Defining and Measuring Development Rates for a Stochastic Resist

Chris A. Mack
Lithoguru.com, 1605 Watchhill Rd, Austin, TX 78703

Abstract

Development rate can be defined microscopically (the development rate at a point) or macroscopically (the propagation rate of an average resist height). In the presence of stochastic noise, these two rates will be different. Using a stochastic resist simulator, the propagation rate of a resist surface is calculated in the presence of stochastic variation in the resist deprotection concentration using a nonlinear development rate model. For both 2D and 3D simulations, the development front propagation rate was fit to semi-empirical expressions. The resulting propagation rate can be more than an order of magnitude higher than for the case of no stochastic noise. The differences between microscopic and macroscopic dissolution rate can have an important effect on how development rate models should be calibrated, depending on their use in continuum or stochastic lithography simulators.

Keywords: Development rate, stochastic resist, lithography simulation, line-edge roughness, LER

1. Introduction

Photoresist development rates are commonly measured to characterize photoresist dissolution behavior. Such data are frequently used to calibrate development models for simulators. An important though frequently unstated assumption of this use is the equivalence of microscopic and macroscopic development rates. Microscopic development rate is the development rate at a point in the resist and is the rate used by simulators (Fig. 1a). Macroscopic development rate is the mean rate at which a large open area of resist develops down and is the quantity measured when development rates are measured (Fig. 1b), for example, when using a dissolution rate monitor. In the absence of stochastic effects that result in surface roughness, these two rates are identical.

In the real, stochastic case, the mean propagation rate of a large open area is a strong function of the stochastic uncertainty of the development rate, especially in regions of moderately low dissolution rates. This paper will explore the impact of stochastic uncertainty in microscopic development rate on the macroscopic development rate through the use of both analytical derivations and Monte Carlo simulations.

2. Theory

Dissolution rate uncertainty will inevitably result from uncertainty in the underlying inhibitor concentration (for example, the concentration of protecting groups in a chemically amplified resist). Consider a simple development rate function[1]

$$r = r_{max}\frac{(a+1)(1-m)^n}{a+(1-m)^n} + r_{min}, \quad a = \frac{(n+1)}{(n-1)}(1-m_{th})^n \qquad (1)$$

where r is the development rate, m is the relative protecting group (inhibitor) concentration, and r_{max}, r_{min}, n, and m_{th} are model parameters. Here, we will neglect r_{min} as small compared to the development rate in the

region of interest. The edge of a photoresist feature will necessarily have a protection level that is near the knee of the development rate curve, so that $m > m_{th}$. Thus, if $n >> 1$, the development rate in this region will be well approximated by

$$r \approx r'_{max}(1-m)^n, \quad r'_{max} = r_{max}\frac{a+1}{a} \tag{2}$$

This development rate expression will be used throughput this paper. While simple, it accurately reflects the non-linear development rate response for the exposure and deprotection levels expected near a photoresist feature edge.

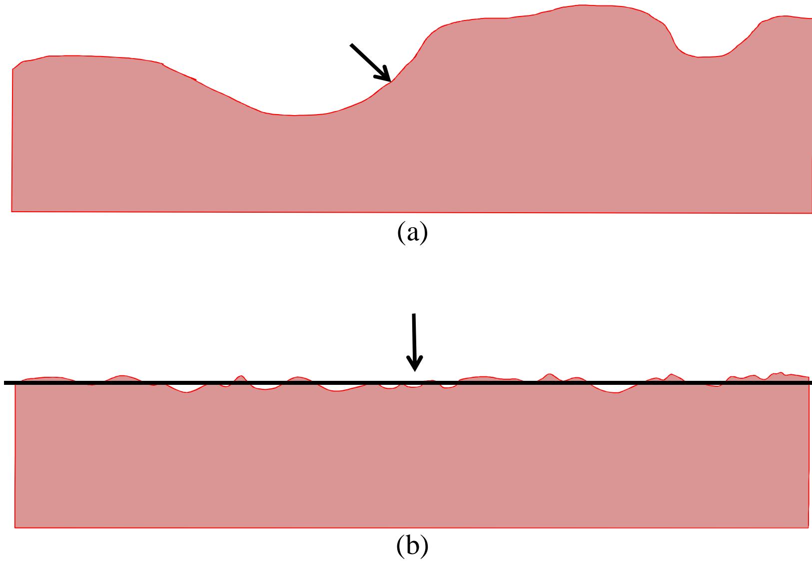


Figure 1. Photoresist development rate is defined in two ways: a) microscopic development – the rate at which a point on the resist surface moves perpendicular to that surface; and b) macroscopic development – the rate at which the mean surface position (the mean development front) propagates.

