Yield Modeling and Analysis

Prof. Robert C. Leachman
IEOR 130, Methods of Manufacturing Improvement
Spring, 2014

1. Introduction

Yield losses from wafer fabrication take two forms: *line yield* and *die yield*. Line yield losses result from physical damage of the wafers due to mishandling, or by mis-processing of the wafer (e.g., skipping or duplicating a process step, wrong recipe, equipment out of control, etc.). Mis-processing is detected either by in-line inspections interspersed through the wafer fabrication process or by an electrical parametric test of a special test pattern on the wafer. This parametric test is almost always performed just before the wafer leaves the fabrication facility to go to the wafer probe area. It is also sometimes performed at one or more points within the wafer fabrication process flow.

Many die yield losses are the result of tiny *defects.* Defects are defined as any physical anomaly that causes a circuit to fail. This includes shorts or resistive paths or opens caused by particles, excess metal that bridges across steep underlying contours causing shorts, photoresist splatters and flakes, weak spots in insulators, pinholes, opens due to step coverage problems, scratches, etc.

It is natural to think of defects as being randomly distributed across the wafer surface, and to speak about the *density* of defects on the wafer surface, i.e., the number of circuit faults per unit area. If we postulate that a die will not work unless it is completely free of defects, then the probability that a die works is the probability that no defects lie within its area. Obviously, the larger the die area, the more the chance it includes one or more defects, and so the less the probability that the die works. Thus wafers with large die printed on them will have a lower die yield than will wafers with small die printed on them, if the two types of wafers are made in the same fabrication process and are subject to the same density of defects.

To fairly compare die yields of products with different die areas made in different factories, it is desirable to find the underlying defect density in each factory. A factory with a lower defect density is capable of producing with a higher die yield.

Not all die yield losses are due to defects. Some mis-processing escapes detection at in-line optical inspections in the fabrication process as well as at parametric test. And some types of mis-processing affect only a portion of the dice printed on the wafer. A prevalent example is *edge loss*. The thickness of films deposited on the wafer is often well-controlled across the central portion of the wafer but poorly controlled near the edge of the wafer, resulting in wholesale die yield losses near the edge. Parametric test and in-line inspections typically are performed on a sample basis and exclude edge die. Hence edge losses show up as die yield loss, even though they are not the result of defects.

For the moment, we will assume all die yield losses are the result of defects in order to develop the theory of defect density models. We will relax this assumption subsequently.

2. The Poisson Model

Suppose the mean number of defects per die is λ_0. According to the Poisson probability distribution function, the probability that a die has k defects is given by

$$P(k) = \frac{e^{-\lambda_0} \lambda_0^{\ k}}{k!} , \quad \text{for } k = 0, 1, 2, \cdots \tag{1}$$

The probability the die works is $P(0)$; the expected die yield is therefore

$$DY = P(0) = e^{-\lambda_0} \ .$$

If the mean defect density is D_0 defects per square centimeter, and the die area is A sq cm, then we should take $\lambda_0 = D_0A$. We therefore write

$$DY = e^{-D_0 A} \ . \tag{2}$$

Equation (2) is called the Poisson die yield model. Given an observed die yield DY, we can infer that the underlying defect density in the fab is

$$D_0 = -\frac{\ln DY}{A} \, . \tag{3}$$

A very useful feature of the Poisson model is the additivity of defects. If the overall defect density D_0 is decomposable into defectivity contributions at different steps or different mask layers, e.g.,

$$D_0 = D_1 + D_2 + D_3 + \ldots + D_n \ ,$$

then the yield loss contribution of each step or layer is easily identified, as the overall die yield has a product form:

$$DY = e^{-AD_0} = e^{-A\sum_{i=1}^{n} D_i} = \prod_{i=1}^{n} e^{-AD_i} \ .$$

Using this product form, one can calculate the yield improvement to be gained from reductions in defect density achieved at various steps or layers. For example, if the defect density in layer j is reduced from D_j to $D_j - \Delta D_j$, then the new die yield is

$$DY^{NEW} = e^{A\Delta D_j} \prod_{i=1}^{n} e^{-AD_i} = e^{A\Delta D_j} DY.$$

Empirically, the Poisson yield model has been found to give accurate yield predictions for small die (when $A \leq 0.25$ sq cm) and when the expected number of defects per die is low (when $D_0A < 1.0$). In the case of large die areas, it tends to underestimate die yield, for reasons that will be explained later. Nonetheless, in almost any situation, it is accurate for estimating small *changes* in die yield as a function of small changes in step-level or layer defect densities.

3. The Binomial Model

Suppose the entire wafer has n total defects on it. Let p be the probability that a random defect lands on a given die. Assume the defects are independent from each other. According to the binomial distribution, the probability that k out of the n defects land on the particular die in question is

$$P(k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \ . \tag{4}$$

In particular, the probability the die works is

$$P(0) = (1-p)^n \ . \tag{5}$$

Suppose the area of the whole wafer is A_w, and suppose the area of the die is A. If the defect density is D_0, then the expected total number of defects on the wafer is $n = D_0 A_w$, while the expected number of defects on the die is D_0A. The probability a particular defect is located within a given die is just the ratio, i.e.,

$$p = \frac{D_0 A}{D_0 A_W} \ ,$$

or $p = A / A_w$. Substituting into (5), the expected die yield is

$$DY = P(0) = \left(1 - \frac{A}{A_w}\right)^{D_0 A_w} \ . \tag{6}$$

Typically, the area of the wafer A_w is much larger than the area of the die A. Moreover,

$$\lim_{A_w \to \infty} \left(1 - \frac{A}{A_w}\right)^{D_0 A_w} = e^{-D_0 A} \ . \tag{7}$$

For A_w an order of magnitude larger than A, (6) closely approximates (2). Thus the Binomial model gives essentially the same numerical answers for die yield as does the Poisson model. Since the Poisson model is mathematically more tractable, it is used in preference.

4. Mixed Distribution Models

Actual data on defects shows that defect and particle densities vary widely from chip to chip, from wafer to wafer, and even from lot to lot. In fact, the defects frequently tend to cluster together. Because of this, the Poisson model tends to underestimate die yield when the expected number of defects per chip is greater than one or when the die area is relatively large. (When the defects cluster together in some die, then other die can be relatively defect-free, thereby increasing the yield compared to the case when defects are more spread out.)

One approach for dealing with this problem is to posit that the defect density D itself varies according to a probability distribution $f(D)$. This was first done by B. T. Murphy of Bell Labs. The expected die yield in this case is expressed as

$$DY = \int_0^\infty e^{-DA} f(D) dD \,. \tag{8}$$

By definition, the distribution $f(D)$ has mean D_0, but beyond that, we don't have much of an idea as to what it should look like. If one assumes D is distributed uniformly between 0 and $2D_0$, (8) simplifies to

$$DY = \frac{1 - e^{-2AD_0}}{2AD_0} \,.$$

If one assumes D is distributed according to a symmetrical triangular distribution extending from 0 to $2D_0$ with peak at D_0, it can be shown that (8) simplifies to

$$DY = \left[\frac{1 - e^{-AD_0}}{AD_0} \right]^2 . \tag{9}$$

Equation (9) is commonly referred to as the *Murphy model* for die yield. Given a die yield DY and die size A, one can numerically solve for the implicit mean defect density D_0 that satisfies (9).

If one assumes D is distributed according to an exponential distribution, i.e.,

$$f(D) = \frac{1}{D_0} e^{-\frac{D}{D_0}} \,,$$

it can be shown that (8) simplifies to

$$DY = \frac{1}{1 + AD_0}\,, \tag{10}$$

which is known as the *Seeds model* for die yield. A variant of the Seeds model, known as the *Bose-Einstein model* for die yield, is a product form

$$DY = \left(\frac{1}{1 + AD_0}\right)^n\,, \tag{11}$$

where n is the number of *critical mask layers*. The idea behind the Bose-Einstein model is that most fatal defects are deposited in certain difficult ("critical") mask layers. For example, metal layers are especially prone to the generation of fatal defects. We would expect that a device fabricated in a process technology with a given number of critical layers (say, four metal layers) will have a lower die yield than a device with the same area fabricated in another technology with fewer critical layers (say, two metal layers). The Bose-Einstein model can be developed assuming die yield in each critical layer is expressed using the Seeds model, and overall die yield is the product of defect-limited yields in all the critical layers.

Finally, if $f(D)$ is assumed to be a Gamma distribution, it has been shown that (8) reduces to a Negative Binomial model, i.e.,

$$DY = \left(1 + \frac{AD_0}{\alpha}\right)^{-\alpha}\,, \tag{12}$$

where α is called the *cluster parameter*. If defect data is available, this parameter can be estimated from the defect data as

$$\alpha = \frac{\bar{\lambda}^2}{\left(\sigma^2 - \bar{\lambda}\right)}\,. \tag{12}$$

Here, $\bar{\lambda}$ is the mean number of defects per die and σ is the standard deviation of the number of defects per die.

By suitably choosing the extra parameter α, the Negative Binomial model (11) can closely approximate any of the other models. For $\alpha \geq 10$, the Negative Binomial model is essentially the same as the Poisson model (2). For $\alpha = 5$, the Negative Binomial closely approximates the Murphy model (9). For $\alpha = 1$, the Negative Binomial closely approximates the Seeds model (10).

A drawback to using the Negative Binomial model for determining defect density is that given only a die yield DY and a die area A, it is not clear what value of α to use in order to determine the underlying mean defect density D_0. If α were somehow given, the mean defect density can be easily computed as

$$D_0 = \frac{\alpha}{A}\left(DY^{-\frac{1}{\alpha}} - 1 \right). \tag{13}$$

5. Practical Defect Density Models

For small die sizes $A \leq 0.25$ sq cm, or for low defect densities $AD_0 < 1$, the simple Poisson model (2) is widely used and is accurate. Moreover, if one is only concerned about the *change* in die yield given a change in defect density at one or several process steps, an analysis using the Poisson model is sufficiently accurate.

