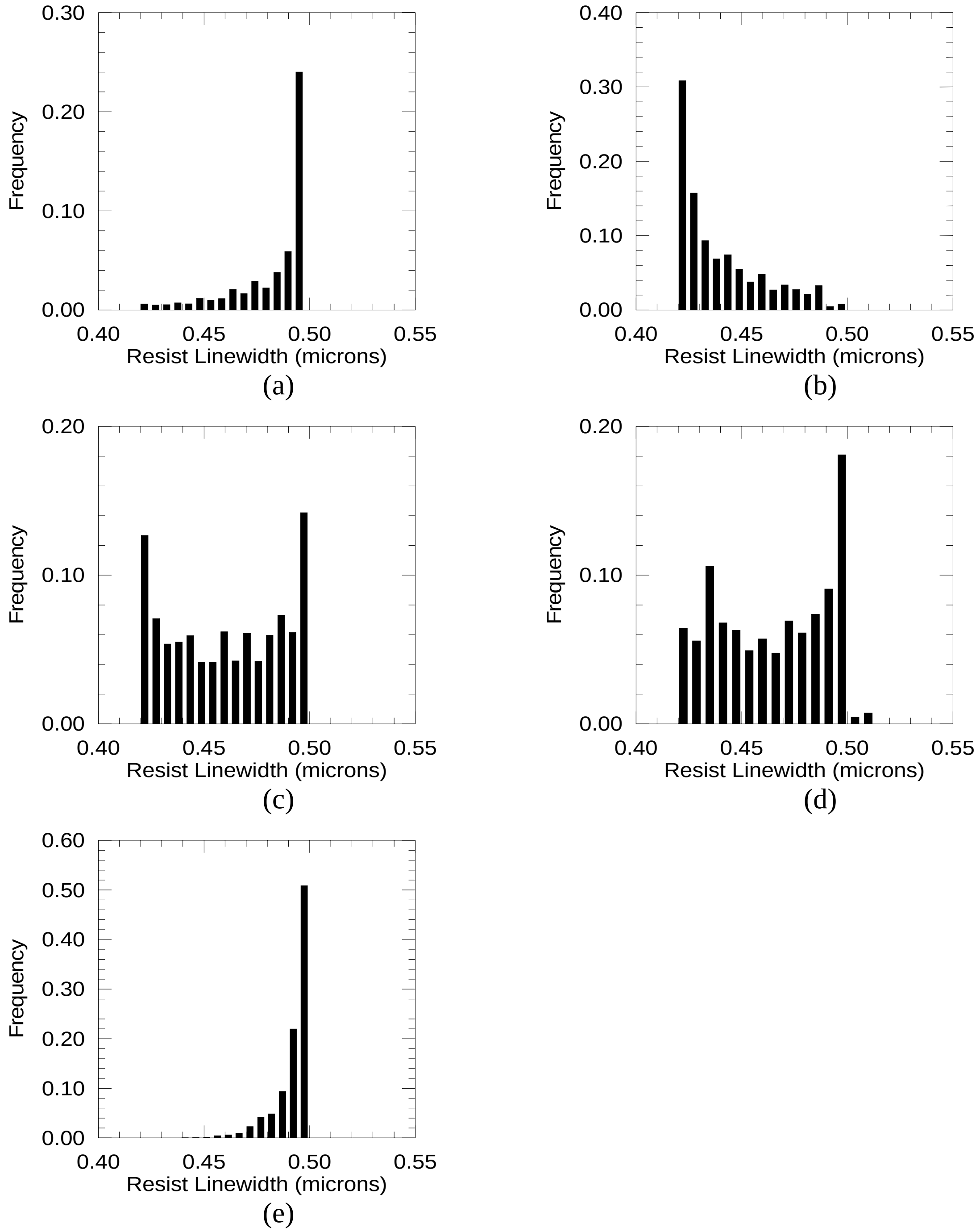


Figure 4. Linewidth distribution for a 0.5 micron process from a swing curve (without TAR) with resist thickness error functions with: (a) Mean at maximum (1.04 μm), 3σ = 1/2 period of swing (0.06 μm); (b) Mean at minimum (0.98 μm), 3σ = 1/2 period of swing (0.06 μm); (c) Mean centered between a maximum and a minimum (1.01 μm), 3σ = 1/2 period of swing (0.06 μm); (d) Mean at maximum (1.04 μm), 3σ = period of one swing (0.12 μm); (e) Mean at maximum (1.04 μm), 3σ = 1/4 period of swing (0.03 μm).

minimum and a maximum of the swing. In Figure 4d, the mean resist thickness is set at a maximum but the 3σ width is one full period. The result again is a bimodal CD distribution, though skewed this time. Finally, in Figure 4e the CD distribution is tightened considerably when the resist thickness distribution is tightened to one-quarter of one swing period.