Suppose that m is a random variable with a normal distribution, so that $m \sim N(\mu, \sigma_m)$. If the randomness of m is the only source of uncertainty in the resulting development rate, a probability distribution function (pdf) for r can be calculated.

$$pdf_r = \left|\frac{dm}{dr}\right| pdf_m, \quad \left|\frac{dm}{dr}\right| = \frac{1}{nr}\left(\frac{r}{r'_{\max}}\right)^{\frac{1}{n}} \tag{3}$$

$$pdf_r = \frac{1}{\sqrt{2\pi}\sigma_m}\frac{1}{nr}\left(\frac{r}{r'_{\max}}\right)^{\frac{1}{n}} e^{-\left[1-\mu-(r/r'_{\max})^{1/n}\right]^2/2\sigma_m^2} \tag{4}$$

Unfortunately, the moments of this distribution (in particular, the mean and the variance) cannot be analytically calculated, making the utility of this pdf expression questionable. The mode, however, does have an analytical form:

$$r_{mode} = r'_{\max}\left(1 - m_{mode}\right)^n, \qquad 1 - m_{mode} = \frac{1-\mu}{2}\left(1 + \sqrt{1 - \frac{4(n-1)\sigma_m^2}{(1-\mu)^2}}\right) \tag{5}$$

This expression gives real roots so long as $\sigma_m < \dfrac{1-\mu}{2\sqrt{n-1}}$. For small σ_m (that is, for $\sigma_m << 1 - \mu$, an approximation that will be made often in this work),

$$1 - m_{mode} \approx \left(1-\mu\right)\left(1 - (n-1)\left(\frac{\sigma_m}{1-\mu}\right)^2\right) \tag{6}$$

so that

$$r_{mode} \approx r\left(\mu\right)\left(1 - (n-1)\left(\frac{\sigma_m}{1-\mu}\right)^2\right)^n \tag{7}$$

Again for the case of small σ_m, the actual pdf of equation (4) is well approximated by a Generalized Gamma Distribution (GGD):

$$pdf_{GGD} = \frac{\beta}{\lambda\Gamma(\alpha)}\left(\frac{r}{\lambda}\right)^{\alpha\beta-1} e^{-\left(r/\lambda\right)^\beta} \tag{8}$$

which has the following properties:

$$\text{mode} = \lambda\left(\alpha - \frac{1}{\beta}\right)^{\frac{1}{\beta}}, \qquad \left\langle r\right\rangle = \lambda\frac{\Gamma\left(\alpha + \frac{1}{\beta}\right)}{\Gamma(\alpha)}, \qquad \left\langle r^2\right\rangle = \lambda^2\frac{\Gamma\left(\alpha + \frac{2}{\beta}\right)}{\Gamma(\alpha)} \tag{9}$$

Matching the mode of the GGD to the mode or r, we find that

$$\lambda = r_{mode}\left(\alpha - \frac{1}{\beta}\right)^{-\frac{1}{\beta}} \tag{10}$$

The GGD can then be conveniently calculated by defining

$$\gamma \equiv \left(\frac{r}{r_{mode}}\right)^\beta\left(1 - \frac{1}{\alpha\beta}\right) \tag{11}$$

so that

$$pdf_{GGD}(r) = \left(\frac{\beta}{r}\right)\left(\frac{\alpha^{\alpha} e^{-\alpha}}{\Gamma(\alpha)}\right) e^{\alpha(1-\gamma+\ln\gamma)} \tag{12}$$