For large die sizes, the Negative Binomial model (11) is the most flexible and potentially most accurate model. However, the extra parameter α needs to be determined by statistical methods or by estimation from actual defect data. Where such data are not available, the Murphy model (9) is frequently used.

6. Models Incorporating Both Random and Systematic Yield Losses

Now suppose that in addition to random defects there are losses independent of die size that we shall term *systematic yield losses.* We decompose overall die yield as $DY = Y_S Y_R$, where Y_R is the *defect-limited yield* and Y_S is the *systematic-limited yield*. If we posit a simple Poisson Model for defects, we have

$$DY = Y_S Y_R = Y_S e^{-AD_0}. \tag{14}$$

Using linear regression, we can determine a best fit of the two parameters Y_S and D_0 to actual data on DY vs. A if we take logs of both sides of the above equation:

$$\ln DY = \ln Y_S - AD_0 . \tag{15}$$

Here, $\ln Y_S$ is the constant and D_0 is the coefficient on the independent variable A. We can determine these unknown parameters from *wafer map* data for a single product using what is known as the *windowing* technique. A wafer map presents the yield by die position on the wafer. A *stacked wafer map*, showing the average yield by die position, is depicted in Figure 1. For the purposes of this section, assume we have a wafer map for a single wafer, i.e., a map showing which dice worked and which dice failed on a given wafer.

The windowing technique is explained as follows. The average die yield for the product vs. the die area of the product constitutes one data point. Now suppose we group the dice printed across the wafer surface into pairs, and pretend that the pair is a single die with area $2A$. This paired single-die only works if both component dice work. From review of

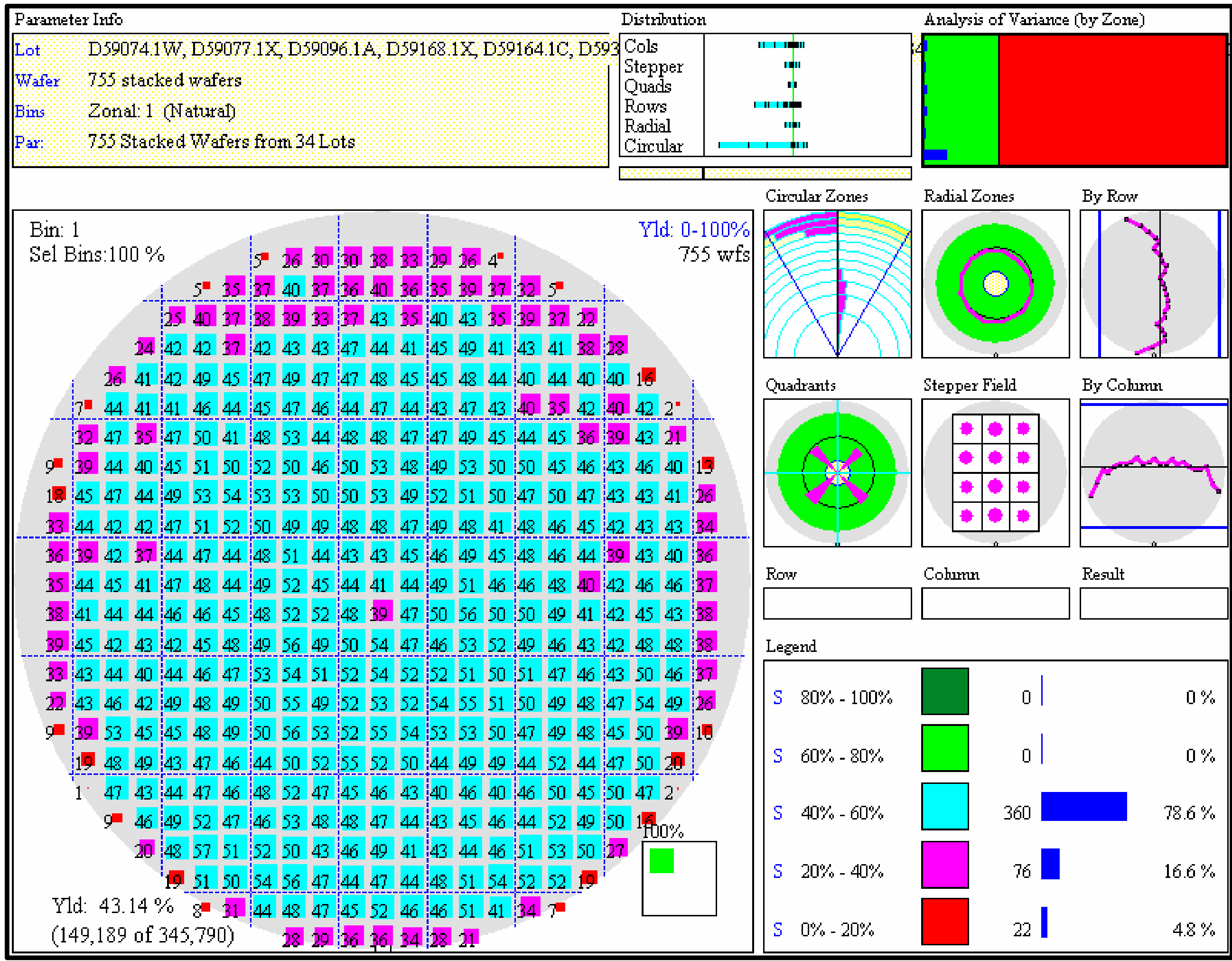

Figure 1. Sample Wafer Map

the wafer map, the die yield of this paired single-die can be identified. This provides a second data point. The procedure can be repeated for die groups of size 3, 4, 5, 6, 7, 8, etc., providing more data points for the regression.

Just as for the simple Poisson model for defect density, any of the compound defect density models could be appended with a systematic-limited yield coefficient Y_S. In practice, a two-parameter model such as the simple Poisson with a systematic-limited yield coefficient is typically sufficient for practical purposes.

We remark that the windowing method simply sorts out yield losses into those that are independent of die area versus those that are dependent on die area. This is *not* equivalent to a decomposition of yield losses by point-defect mechanisms vs. other mechanisms. For example, edge losses will be larger for wafers with larger-sized dice printed on them than for wafers with small dice. Thus losses from some of the non-defect mechanisms will be included in the D_0 parameter rather than in the Y_S parameter.

7. Models with Baseline Random Yield Loss and Systematic Yield Loss

A different and useful decomposition of die yield stems from an SPC-type viewpoint. Suppose we posit that the truly random die yield losses all must come from a stable, stationary system of chance causes that we classify as *baseline defects.* There may be occasional *excursions* (significant additional yield losses from this baseline) when the process or equipment drifts out of control, either misprocessing or depositing excessive particles. Losses from such excursions, as well as chronic losses that are not randomly distributed, are termed *systematic yield losses*. Systematic yield losses have an observable signature. It could be a signature over time (e.g., one or several lots with exceptionally poor yields), or a spatial signature (e.g., certain die positions on the wafer or certain wafer positions within the lot with much lower than average yields). Edge losses are a good example of a systematic yield loss with a spatial signature.

Suppose we abstractly collect all systematic losses into a single term $1-Y_S$ and all random losses into a single term $1-Y_R$. (Y_S is known as the *systematic limited yield*, Y_R is the *baseline defect-limited yield*.) The overall die yield is $DY = Y_R Y_S$. Improvement of baseline yield requires fundamental improvement in the cleanliness of the process and equipment. Improvement of systematic yield requires improved process execution and/or improved process monitoring and control (to detect excursions from baseline losses and react to contain losses from such excursions). Typically, faster progress in yield improvement can be made on the systematic side, whereby signatures can be analyzed to help determine root causes and engineering projects to mitigate specific systematic yield losses can be formulated and carried out. It is therefore helpful to know the $Y_R – Y_S$ breakdown of overall die yield, as well as the breakdown of $1-Y_S$ into its many component losses.

Some insight for the decomposition of yield into Y_R and Y_S components may be gained by viewing a *wafer yield histogram* in addition to the *wafer yield map*. An example wafer yield histogram is presented in Figure 2. From a large sample of wafers, the number of

good die per wafer vs. the number of wafers achieving that yield is plotted. If yield losses were solely due to the stable system of chance causes, then by the Central Limit Theorem, the histogram should present a normal distribution. But it has a long left tail, indicating there are significant excursion losses. The overall histogram reflects a juxtaposition of the normal distribution for the baseline losses plus the excursion and systematic losses.

If yield losses were solely due to stationary random baseline defects, on the wafer map we should see a Poisson distribution of yield losses, which, for a large number of die per wafer such as here, should look like a normal distribution. There should be no spatial correlation of yield across the wafer. But that is not what we see. Note the poor edge yield and the poor yield in the dead center of the wafer; those are clearly systematic problems.

Suppose we looked at a wafer map for a wafer that to the best of our knowledge was not involved in any excursions. Suppose we focus on the best-observed-yielding die site in that map, probably located near the center of this wafer map. We will certainly ignore the die sites near the edge that exhibit edge losses, and we will ignore the poor-yielding die site in the center of the map in Figure 1, as well as any other die sites exhibiting a spatial signature. In Figure 1, the best-yielding die site exhibits a die yield of 54%, and there is only one die site achieving this yield.

For baseline random defects, die yield is well-characterized by a Binomial distribution. (Recall that a Poisson model and a Binomial model are equivalent when the number of die per wafer is sufficiently large.) For Binomial die yield, the distribution exhibited by the wafer histogram should be a normal distribution, if the wafer sample is sufficiently large. For such a distribution, a span of 6σ should contain (almost) all observations and the peak should be centered 3σ from the maximum die yield.

However, the overall yield distribution as seen in Figure 2 is a juxtaposition of the baseline random defect-limited yield and the systematic mechanisms-limited yield. We might expect that excursions add a long left-hand tail, while chronic systematic losses (such as edge losses) shift the whole distribution to the left.