Figure 5 compares the distributions resulting from swing curves with and without TAR. In both cases the nominal resist thickness was set at a maximum of the swing and the 3σ error was 0.06μm (half of a period). As expected, the use of a top layer antireflection coating, which significantly reduces the swing curve amplitude, will significantly improve the resulting CD distribution.

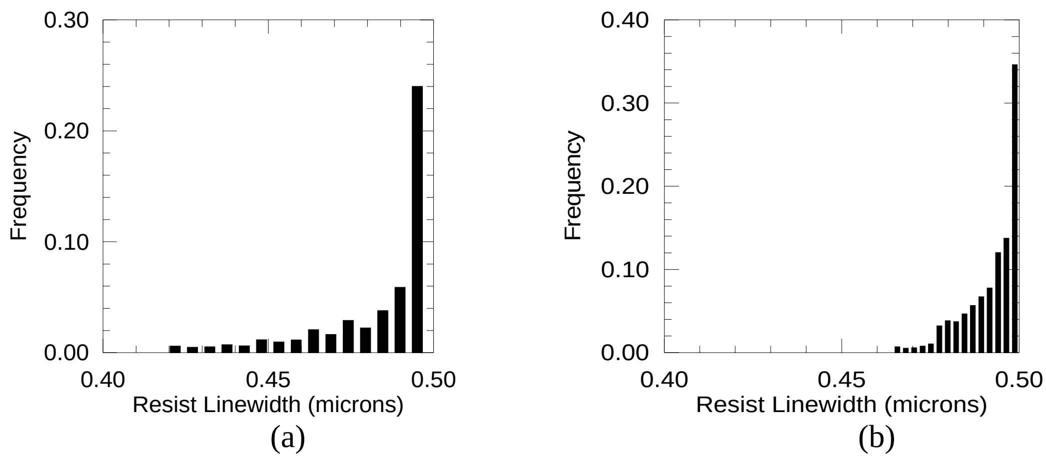


Figure 5. Linewidth distribution with resist thickness error functions with: mean at maximum, 3σ = 1/2 period of swing for (a) swing curve without TAR and (b) swing curve with TAR.

2-D Input

The above examples showed two one-dimensional input variable cases. But what if resist thickness and exposure energy were varying independently, but at the same time? For such a case, the 1-D analysis method can be easily extended to two dimensions. If the two input errors are independent (as is usually the case), their individual 1-D probability functions can be multiplied together to obtain a 2-D probability function. It should be noted that this assumption of independence of the input errors is not a requirement, and is used in this paper only for convenience.

Figure 6 shows such a 2-D input error function assuming both exposure energy and resist thickness errors are normally distributed. The PROLITH/2 lithography simulator was used to map out the entire two dimensional process response space, which is also shown in Figure 6. Note that the two dimensional response is *not* the product of the two one dimensional responses. Careful examination of Figure 6 shows that the swing curve amplitude decreases with increasing exposure dose showing that the two variables are coupled in their impact on CD. The calculation of the CD distribution is carried out in the same manner as in the 1-D case using a mean exposure energy as the nominal dose to size with a 3σ of 10%, and a mean resist thickness set at a maximum of the swing curve with a 3σ of half a period.

Looking at the result (Figure 6), the distribution is skewed due to the swing curve, but is somewhat spread out due to the exposure errors. One extremely important result is that the mean of the CD distribution is not the nominal CD, even though the nominal CD would occur for the mean of the exposure and resist thickness input distributions. The reason for this is the non-linear nature of the process response.