As we shall see, the assumption of small σ_m is equivalent to saying $\alpha >> 1$, so that Sterling's approximation to the Gamma function can be used:

$$\Gamma(\alpha) \approx \alpha^{\alpha} e^{-\alpha} \sqrt{\frac{2\pi}{\alpha}} \left(1 + \frac{1}{12\alpha} + \frac{1}{288\alpha^2} + ...\right) \tag{13}$$

giving

$$pdf_{GGD}(r) \approx \left(\frac{1}{r}\right)\left(\frac{\beta\sqrt{\alpha}}{\sqrt{2\pi}}\right) e^{\alpha(1-\gamma+\ln\gamma)} \tag{14}$$

The values of α and β can now be determined by empirically fitting the GGD to numerical evaluations of equation (4). Excellent fits are obtained when

$$\alpha = \left(\frac{1-\mu}{3\sigma_m}\right)^2, \qquad \beta = \frac{3}{n} \tag{15}$$

This leads to an approximate pdf for r of

$$pdf(r) \approx \frac{1}{\sqrt{2\pi}}\left(\frac{1}{r}\right)\left(\frac{1-\mu}{n\sigma_m}\right) e^{\alpha(1-\gamma+\ln\gamma)} \tag{16}$$

where

$$\gamma = \left(\frac{1-m}{1-m_{mode}}\right)^3 \left(1 - 3n\left(\frac{\sigma_m}{1-\mu}\right)^2\right) \approx \left(\frac{1-m}{1-\mu}\right)^3 \left(1 - 3\left(\frac{\sigma_m}{1-\mu}\right)^2\right) \tag{17}$$

Using the small σ_m approximation for γ [the right-hand side of equation (17)] requires a small adjustment to the scaling for the pdf, giving

$$pdf(r) \approx \frac{1}{2.54}\left(\frac{1}{r}\right)\left(\frac{1-\mu}{n\sigma_m}\right) e^{\alpha(1-\gamma+\ln\gamma)} \tag{18}$$

Note that γ, and the exponential term in the pdf, is not dependent on the development rate function parameters. A plot of equation (18) compared to equation (4) is shown in Fig. 2 for the case of $n = 10$.

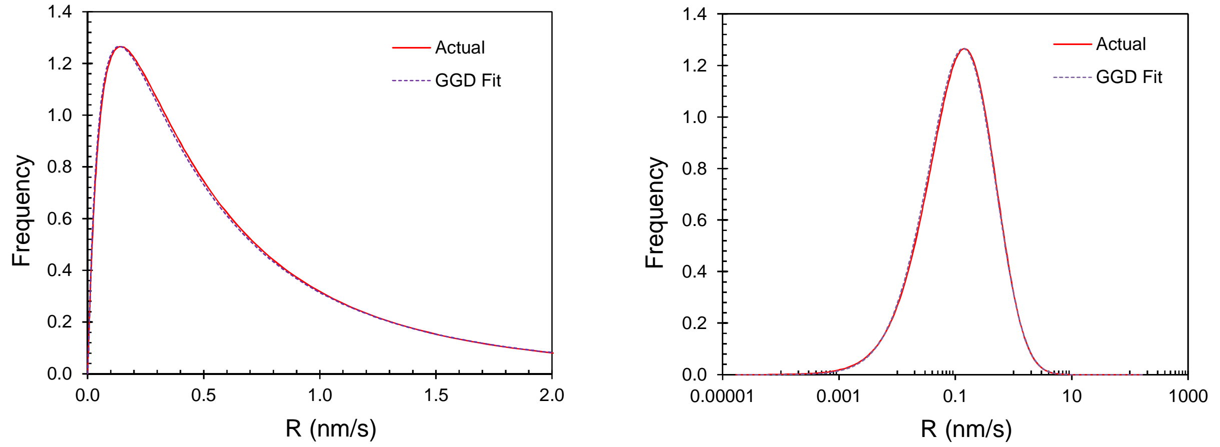


Figure 2. Comparison of the numerical calculation for the pdf_r (labeled as "actual") compared to the approximate GGD form from equation (18). The two graphs differ only in the use of linear and log scales for the x-axis.