We can argue that the only way the best-observed yield is achieved is when no systematic losses are present and we are witnessing a point that is at the right-hand edge of the (unseen) normal distribution for the baseline random yield. We therefore could expect the distance between the mean of the baseline random yield distribution and the maximum die yield identified on a stacked wafer map to be approximately 3σ of the distribution that results from the baseline random defect-limited yield, as long as the number of die per wafer is sufficiently large. We can use this observation to estimate Y_R as follows.

For a Binomial model with mean Y_R, the standard deviation of the average yield is given by

$$\sigma = \sqrt{Y_R(1-Y_R)/m}\ , \tag{16}$$

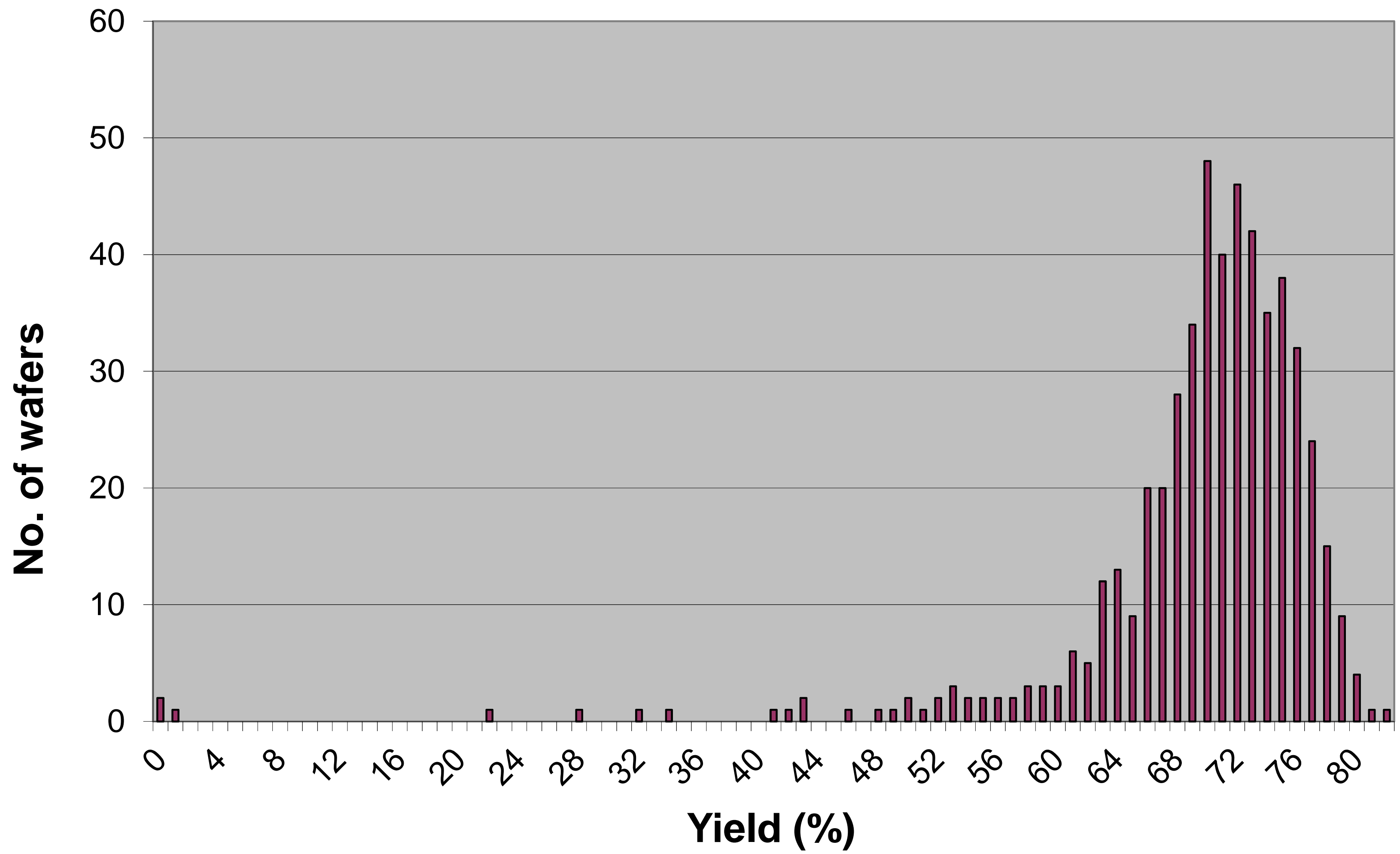

Figure 2. Example of Yield Histogram

where m is the total number of wafers in the stack. If we let MY denote the maximum die yield that is observed. The difference between MY and the unknown Y_R depends on the number of die sites that were considered and at how many die sites MY was observed. For example, suppose there were 300 die sites considered (die sites subject to edge loss or other chronic systematic loss mechanisms are ignored), and suppose MY was observed at 2 die sites. Using a normal approximation, that would suggest MY occurred at $\Phi^{-1}(1-2/300)$ standard deviations above Y_R where Φ denotes the cumulative unit normal density function. More generally, if we apply the normal approximation so that the distance from Y_R to MY is $\Phi^{-1}(1 - l/n)\sigma$, where l is the number of times MY appears on the stacked wafer map and n is the number of die sites considered on the stacked wafer map, then we estimate MY occurs at $Y_R + k\sigma$ where $k = \Phi^{-1}(1 - l/n)\sigma$. In particular, if MY were observed at 2 out of 88 die sites, that would suggest that MY is two standard deviations above Y_R, and if MY were observed at 1 out of 714 die sites, that would suggest that MY is three standard deviations above Y_R.

Using this approximation, we have

$$MY - Y_R = k\sigma = \Phi^{-1}(1 - l/n)\sqrt{Y_R(1 - Y_R)/m} \quad , \tag{17}$$

which may be solved using the quadratic formula to find Y_R. Once Y_R is determined, we can divide it into DY to determine Y_S.

This binomial-sigma method will result in a smaller systematic mechanism limited yield than the Y_S term computed using the windowing method. The virtue of this approach to match maximum-observed-yield is that edge losses and defect excursions are excluded from the determination of Y_R (to the extent that they do not contribute to the die sites exhibiting maximum yield). That is, defect losses are sorted out into baseline losses present across all die sites on every wafer from losses with a spatial or temporal signature.

As an example, the wafer map in Figure 1 provided the average yields by die site over a large group of wafers, in this case, 755 wafers. As for candidate die sites, we ignore the top three rows of dice, seemingly subject to some systematic mechanism. Starting in the fourth row and ignoring die subject to edge loss, we have one row of 15, one row of 17, two rows of 19, then 10 rows of 21, 2 rows of 19, one row of 17, one row of 15, one row of 13, and one row of 9, ignoring the bottom row subject to edge loss. This makes for a total of 372 candidate die sites for observing the maximum die yield. The maximum observed die yield is 57%, occurring at only one site out of the 372 candidate die sites. Then $k = \Phi^{-1}(1 - 1/372) = 2.78$. Solving (17), we obtain the estimate $Y_R = 51.8\%$. The average die yield is 43.1%, implying $Y_S = 83.2\%$.

8. Findings of the SMLY Survey

During 2002-2003 the author and Dr. C. Neil Berglund undertook a study sponsored by International Sematech of systematic yield loss mechanisms. Yield learning trends, prevalent yield loss mechanisms, yield analysis and yield improvement practices were

surveyed at fabrication facilities operated by five Sematech member companies, including one in Asia, two in Europe and two in the USA.

Yield Learning Trends

The participants in the SMLY survey provided die yields for selected high- and moderate-volume devices from the 350nm, 250nm and 180nm technology nodes. Die yields were broken down into line yield through die sort (i.e., wafers that survive the die sort process to count towards die yield statistics) and the die yield of sorted wafers. Figures 3, 4 and 5 display yield loss vs. calendar time for selected devices at the 350nm, 250nm, and 180nm technology nodes. Most of these are logic devices, although SRAMs and flash memory devices also are represented (yield loss after repair in these cases). Die sizes varied from slightly larger than 0.2 cm^2 to more than 2 cm^2.

To compensate for differences in die size at the participants, we needed to normalize the random yield losses that depend on the die area. As will be explained below, we found over the entire process life for the devices produced at all participants that yield losses due to systematic categories were comparable to random defect losses. We therefore fit the following Poisson model to the data provided by the participants to predict their die yield for a normalized die size:

$$Y = Ys * Yr = [\ (1 + Y)/2\ ] * Yr = [\ (1 + Y)/2\ ] * \exp(-AD_0)$$

where Y is the reported die yield (excluding sort line yield), A is the reported die area, and D_0 is the derived Poisson defect density. As may be seen, this model assumes exactly half of total die yield losses stem from systematic mechanisms that are independent of die area, while the other half are to be explained by a fatal defect density D_0.

Using the value of D_0 derived as above, we predicted the ***yield loss*** for a 0.5 cm^2 device in each process technology reported by the participants as

$$YL = 1 - SLY * [\ (1 + Y)/2\ ] * \exp(-0.5*D_0)$$

where Y is the reported die yield, SLY is the reported sort line yield and D_0 is the defect density calculated as above.