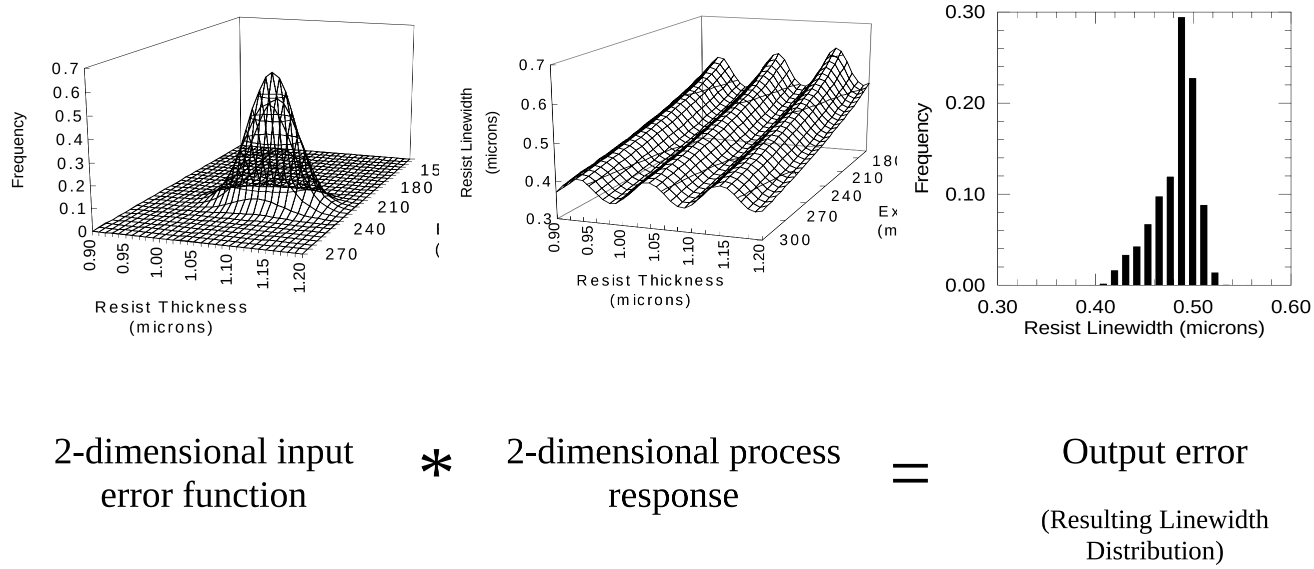


Figure 6. Example of the calculation of an output error distribution from a normal input error function and a simulated process response for a two-dimensional case.

4-D Input

Finally, to show the application of this methodology to more complicated (and more realistic) situations, four input variables were used: focus, exposure energy, resist thickness, and the development parameter R_{max}. The meaning of the first three terms are self-evident and R_{max} is the maximum development rate of the photoresist corresponding to completely exposed positive resist. Changing R_{max} scales the development rate and could be used to account for a variety of process variations which affect photoresist development properties.

First, PROLITH/2 was used to simulate the 4-D process response space centered around the baseline process of Table I Study (a). Using the nominal process as the center of the 4-D error function with 3σ values given in Table II Study (a), the CD distribution was calculated and is shown in Figure 7a. The shape is characteristic of the skew caused by the swing curve, but spread out considerably due to the other variables. The mean of the resulting CD distribution is 0.480 μm with a calculated yield of 87.7% based on a ±10% CD specification. The fact that the mean is less than the nominal linewidth indicates that the nominal process is not “centered” despite the non-statistical evidence to the contrary.

Parameter	Study (a)	Study (b)
Projection System:		
Nominal linewidth	0.50 μm	0.40 μm
Pitch	1.00 μm	0.8 μm
Wavelength	365 nm	365 nm
Numerical aperture	0.54	0.58
Image reduction ratio	5:1	5:1
Partial coherence	0.5	0.48
Focal position	-0.3 μm	-0.3 μm
Nominal exposure dose	210 mJ/cm^2	110 mJ/cm^2
Substrate	Silicon	Silicon
Layer 1	Oxide, 30 nm	Titanium Nitride, 62 nm
Layer 2	Polysilicon, 350 nm	Aluminum, 300 nm
Post-Exposure Bake:		
PEB diffusion length	60 nm	65 nm
Development:		
Time	60 seconds	60 seconds
Maximum development rate	150 nm/s	125 nm/s
Minimum development rate	0.05 nm/s	0.008 nm/s
Threshold PAC concentration	-100	0.10
Selectivity	6.0	5.10
Resist System:		
Thickness	1.04 μm	0.90 μm
Parameter A	0.800 μm^{-1}	0.930 μm^{-1}
Parameter B	0.300 μm^{-1}	0.073 μm^{-1}
Parameter C	0.016 cm^2/mJ	0.019 cm^2/mJ
Index of refraction	1.75	1.70