The advantage of using the Generalized Gamma Distribution is the ability to calculate the mean and variance of the development rate analytically. Applying Sterling's approximation to equations (9),

$$1+\left(\frac{\sigma_r}{\langle r\rangle}\right)^2=\frac{\Gamma(\alpha)\Gamma\left(\alpha+\frac{2}{\beta}\right)}{\left[\Gamma\left(\alpha+\frac{1}{\beta}\right)\right]^2}\approx\frac{\left(1+\frac{2}{\alpha\beta}\right)^{\alpha+\frac{2}{\beta}-\frac{1}{2}}}{\left(1+\frac{1}{\alpha\beta}\right)^{2\alpha+\frac{2}{\beta}-1}} \tag{19}$$

For $\alpha\beta >> 1$,

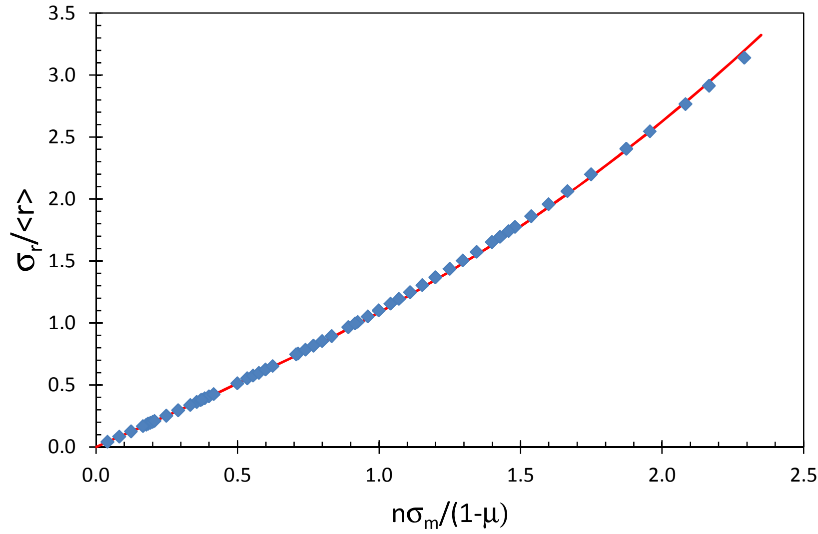
$$\left(\frac{\sigma_r}{\langle r\rangle}\right)^2\approx\left(\frac{n\sigma_m}{1-\mu}\right)^2+O\left(\frac{n\sigma_m}{1-\mu}\right)^4 \tag{20}$$

The accuracy of this approximation will now be tested with simulations, along with simulations of the development front propagation rate (i.e., the macroscopic development rate).