As may be seen in the figures, yield loss learning rates are quite comparable across the participants. The slopes of the curves for the various participants at the 350nm node are quite parallel, and the curves of the various participants line up remarkably well at the 250nm node, almost as if all of the data came from a single fab. At the 180nm node, data is more scant, and more disparity seems to be present in the learning rates of the participants. (Although the pace of yield learning may seem to be less at the 180nm node, this is only because the time scale is stretched out in Figure 5 compared to that in Figures 3 and 4.) One may also observe from these figures that the participants seem to be unable to drive yield loss for a 0.5 cm^2 device down to 10 percentage points, even at

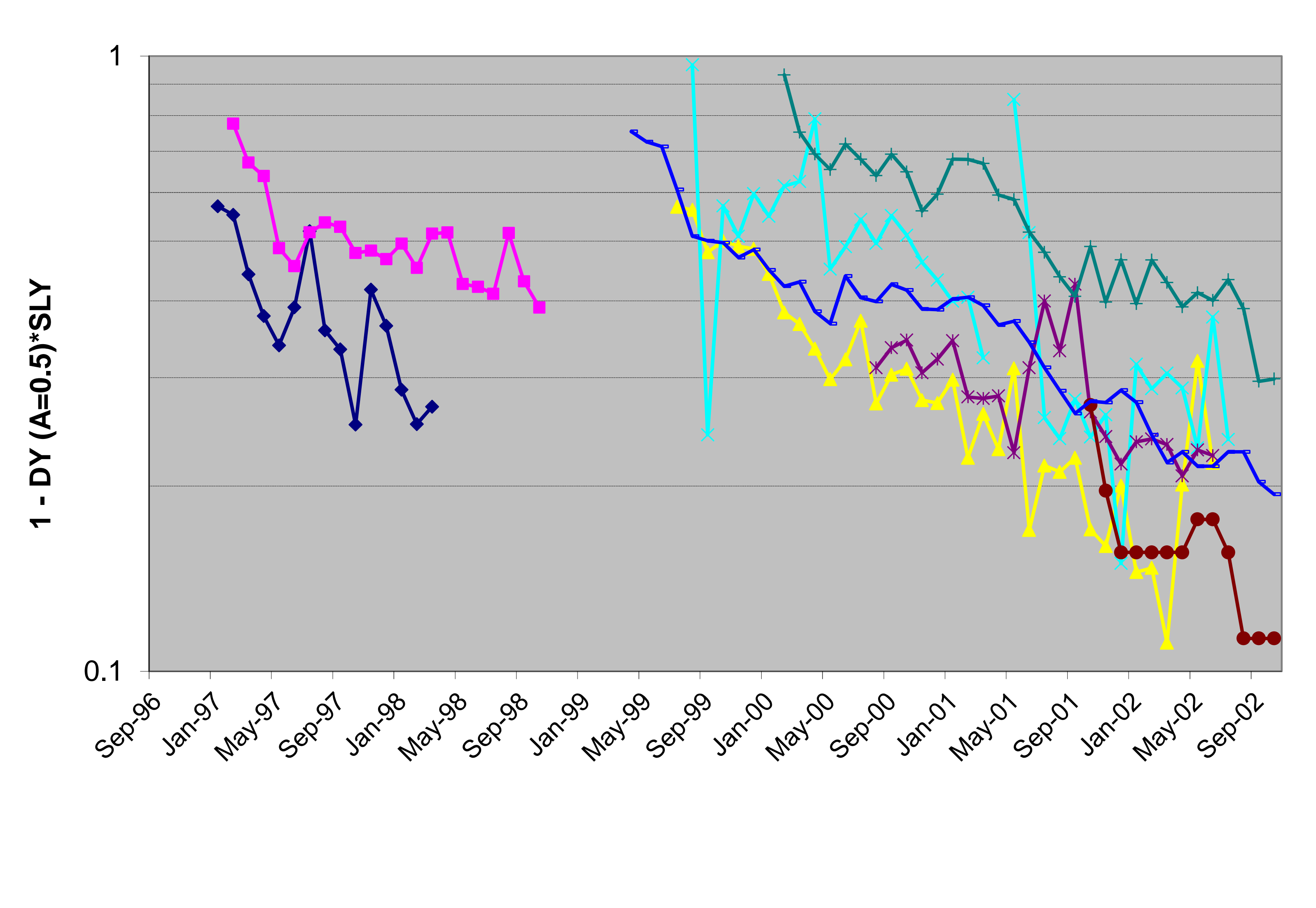

Figure 3. 350nm Yield Loss Curves

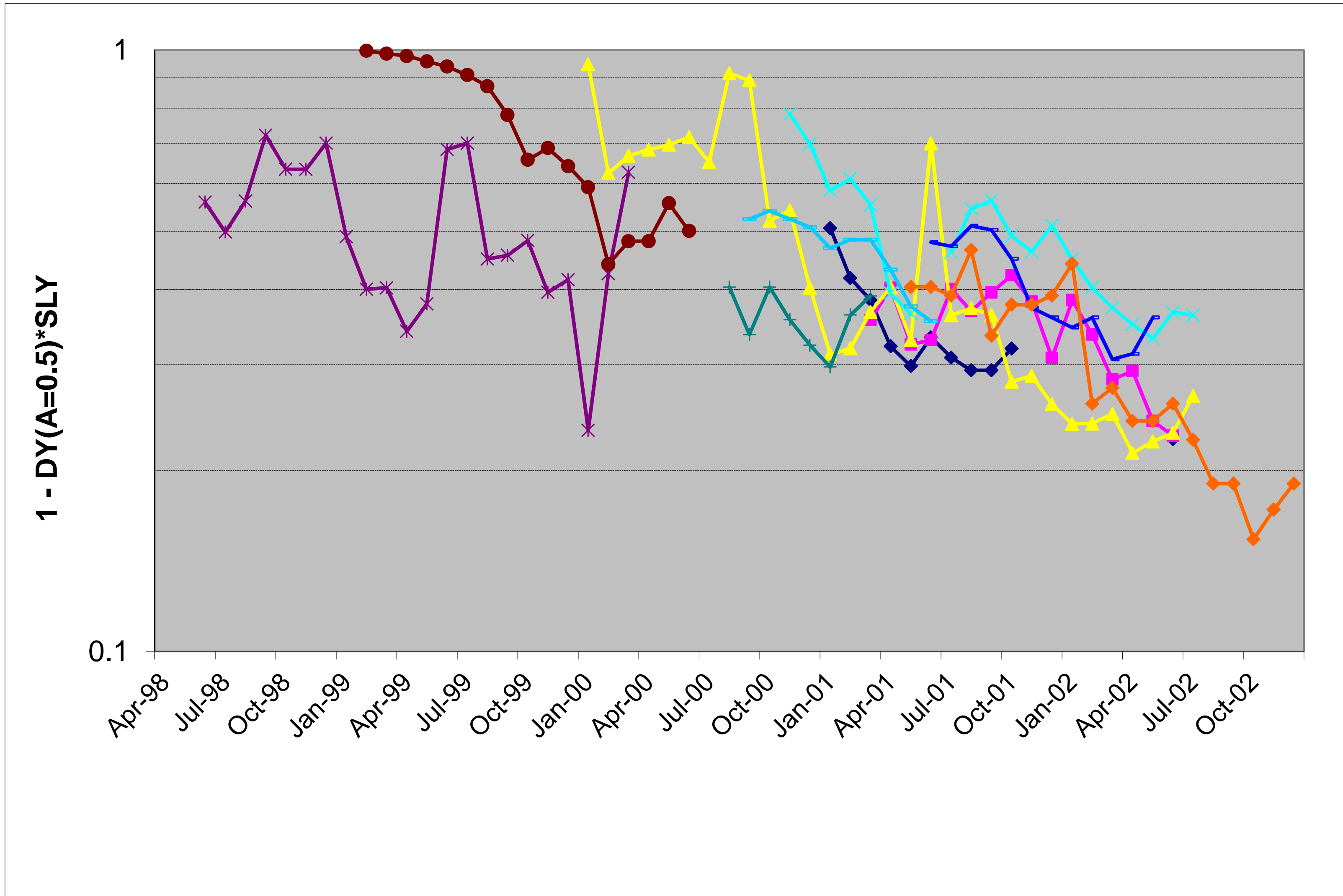

Figure 4. 250nm Yield Loss Curves

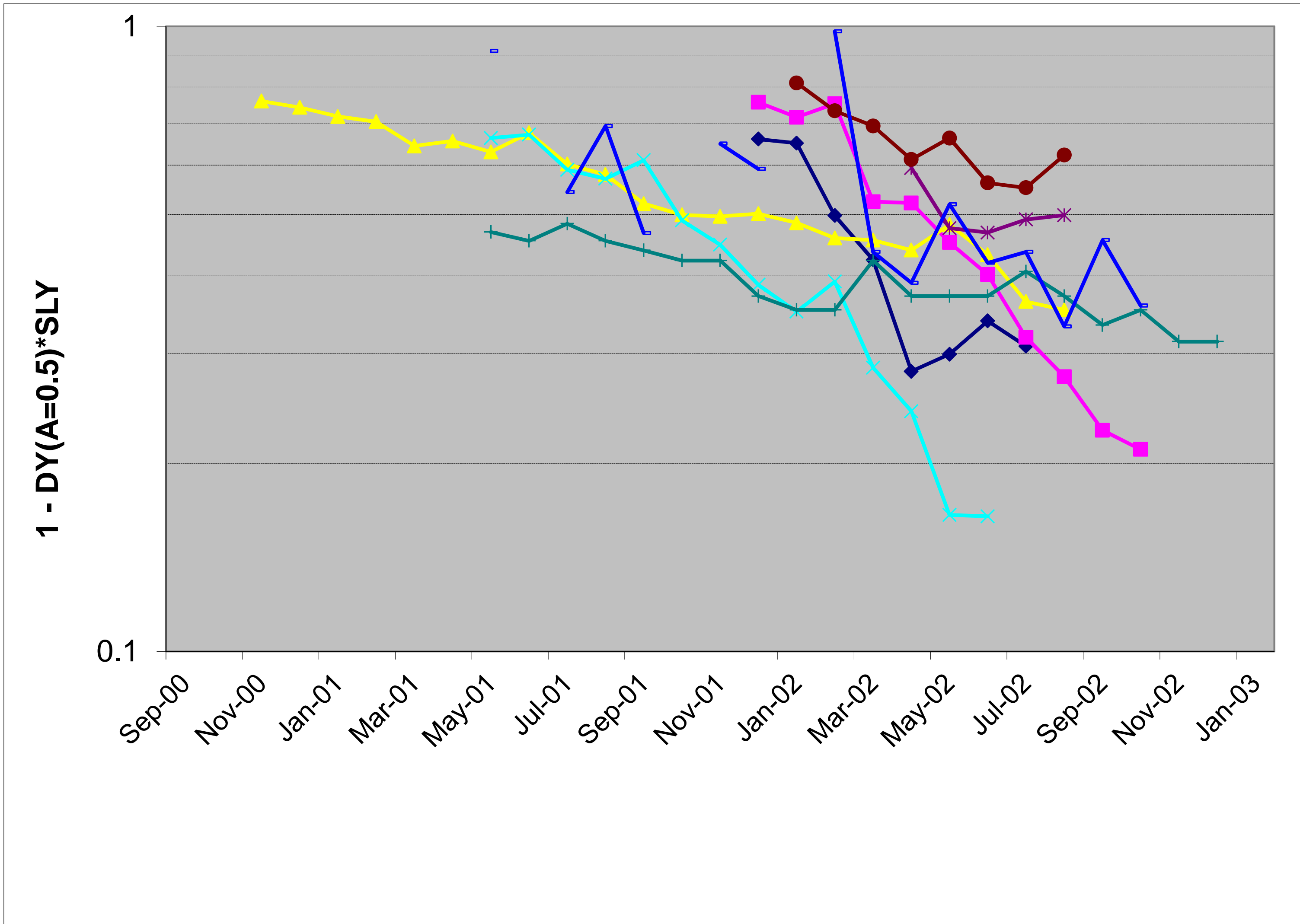

Figure 5. 180nm Yield Loss Curves

process maturity. Note that the maximum die yield achieved at maturity seems to be declining through the generations – more on this point below.