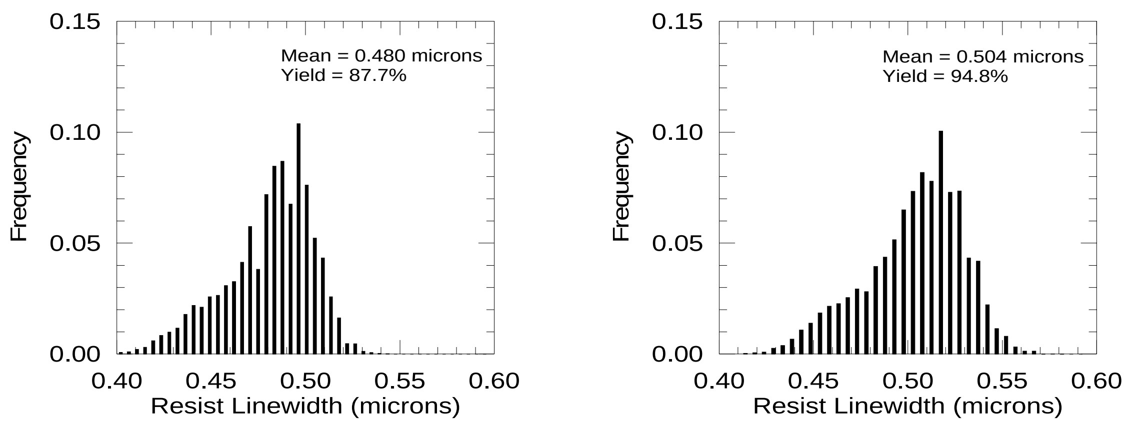
Table I. PROLITH/2 lithography simulation parameters used as a baseline for the simulation of (a) a 0.5 micron process with nominal at a maximum of the swing curve, and (b) a 0.4 micron process with nominal at a minimum of the swing curve.

Since the mean linewidth of the nominal process is undersized by 4%, could shifting the process result in an improvement in yield? Such a question can be easily answered. The most difficult part of the yield calculation is the calculation of the process response space, which is independent of the error input functions. The errors can be easily manipulated and a new yield calculated using the already calculated process response space. To find the process which optimizes the CD limited yield, the mean of each of the four input errors was varied, keeping the 3σ values constant. Yield was calculated for each process condition. The best result, shown in Figure 7b, used the same process as the nominal but with the exposure energy reduced by about 8%. This "under-exposure" shifted the mean linewidth up to the nominal linewidth and resulted in a significant increase in yield (to 94.8%).

Parameter	Mean (a)	3σ (a)	Mean (b)	3σ (b)
Exposure energy	210 mJ/cm^2	21 mJ/cm^2	110 mJ/cm^2	11 mJ/cm^2
Resist thickness	1.04 μm	0.06 μm	0.90 μm	0.05 μm
Focus	-0.3 μm	0.5 μm	-0.3 μm	0.5 μm
Development R_{max}	150 nm/s	30 nm/s	125 nm/s	25 nm/s

Table II: Input error distributions for the nominal process of the four-dimensional test cases of (a) a 0.5 micron process with nominal at a maximum of the swing curve, and (b) a 0.4 micron process with nominal at a minimum of the swing curve.

Note that in the case where the nominal resist thickness is at a maximum of the swing curve, *under-exposure* causes the yield to increase. Alternatively, when working with a nominal at a minimum of the swing curve, an *over-exposure* would have the same effect. To illustrate this point in a different, more interesting way, a study was performed with a process which has a nominal resist thickness at a minimum. Using the process parameters in Table I(b) and the error distributions in Table II(b) for resist thickness, R_{max} and focus, the yield was calculated several times using different values for the nominal exposure energy. For this study, the 3σ of the exposure remained 10% of the nominal for each data point. The results of this study, shown in Figure 8, illustrate that when working at nominal settings at a minimum of the swing curve, an over-exposure can cause an increase in yield. In the case of this 0.4 micron dense lines/spaces with an i-line process, an over-exposure of 4.5% caused an increase in yield of 5.2%.



(a) (b)

Figure 7. Linewidth distribution for four input variable case: mean resist thickness at a maximum, mean focus at nominal, mean developer at nominal and (a) mean exposure at nominal (210 mJ/cm^2) and (b) mean exposure at 8.1% below nominal (193 mJ/cm^2) with yield based on CD specification of nominal ±10%.