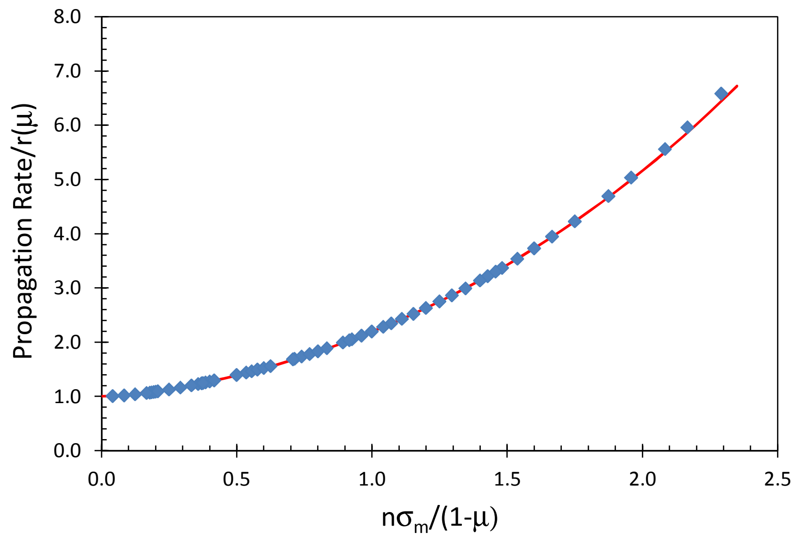
3. Monte Carlo Simulations

Simulation was used to predict the resist height as a function of development time for an open-frame exposure/development in the presence of stochastic dissolution-rate noise.[2,3] The development model discussed above in equation (2) was used, with $m \sim N(\mu, \sigma_m)$ selected from a random number generator for each grid point (all grid points remaining uncorrelated). The grid size was set to 1 nm. For 2D (1+1) simulations, the simulation width was 2048 grid elements and the resist thickness was 4096. For 3D simulations (2+1), the widths in x and y were 512 grids and the resist thickness was 1024. The parameters were r_{max} = 200 nm/s, m_{th} = 0.5, and n was varied between 5 and 15. The development time was adjusted in each case so that approximately 1000 time steps would allow the front to reach the bottom of the resist.

For each combination of μ and σ_m evaluated, the mean and standard deviation of the microscopic development rate were calculated, and the development front propagation rate was determined by fitting the average resist surface height versus time with a straight line. An example of the 2D simulation results, for $n = 10$, is shown in Fig. 3. An example of 3D simulation results is shown in Fig. 4.



(a)



(b)

Figure 3. 2D simulation results for $n = 10$. Each data point is the average of 500 trials. μ was varied between 0.72 and 0.76 and σ_m was varied between 0.005 and 0.055.

The simulation results were fit to semi-empirical expressions (guided by the results of the theory section above). Equation (20) predicts a linear variation of the relative development rate uncertainty with

relative deprotection uncertainty, but for higher amounts of uncertainty a higher-order term was required. For 2D and 3D simulations and for all values of n,

$$\left(\frac{\sigma_r}{\langle r\rangle}\right)^2 \approx \left(\frac{n\sigma_m}{1-\mu}\right)^2 + (0.039n - 0.21)\left(\frac{n\sigma_m}{1-\mu}\right)^4 \tag{21}$$

The front propagation rate, r_{prop}, was also fit to empirical equations.

$$\text{2D:} \quad r_{prop} \approx r(\mu)\left(1+\left(\frac{n\sigma_m}{1-\mu}\right)^{1.5} + 0.016n\left(\frac{n\sigma_m}{1-\mu}\right)^3\right) \tag{22}$$

$$\text{3D:} \quad r_{prop} \approx r(\mu)\left(1+2.2\left(\frac{n\sigma_m}{1-\mu}\right)^2 + (0.04n-0.188)\left(\frac{n\sigma_m}{1-\mu}\right)^4 + (0.0016n-0.008)\left(\frac{n\sigma_m}{1-\mu}\right)^6\right) \tag{23}$$

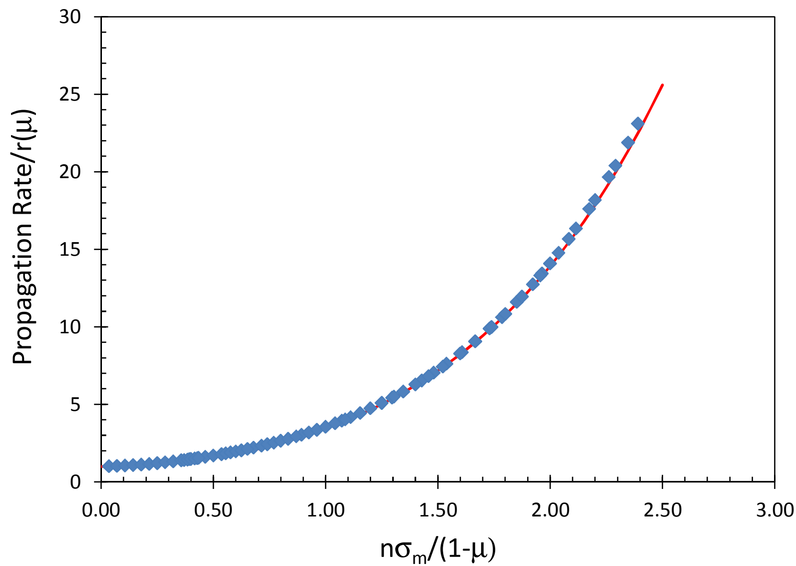


Figure 4. 3D simulation results for n = 10. Each data point is the average of between 4 and 50 trials. μ was varied between 0.72 and 0.77 and σ_m was varied between 0.001 and 0.055. The resulting $r(\mu)$ went from 0.08 to 1.92 nm/s, and the propagation rate varied from 0.1 to 6.7 nm/s.