We obtained from the participants their dates for when development of the baseline process technology at each node was initiated, when the first wafer start of a salable product or product-like test vehicle was made at each node, and when the baseline process was qualified for mass production. Both the time of first wafer-start and the elapsed time from first wafer-start until qualification for mass production exhibited considerable variation across the participants. Notwithstanding this variation, we sought to find the industry-wide trend in the time required from first wafer-start in a technology until certain yield loss milestones were achieved. Table 1 summarizes what might be termed "envelope performance," measuring the elapsed times from when any participant made a first-wafer start in a new technology node until any participant reached specific yield loss milestones for that node.

Table 1
Elapsed Time to Reach Yield Loss Milestones

Node	**Date of first wafer-start**	**Elapsed time (months) until yield loss was reduced to:**			
		50%	**40%**	**30%**	**20%**
350nm	Nov-95	16	17	27	55
250nm	Apr-97	22	24	45	61
180nm	May-99	24	31	36	36

Note the disturbing trend that the elapsed time to bring yield losses down to 50% and 40% is steadily increasing across the technology generations. On the bright side, at the 180nm node, the participants achieved a substantial reduction in the time required to reach the 20% yield loss milestone.

To investigate this further, we numerically analyzed the individual yield learning rates of the participants, fitting the function

$$YL(t) = YL(0)^{*} \exp(-\alpha t)$$

to the curves depicted in Figures 3, 4, and 5. Here, t is measured in months and α is the average slope of the curve, i.e., the yield learning rate. We computed α for all curves with at least 5 months of data and then averaged them across all devices and participants. The results are summarized in Table 2.

The results in Table 2 reveal an encouraging development of accelerated yield learning at the 180nm technology node, rising from 4 – 4.5% per month at the 350nm and 250nm nodes to 6.5% per month at the 180nm node.

Table 2
Yield loss learning rates at various technology nodes

Node	Rate of Reduction in Yield Loss					
	Exponential learning rate (α)			Equivalent percentage reduction in yield loss per month		
	Worst	Best	Average	Worst	Best	Average
350nm	0.015	0.080	0.045	1.5	7.7	4.4
250nm	0.020	0.049	0.041	2.0	4.8	4.0
180nm	0.021	0.128	0.067	2.1	12.0	6.5

This result seems intriguing in light of the declining die yield at process maturity observable in the figures. To help understand this apparent contradiction, we decided to incorporate a larger data set. The Competitive Semiconductor Manufacturing surveys carried out from 1991 through 2000 at the University of California at Berkeley developed a large database from semiconductor fabrication companies worldwide that can be used to examine yield trends in the industry (Leachman, 2002). The yield data from the SMLY survey was augmented by the CSM data. The consolidated results are depicted in Figure 6, where the overall yield (die yield times sort line yield) for a representative group of companies fabricating random digital logic (i.e., exluding memory) process technologies is shown for the 500 through 180 nm generations, covering the time frame 1995 through 2002. While more advanced process technologies were in production at many companies during these surveys, the requirement that the data cover a large portion of the entire production life of a technology limits our data to the 180 nm and earlier generations.

In every technology generation the yield data from both the CSM and SMLY studies shows considerable variation from company to company, not only in its characteristics but also in the timing of process introduction. Because data from 11 different companies were available, each covering at least three of the four technology generations covered here, attempts to include yield history from all companies in Figure 5 were found to be too confusing to expose the trends determined from detailed studies of the data. Instead, for each generation, only three or four companies' yield histories determined to be representative of that generation are shown. Even tempering conclusions based on the large variations observed from company to company, the data exhibits two disturbing trends. First, the yield at initiation of manufacturing has declined sharply in moving from the 500 nm generation to the 180 nm generation; and second, the yield at process maturity has decreased with each successive process generation. The improvement in learning rates summarized in Table 1 is evident in Figure 6 as well (in the sense that the yield learning curves are getting steeper for more recent technologies), but evidently this improvement was not enough to make up for the worse starting points and shortened

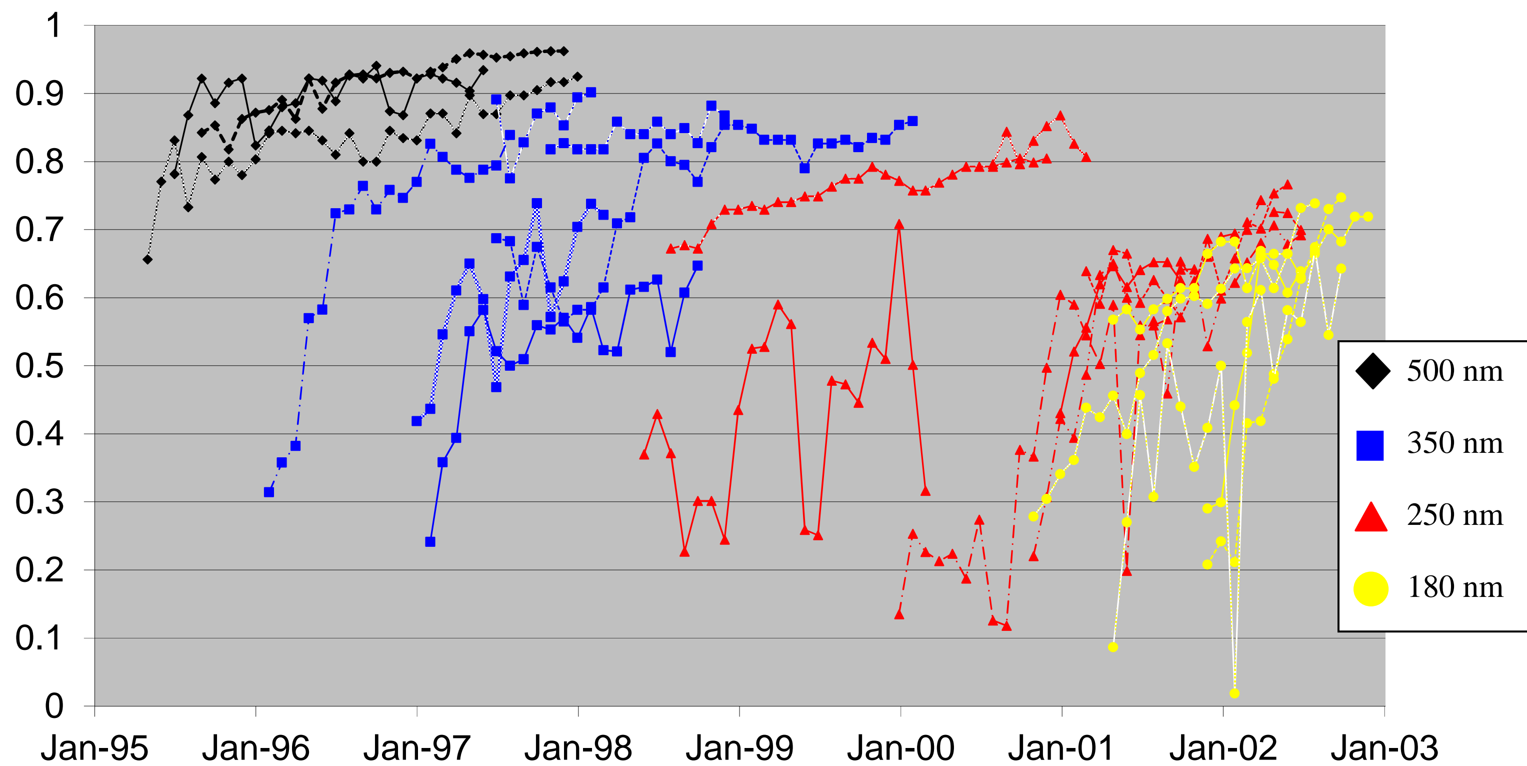

Figure 6. Random Logic Die Yield Trends at Selected CSM and SMLY Survey Participant

product lives. By and large, die yield is worse over the entire process life in each succeeding generation.

Systematic Mechanisms vs. Random Defect Limited Yield

The methodology introduced in Section 7 was applied to wafer maps furnished by the SMLY Survey participants for selected devices. Results are summarized in Table 3. Note that, in every case except one 350nm process technology, systematic losses dominate baseline random losses, even for mature technologies. This is not to say losses from mis-processing outweigh losses from defects; rather, it is to say that once defect excursions are separated from baseline contamination and viewed as systematic limiters, then losses from these and other systematic mechanisms in aggregate are more serious than baseline losses.