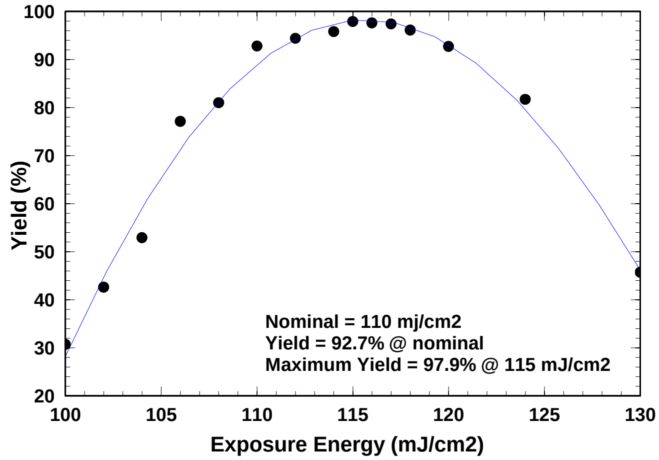


Figure 8. A "Dose-Yield Optimization Curve" showing yield as a function of exposure energy for 0.4 micron dense lines/spaces with an i-line process with nominal exposure of 110 mJ/cm^2.

In order to determine the effects of feature size on yield, a study was performed using the parameters listed in Table I(b) and Table II(b). For a nominal 0.4 micron i-line process, the yield was determined for dense lines/spaces at various linewidths. The resulting data is shown in Figure 9. As expected, the yield of a 0.4 micron process drops rapidly when linewidths fall below 0.4 microns, while the yield increases less precipitously when linewidths increase. It should be noted that these yield values were calculated using a nominal exposure of 110 mJ/cm^2, and that, as shown above, it may be possible to enhance these yield values by changing the nominal exposure.

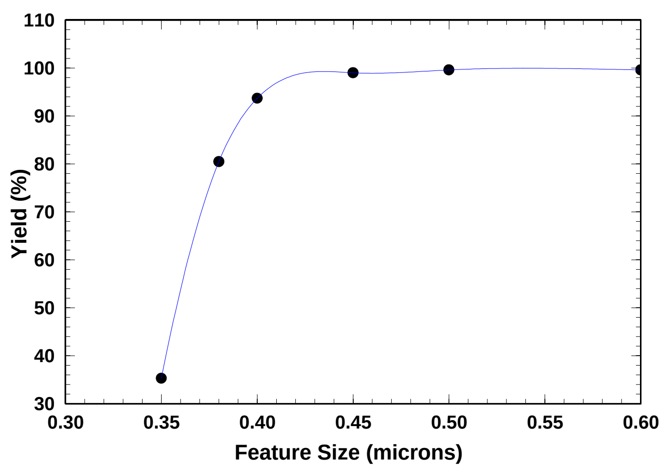


Figure 9. Yield as a function of feature size for a 0.4 micron i-line process of dense lines/spaces of various widths.

In addition to dense lines/spaces, CD distributions and yields were determined for isolated lines and contacts for the 0.4 micron i-line process whose parameters are described in Table 1(b) and Table II(b). To better understand the effects of different features on yield, it is helpful to compare the CD distributions of the three feature types. Figure 10 shows the distributions of dense lines, isolated lines and contacts for this 0.4 micron process. As described earlier in this paper, a tight CD distribution indicates that the process is less sensitive to process errors than a process whose distribution is more widely spread. By inspecting the graphs in Figure 10, it is apparent that the isolated lines have the most compact distribution, indicating that, as expected, isolated lines are the least sensitive to process errors. The CD distribution of the contacts is the most widely spread, indicating these features are the most sensitive to process errors. This information seems to confirm our knowledge that isolated lines are the easiest to print and contacts are the most difficult. The calculated yield further confirms this conclusion. Based on a $\pm 10\%$ CD specification, the yield was 99.0% for the isolated lines, 93.7% for the dense lines, and 81.1% for the contacts.