4. Discussion and Conclusions

Lithography simulators, whether making the continuum approximation or performing stochastically, require a model relating the dissolution rate at a point in the resist to the level of deprotection at that point. The data to calibrate such a model universally comes from measurements of film thickness versus development time for large open-frame exposures. As the simulations performed in this work show, the

average development front propagation rate, as might be measured by a dissolution rate monitor, can be many times larger than the development rate at the mean deprotection concentration when stochastic variations in deprotection rate are present. Figure 4, for example, shows propagation rates more than 20 times higher than the expected rate. A summary of this effect is shown in Figure 5.

The reason for this unexpected behavior is the dramatically skewed probability distribution for development rate that arises from a normally distributed deprotection level when the dissolution rate dependence is highly nonlinear (that is, for a high value of n). This observation gives rise to two concerns. When using dissolution rate data to calibrate a continuum model, differences in the variance of the deprotection level will affect the data and thus the model fit. If the continuum simulator is to be used for cases where the variance in deprotection level matches that of the dissolution rate measurements, all is well. But if not, there will be some concern as to the faithfulness of the simulation results. When using dissolution rate data to calibrate a stochastic model, the calibration process should include stochastic effects as well. Otherwise, a continuum model calibration process could greatly overestimate the dissolution rate values to be used in a stochastic simulator near the knee of the development curve.

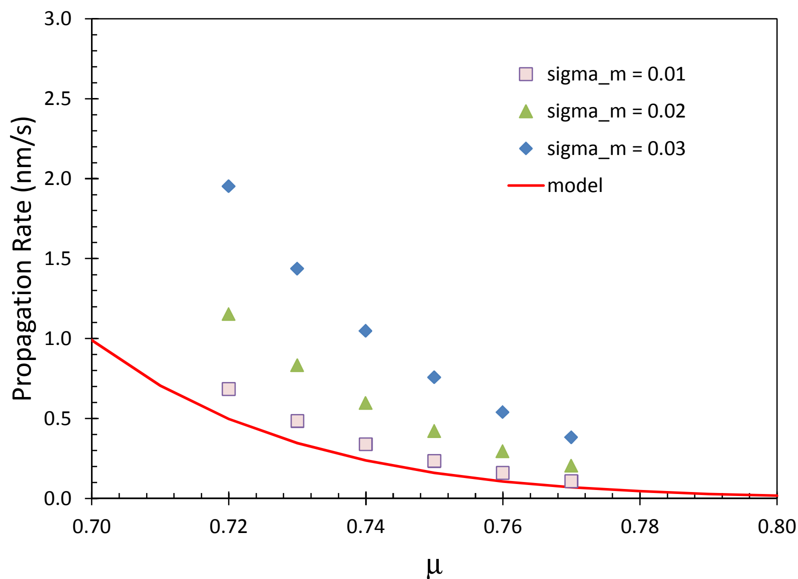


Figure 5. Comparison of 3D simulation results of the development front propagation rate (for $n = 10$) for different σ_m. The "model" curve corresponds to $\sigma_m = 0$.

[1] C. A. Mack, Fundamental Principles of Optical Lithography, John Wiley & Sons (2007), p. 260.
[2] C. A. Mack, "Stochastic Modeling in Lithography: The Use of Dynamical Scaling in Photoresist Development", *Journal of Micro/Nanolithography, MEMS, and MOEMS*, Vol. 8, No. 3 (Jul-Sep 2009) p. 033001.
[3] C. A. Mack, "Stochastic modeling of photoresist development in two and three dimensions", *Journal of Micro/Nanolithography, MEMS, and MOEMS* , Vol. 9, No. 4 (Oct-Dec, 2010) p. 041202.