Table 3
Systematic and Baseline Random Yields Derived for Selected Devices at the SMLY Participants

Company	Process	Die Size (cm^2)	Overall Die Yield Y	Y_S	Y_R	D_0 cm^{-2}	Y/Y_{middle}
A	250 nm	0.75	57%	70%	82%	0.27	85%
A	250 nm	0.109	84%	87%	97%	0.29	98%
A	350 nm	0.388	71%	80%	89%	0.31	90%
B	130 nm	0.65	43%	83%	52%	1.01	85%
C	350 nm	0.177	88%	89%	98%	0.11	98%
C	350 nm	0.076	93%	94%	99%	0.19	98%
D	220 nm	0.56	76%	85%	90%	0.19	96%
D	180 nm	0.288	36%	55%	66%	1.44	85%
E	180 nm	0.66	82%	88%	93%	0.12	97%
E	250 nm	0.414	92%	96%	96%	0.09	99%
E	350 nm	0.673	94%	98%	96%	0.06	99%

The last column of Table 3 expresses the ratio of the overall die yield Y to the average die yield in the middle of the wafer Y_{middle}; this ratio serves to quantify the magnitudes of the edge losses. As may be seen, there is considerable variation in the amount of edge loss.

General trends in systematic yield losses reported by the participants are as follows:

(1) Back-end process steps contributed increasing yield loss as the participants transitioned through the 0.35um, 0.25um, and 0.18um technology nodes, both in terms of loss from particulate contamination as well as loss from malformed structures in the die.

(2) Photo-related losses also increased through the technology nodes. The usage of low-k lithography, involving mitigation measures such as optical proximity correction, phase-shift masks, tiling, lot-to-lens dedication, etc., increased through the generations. The adequacy of the accuracy in leveling and tilting of photo shots decreased. Missing patterns and islands are an example systematic yield loss mechanism. Such losses are typically intermittent, and they may have a spatial signature across the exposure field or across the die.

Prevalent Systematic Yield Loss Mechanisms

Edge loss was the dominant systematic mechanism across all participants. This is a difficult process integration issue, requiring tuning of photo, etch, deposition and CMP recipes.

Patterns or islands intermittently missing from photo shots. This is a leveling issue, whereby the topology of the top of the wafer is not level with respect to the lens or a reticle enhancement technique issue, whereby OPC makes the photo resolution more difficult. This issue was prominent for large die such as microprocessors.

Voids or blisters in very thin films, migration through very thin films, or poisoned vias. This issue was prominent for low-power wireless devices.

Lifting of entire films or layers. This problem is sometimes caused by too-aggressive cleaning steps.

Particle excursions. While there is a background level of contamination any given device experiences, much of the yield loss from contamination was associated with abnormal levels of particles. Such abnormal levels may be detectable and therefore losses from them can be contained. For this reason, some participants viewed particle excursions as a systematic loss.

Types of Spatial Signatures in Systematic Yield Loss

Edge loss was the most common spatial signature among the participants. This has to do with poor topology near the edge of the wafer in terms of incomplete films and/or certain films deposited onto the backside of the wafer and susceptible to delamination. Poor edge topology may involve deficiencies in lithography, etch, CMP and/or film deposition.

Other commonly-cited signatures on wafer maps for the various devices in the various technologies include the following:

- Edge losses heavier at certain angles, e.g., “beards” and “bananas”

- Exposure-field loss signatures (e.g., “checkerboards”)

- Photo-related losses steadily increasing towards the edge of the wafer

- Test loss signatures (e.g., whole rows of failed die)
- "Stalactites" and "fingers" (reflecting lifting of films or layers, sometimes caused by too-aggressive cleaning steps)
- Center loss, "bull's eye"
- Doughnuts, concentric rings (sometimes resulting from out-gassing or voids, sometimes resulting from improper equipment calibration)
- Lower-yield for first wafer in lot (sometimes resulting from excessive particle counts)

Evolution of Yield Loss over the Process Life

In the development phase, it is typical that large systematic yield loss mechanisms are identified and reduced, including most of the front-end, transistor-related issues. After transfer and during ramp, systematic losses continue to dominate. As the larger mechanisms are mitigated, other systematic losses and random losses are revealed. These losses could not be seen until the larger losses were mitigated. Many systematic losses are not discovered until later even though the problems have been present since the start of the process life.

Front-end systematic yield problems tend to dominate during development and early production, while back-end systematic and random yield problems tend to dominate later in the process life. These back-end losses were always there, but again the early-life front-end problems obscured them.

Identifying, Quantifying, and Prioritizing Components of Yield Loss

All participants employed the same general mix of techniques for identification of yield losses, summarized as follows. While this mix was present at all participants, the extent of application and the relative sophistication of each technique varied among the participants.

(1) *In-line inspection*. Optical and digital-image-scanning equipment are used to inspect a sample of production lots at various points in the process flow. A sample of wafers is scanned from each selected lot, and wafer order is preserved across inspection points so that defects added since last inspection point may be noted. Some or all of the defects on a subset of the inspected wafers are reviewed using review stations. The defects are reviewed and classified. Defects may be particles or they may be malformed structures, e.g., missing patterns. End-of-line wafer yield maps are overlaid on the maps of observed defects to establish "kill rates" for in-line observed particulate and non-particulate defects. In-line inspection and review equipment are also

used to try to recognize in-line defects or signatures corresponding to yield-loss signatures revealed by end-of-line test results (see (4) and (5) below).

(2) *E-test*. An electrical test is performed at a single in-line point (typically, just after first Metal Layer) on a sample of scribe-line structures on the wafer. A more complete test of scribe-line structures is performed at the end of the line. These tests provide data on electrical performance that may be correlated with die sort yields.

(3) If the device is a memory device or if it has embedded memory with access circuitry permitting rastering, *low-yield wafers or lots are rastered* to identify the nature of prominent failures in the memory cells that can be rastered.

(4) *Wafer maps*. Die yields by die site position on the wafer are reviewed, especially for low-yielding lots and for stacked wafer maps of one or several low-yielding lots. Spatial signatures are researched. Fail-bin maps are similarly analyzed. Systematic yield losses are estimated for the identified signatures.

(5) *Yields by wafer position in lot, by wafer quadrant and by rings on the wafer are examined for any signatures*. Systematic yield losses are estimated for the identified signatures.

(6) Wafers or lots contributing to the signatures identified in (4) or (5) above are subjected to a *commonality analysis* to research potential equipment and recipe sources of the yield loss. Other forms of commonality analysis consider temporal characteristics, e.g., queue time between steps, time between preventive maintenance and processing, etc. Most participants routinely perform commonality analysis on all lots achieving yields below a certain threshold.

(7) Wafers contributing to the signatures identified in (3), (4) or (5) and other low-yielding wafers may be subjected to *physical failure analysis* to establish root cause of particular die or cell failures.

(8) *The yield impacts from each of the observed mechanisms are tallied from kill rates* applied to observed defect rates for each inspection point, and from end-of-line estimated systematic losses by signature. The relative yield impacts from each of the identified mechanisms become the basis for prioritizing yield improvement efforts. Commonality analyses and failure analyses serve to facilitate assignment of process engineering staff to the yield problems.

(9) *Defect losses are categorized by "baseline" (i.e., appears in every lot) vs. "excursion"*. Excursion losses suggest weakness in process control, while baseline losses require fundamental process or equipment improvement.

(10) *The total "bottom-up" yield loss developed in (8) is compared to actual overall yield loss* in order to quantify the amount of yield loss not yet

explained. Also, the "bottom-up" yield loss is reduced by the amount to be gained from known solutions in-progress, and this adjusted amount is also compared to target yield and/or to overall yield loss to quantify the amount of yield improvement still needed to be identified.

We wish to emphasize that fabs tend to focus their yield improvement efforts on the mechanisms that they are able to *confirm* as causing yield losses. By "confirm," we mean they have found physical evidence of the mechanism so that they need not rely solely on statistical correlations of engineering data that suggest the presence of the mechanism. (One participant stated this point quite succinctly: "We work on the problems that we can see.") The nature of their methodology, emphasizing in-line defect inspection and end-of-line rastering of memories and/or failure analysis, tends to confirm particulate contamination mechanisms more readily than non-particulate mechanisms. Non-particulate mechanisms tend to be more difficult to detect with in-line inspection, they tend to kill entire memory blocks rendering them un-rasterable, and they are more difficult to sleuth in failure analysis laboratories. We believe this is an important reason why systematic mechanism yield losses are perceived as less than defect mechanism losses by some participants.

Interface with Non-Manufacturing Departments

Systematic yield losses often were the result of issues with product design, technology development and transfer, or product testing. We therefore reviewed practices of the participants in these areas and highlighted systematic yield losses over the technology generations stemming from interfaces with the non-manufacturing departments. Our major findings are highlighted below.

<u>Interface with Product Design</u>

At all participants, *design rules* are set relatively early during technology development, and once issued become very difficult to modify or change. Problems that show up later as a result of the design rules are usually resolved by process improvements, not design rule changes. However, increasingly there are manufacturability issues that are identified during process introduction and ramp that could be effectively addressed during design. As a result all participants have developed a complementary set of (often lengthy) "design for manufacturability guidelines" for each new process that is added to as information is obtained. Most participants felt that the designers in their companies poorly implemented these guidelines.

We found that numerous mask changes are typically made in the early life of lead products in each technology. Some mask changes are made to correct design flaws, to improve reliability or to implement design enhancements, but some are made to address specific yield problems identified by detailed failure analysis. The number of mask changes made for yield improvement is increasing over the technology generations.

At some point in the three technology nodes included in this survey, all but one of the participants implemented "lot-to-lens" dedication. This is the practice of running all

critical layers on a wafer through the same stepper or scanner. The reason for doing so is that yield was found to improve, and in fact to improve sufficiently to more than offset the manufacturing efficiency that results from this practice. However, the participants did not report any specific yield loss signature associated with this scheme. Instead a large number of the non-particulate fail bins improved to one degree or another when lot-to-lens dedication was implemented. Similar observations were made when a mask was replaced or several identical mask sets for the same device were run in the same fab. Even though all masks met specifications, yields were found to differ, again usually with no specific and unambiguous loss signature. The conclusion we reach is that there must be an increasingly strong yield loss mechanism or mechanisms whose root causes lie within the die (or at least within the stepper field). Furthermore, these yield loss mechanisms are evidently not adequately addressed either in the process and mask specifications or in the design rules.