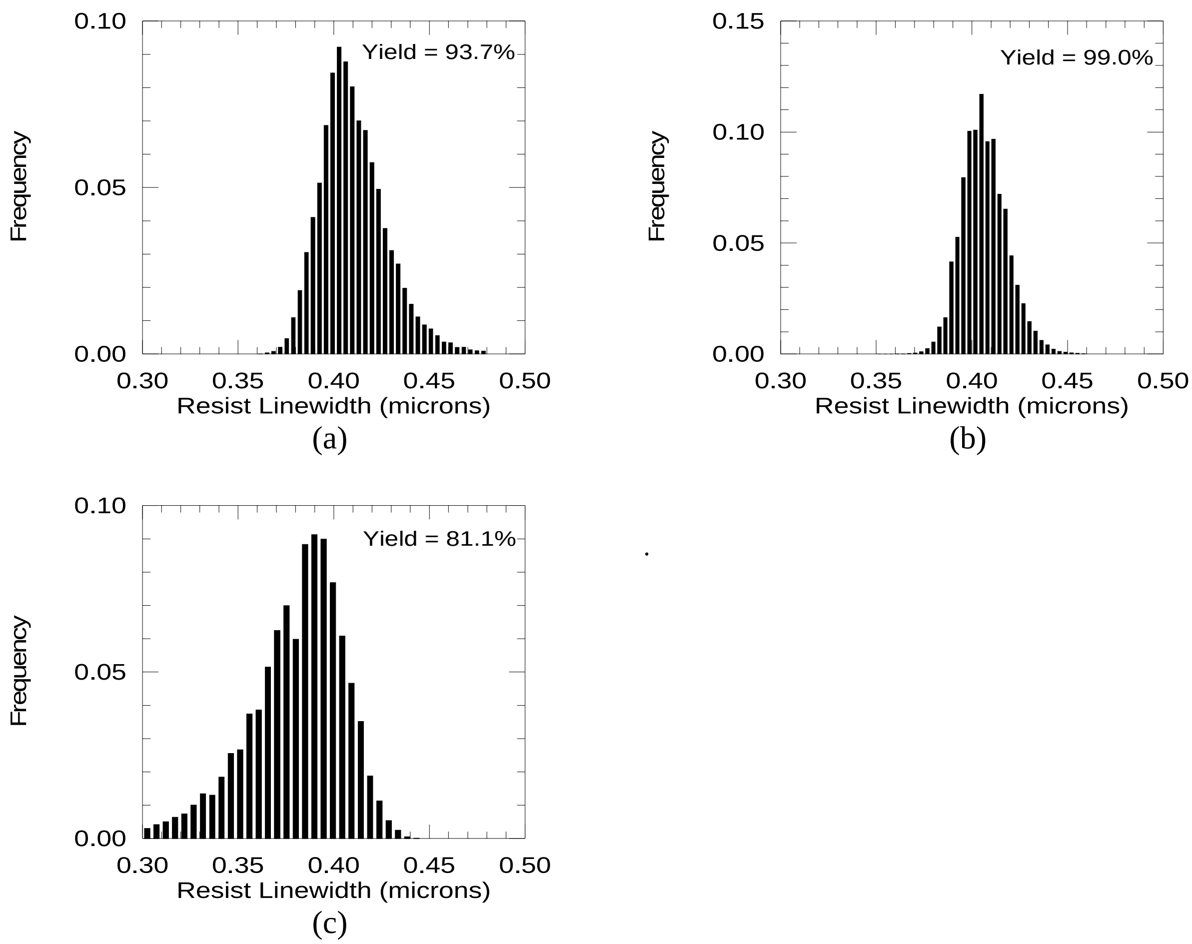


Figure 10. CD distributions of (a) dense lines/spaces, (b) isolated lines, and (c) contacts for a 0.4 micron i-line process with mean resist thickness at a minimum of the swing curve.

IV. CONCLUSIONS

Several important conclusions can be drawn from this work. First, lithography simulation tools can be used to conveniently calculate large, multi-dimensional process response spaces. Second, a knowledge of the error distribution for each input variable can be combined with the process response to give a predicted CD distribution of the output. This distribution can be further analyzed to give a single yield number to characterize the quality of the process. With such a method for predicting linewidth distributions and CD limited yield, many applications are possible. Different processes can be compared using yield as the metric, both for use in Cost of Ownership models and for process optimization studies. Further, by using this methodology it is possible to find a "yield maximum" for a given process. This yield maximum may not coincide with the nominal parameter settings.

There are many opportunities for future work. First, in this paper, only normal input errors were used. This limitation is not necessary and was only imposed for convenience. Any arbitrary input error distribution can be used, including distributions determined by historical data. Second, other process parameters can be investigated either alone or in combination with the parameters used in this paper. Third, other outputs can be used. For example, yield could be defined based on a combination of CD, resist sidewall angle, and resist thinning. Fourth, an investigation of the effects of different types of errors (i.e., within die vs. die-to-die) on the yield can be performed.

REFERENCES

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