Another important trend from the point of view of interface between design and manufacturing is the increasing application of resolution enhancement techniques. At some point in the three technology nodes covered by our study, all participating fabs implemented optical proximity effect corrections (OPC), phase-shift masks (PSM), and the addition to designs of tiling or metal squares to low-density regions (for improved post-CMP uniformity of film thicknesses). We expected that non-optimum implementation of these schemes would show up as systematic yield loss contributors with their own specific signatures. The schemes that are implemented in the mask, such as tiling, OPC and PSM, were defined and specifications created during the process development phase. We heard of only one instance at the participants where the design rules and associated algorithms for these schemes were improved for yield impact as the process technology matured, even though it might be expected that there would be considerable opportunity for optimization resulting from specifically tracking and monitoring the yield impact of these schemes.

In fact, the participating fabs rarely highlighted any systematic yield loss mechanisms due to process and electrical parameter variations within the die (what we shall term "intra-die systematics"). We believe that this is not because such loss mechanisms were not there, but rather because of the methodology they practiced to identify and prioritize systematic yield losses.

Fab yield management methodology for systematic mechanism limited yields focuses on inter-die and inter-wafer correlations. The sampling scheme used for parametric measurements and in-process control is sufficiently sparse that the measurements are not sensitive to many if not all intra-die error contributors. Similarly, sort yield data is inherently a sampling scheme with die periodicity. Any intra-die correlations will only be observed through failure analysis.

As long as the intra-die contributions to yield loss are small, existing fab methodology for yield management can address the issue satisfactorily through improving the inter-die process control to compensate, and there will be no necessity to separate the intra-die from the inter-die contributions. However it is apparent that the increasing contribution of intra-die parameter variations is a major contributor to the decrease in process windows and the increase in systematic yield loss observed in recent years.

Implementations of OPC, PSM, and tiling were not carried out by the product designers, but rather by a separate mask layout department serving as an interface between design and manufacturing. Uniformly, we found that the engineers in both the manufacturing lines and design groups that we interviewed were not familiar with the activities of these interface departments. The lack of communication on intra-die systematics between the design community and the fabs is an important issue that needs to be addressed if increases in systematic yield losses are to be contained.

<u>Interface with Technology Development and Transfer</u>

When many module changes are made at once in a new technology generation, integration gets exponentially more difficult. When the generation change features a materials change (e.g., interconnect transitioning from W to Cu), integration also gets much more challenging. To the extent that integration challenges can be paced and staggered through the technology generations, this mitigates systematic yield losses. For example, one participant did a half-generation between 250nm and 180nm at which interconnect was transitioned from W to Cu. The participant also changed stepper/scanner makes in between technology generations rather than at the generation change. Staggering these changes away from the main technology nodes avoided many systematic losses.

We heard about a number of events of systematic yield loss resulting from differences between source (development) fab and recipient (production) fab in terms of equipment makes and models, settings of equipment constants, calibration of equipment, process recipes, resists used, solvents used, etc. To the extent practical, application of the "copy exactly" policy mitigates such problems. This applies not only to specifications at time of transfer but also during development and subsequent to transfer if parallel production is pursued. It also applies to the propagation of "best-known methods" across multiple production fabs to standardize process execution and control to mitigate avoidable systematic yield losses.

We also heard about events where there were systematic yield crashes resulted from seemingly innocuous changes in DI water supply, solvents, cleaning or stripping intensities, etc. Clearly, any change in process chemicals or in application intensity must be carefully checked out to guard against systematic yield loss.

It is important to note that with one significant exception the methodology for systematic mechanism yield loss practiced by the technology development groups is virtually identical to that practiced after transfer to production by the fabs, except that the magnitudes of the systematic yield losses are larger. Signatures are identified and their yield loss prioritized, then the root cause of each major signature is determined through failure analysis, and finally corrective actions are taken. As each major yield loss contributor is eliminated, other systematic mechanisms become exposed. Eventually the yield reaches a point where transfer to manufacturing occurs and production ramp begins. Thereafter the fabs continue to identify and correct remaining systematic yield loss mechanisms using similar methodology.

The significant exception in yield management methodology that the technology development workers have is that they also have control over the intra-die systematics through mask layout during process development. Thus they can address systematic yield loss issues not only through the techniques available to the fabs, but also by modifying masks such as through modifications to OPC or tiling, or through modifying schemes for implementing PSM. Once the process is transferred to the fabs, however, such options for yield loss reduction do not seem to be available.

Interface with Product Testing

There were many cases where yield loss excursions were detected that were traceable to testing (i.e., die sort) equipment and software. Examples were reported of probe cards not making proper contact, tester problems, and test program errors that resulted in erroneously-reported yield loss. Full-functional fails of this kind typically exhibited a spatial signature that was readily detectable (zero wafers, or zero rows or checkerboards in the case of parallel testing). More subtle were timing problems that led to false-fails. Generally, test programs were written based on timing assumptions established during the early phases of technology development. Frequently, actual timings revealed after start of mass production were shifted to some extent from these assumptions, in some cases necessitating modification of the test program to mitigate test marginalities or even false fails. This kind of activity was pursued by product engineers.

Summary of Best Practices

There were many distinctive and effective practices noted across the participants. In this section we have highlighted the ones we found to be most significant.

(1) *Develop a complete-as-possible "bottom-up" explanation of yield losses, both random and systematic.* Average kill rates applied to average numbers of added defects observed at in-line inspection points are summed to express total yield losses observed by in-line inspection, separately computed for baseline defect counts and for excursions. Yield within each observed spatial signature on wafer maps is compared to average yield to develop loss associated with that signature. Yields in all other correlatable signatures (e.g., first-wafer-in-lot, particular wafer quadrant, wafer edge) are also compared to average yield to develop losses from those signatures. The bottom-up total of losses with known mechanisms is compared to actual total yield loss as well as to target loss. The bottom-up total is also adjusted to exclude loss from mechanisms with known solutions in order to track what losses would be left after solutions are implemented, once again comparing to actual and target overall loss. Table 6 provides an example "bottom-up" analysis. Figure 7 illustrates the notion of comparing identified yield losses and identified solutions with the yield target.

(2) *Use in-line inspection as intensively as possible.* In-line inspection equipment, review equipment, operator time, and technician time are always

fully utilized in order to gather as much data as possible. For example, during demand downturns, the number of in-line inspection points and the sampling rate are increased to gather more information and improve precision. The amount of inspection is set much higher for lead products in new technologies than for mature products and technologies.

(3) *Adjust in-line inspection to detect defects or signatures of losses found at end-of-line.* When new signatures or excursions are detected at end-of-line that are unrelated to defects observed by in-line inspection, an effort is made to adjust in-line inspection recipes and inspection points to be able to detect evidence of the loss mechanism in-line. Often times, adjustment of the inspection equipment was made enabling in-line monitoring to "see" such losses so they could be more easily traced to root cause as well as enabling better containment of the losses. For example, poisoned vias could be detected by re-focusing the inspection equipment to view the bottom of the vias rather in lieu of the standard focus to search for surface defects.

Table 7
Example Bottom-up Analysis of Yield Losses

Mechanism	**Average yield loss**
E Test	3%
Missing islands at metal layers (scanner leveling issue)	3%
Missing islands at implant layers	2%
Wafer edge losses (voids, peeling)	6%
Defect excursions (specific source tools identified)	7%
Subtotal, identified systematic mechanisms	**21%**
Contact pattern fails	3.8%
M1 – M5 pattern fails	3.8%
Poly 1 pattern fails	1.5%
Poly 1 particles	0.5%
STI HDP particles	0.3%
[long list follows of identified baseline particle mechanisms]	
Subtotal, baseline random defects	**23%**
Total identified yield losses	**44%**

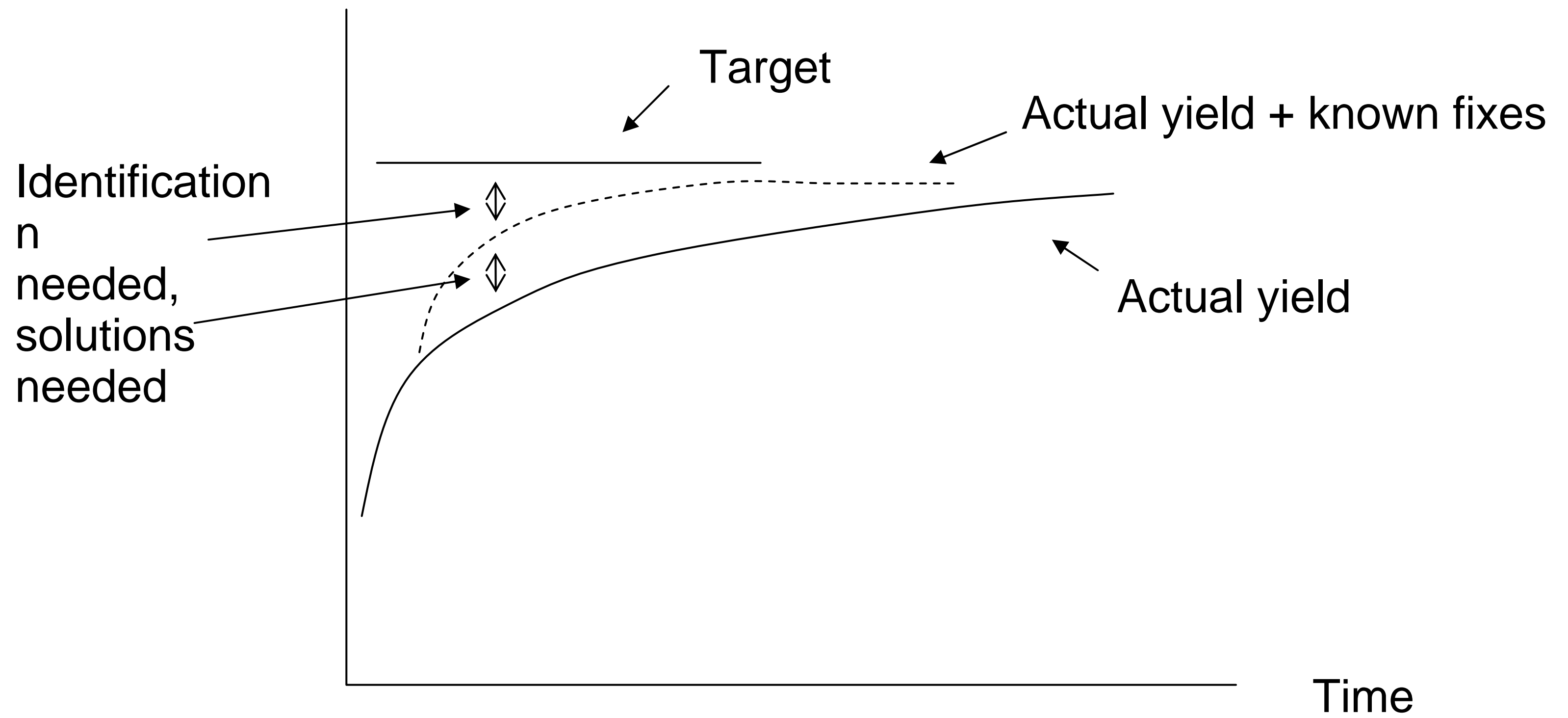

Figure 7.
Example tracking of yield progress

(4) *Perform statistical commonality analyses of lots or wafers with common low-yield signatures or common low-yield excursions.* Root-cause determinations were often accelerated by effective commonality analyses identifying common machines or chambers or common temporal characteristics such as queue time between steps, time since preventive maintenance, etc.

(4) *Find all spatial signatures of yield loss on the wafer maps and/or fail-bin maps to help reveal systematic problems*. A variety of analyses should be performed in this regard, e.g., yield signature by wafer position in the lot, by quadrant of the wafer, by rings of wafer area. As noted earlier, edge loss is perhaps the most predominant form of systematic yield loss. Studying yields in rings across the wafer, and comparing to the topology of the wafer and the films deposited on it across those rings, can help to reveal the root causes contributing to edge losses.

(5) *Perform yield analysis by block within the die*. We were struck by the remarkable variation of yield sensitivity across different kinds of devices in the same basic technology and across different blocks within the same device, e.g., memory cells, access circuitry, random logic blocks, analog blocks, etc. In a number of instances, important insight towards the resolution of yield problems was gained by assessing the yields of individual blocks within the die.

(6) *Reduce organizational boundaries and centralize yield responsibility.* We encountered a variety of organizational structures across the participants. In general, the fewer the hand-offs and the lesser the division of labor for uncovering yield losses, containing them and eliminating them, the better the results. Examples: Defect engineering made a part of device engineering and/or integration engineering rather than a part of process engineering. Failure analysis made a part of product or device engineering. We also saw great examples of personal interdisciplinary growth, whereby defect engineers or integration engineers learned considerable amounts of process and device technology on their way to determining root cause of losses and to developing effective process control procedures. These more multi-dimensional individuals accelerate yield improvement.

Moreover, we saw cases at the participants where different engineers were responsible for different kinds of yield loss (e.g., defectivity vs. systematic mechanisms). The decentralization of yield responsibility seemed to retard yield improvement. Generally, the more integrated the organization and the more centralized the responsibility for yield, the better results we saw.

(7) *Modify yield modeling efforts to exclude areas with systematic signatures when determining defect densities.* As we have argued earlier, typical top-down yield models are not oriented to separate losses with signatures from baseline losses present on every die site on every wafer. As a result, not all losses with spatial signatures are classified as losses stemming from systematic mechanisms. This raises the risk of a curtailment of efforts to

search for and mitigate systematic mechanisms once the modeled systematic losses are eliminated. Wiser participants excluded areas with systematic signatures from their yield models. For example, one participant made a 3mm wafer edge exclusion when calibrating its yield model, i.e., defect density calculations were based solely on yields achieved inside the excluded ring. As noted earlier, we suggested the alternative of fitting the binomial model to the observed Y_{MAX} in order to find Y_S and D_0. In general, application of the "windowing" scheme should be made only to areas of the wafer known to be free of systematic mechanisms. That is, this scheme should be applied only to areas of the wafer and to wafer positions within the lot that exclude spatial signatures, subsequently scaling Y_S to fit the overall yield.

(8) *To the extent practical, make technology transfers and parallel production reflect "copy exactly" of best-known methods. To the extent practical, stagger (through time) major materials changes and major module changes so as to minimize integration challenges and consequent systematic yield losses.* We saw numerous cases of systematic yield losses resulting from not copying exactly. We also found that technology nodes featuring more dramatic and simultaneous changes experienced more systematic yield loss.

(9) *Changes in process chemicals or chemical treatments should be carefully screened for systematic yield loss.* We saw several cases where seemingly innocuous changes to DI water, solvents, cleaning, or stripping recipes caused systematic yield crashes.

Recommendations

Suggestions for Fabrication Management

The bottom-up analysis of yield loss mechanisms described in this report is the primary methodology for quantifying yield loss mechanisms and prioritizing yield improvement efforts. This analysis should make a clear separation of the baseline, truly random defects (present on every die site and on every wafer, with no apparent spatial or temporal signature) from systematic mechanisms. A tally of bottom-up losses should be compared to top-down estimation of losses and to yield targets in order to gauge what fraction of total loss has been found and what fraction of needed yield improvement is accounted for by known solutions. The top-down estimation should be based on a decomposition of overall yield losses into baseline losses with no spatial or temporal signature vs. losses with such signatures, i.e., *systematic losses.*

In-line inspection equipment are traditionally focused to detect particles. This equipment also can be used to detect non-particulate mechanisms, and this should be done so whenever practical. This equipment should always be used up to its capacity to learn as much as possible about yield loss mechanisms, both particulate and non-particulate.

To the extent practical, transfers of process technology need to follow the "copy exactly" principle. This has implications for R&D organizations (i.e., they need to develop technology that *can* be copied exactly by production fabs, considering the equipment sets and equipment settings, gases, and chemicals that they have access to) as well as for production fabs (i.e., avoiding unnecessary risks of serious systematic loss mechanisms associated with adapting different equipment, chemicals, solvents, etc.).

Yield improvement efforts need to be shared and centralized. Yield improvement is inherently interdisciplinary. Effective yield engineers always have undergone substantial personal growth taking them well beyond their disciplinary training; in short, they need to be successful at learning on the job. Given the difficulty and dynamic nature of classifying loss mechanisms into systematic and defect categories, organizations that segregate engineers by systematic and defect mechanisms seem less effective.

We found evidence the intra-die loss mechanisms are growing in seriousness. By and large, such mechanisms are not effectively treated by the participants; in most cases, they are not even identified. Devices now may incorporate many IP blocks designed by different groups within and/or outside the company. Yields by block need to be analyzed. Even within a block, yield may vary, due to ineffective resolution enhancement techniques or simply because certain aspects of the design are marginally incompatible with the capabilities of the process technology. At present, the participants' yield management methodology for systematic mechanism limited yields focuses on inter-die and inter-wafer correlations. New sampling schemes and measurements for in-line process control, new statistical analyses, and new testing schemes need to be devised and implemented to analyze intra-die loss mechanisms. Coordination and communication between designers, engineering staff implementing resolution enhancement techniques, and yield engineering staff also needs improvement.

Suggestions for the IC design community

Given the growing evidence of yield loss from intra-die systematic mechanisms, it is useful to ponder what actions the design community might take to alleviate fab yield problems. One obvious action is for the design community to enhance their device simulation capability to include not only the conventional inter-die and inter-wafer parameter variations to be faced in production, but also to include informed values for parameter variations within their chips due to intra-die process issues. In most cases today simulations inherently assume that all parameter variations occur uniformly across the die and ignore intra-die variations. A few leading-edge companies are beginning to also model intra-die variations and modify their designs based on the results. But even in these cases the actual parameter variations experienced in production over time as the process matures are rarely tracked or included in the design methodology (Berglund, 2003). It would appear that there is significant opportunity here for designers to optimize IC designs such that they are less sensitive to intra-die parameter variations, thereby improving acceptable process windows for these parameters in the fab and reducing any associated inter-die and inter-wafer systematic mechanism yield losses.

A second possible action is for the design community to proactively include in their methodology the fact that design rules must change over life of a process. Today the

Design for Manufacturing Guidelines issued by fabs are rarely incorporated formally into revised design rules. In fact changes in design rules for a process after initial product qualification are difficult and rare. At the same time the fabs are continually identifying and quantifying yield improvement actions that could be taken through IC design. These are documented in their guidelines and could significantly improve yield and mitigate or eliminate yield loss mechanisms. Complementing these actions by proactive design community actions through design rule changes, implemented on new designs (and on redesigns of existing products when economically justified), seems to have considerable potential.

A third and more radical possible action is for the design community to acknowledge that yield improvement is no longer the sole responsibility of the fabs, and that expecting to achieve yielding integrated circuits simply by following a set of design rules is no longer an approach that will result in predictable success. Instead the design community might accept that they are an integral part of the yield improvement methodology throughout the process life, and work with the fabs to identify, develop and implement a new methodology for attacking the intra-die yield loss mechanisms. Actions might include defining new in-process monitors for intra-die parameter variations and helping relate such measurements to both yield and IC design, and modifying designs for improved testability and identification of yield loss mechanisms.

References

Berglund, C. Neil, "Trends in systematic non-particle yield loss mechanisms and the implications for IC design," *Proceedings of the SPIE Conference*, 2003, Sunnyvale, CA.

Cunningham, J. A., "The use and evaluation of yield models in integrated circuit manufacturing," *IEEE Transactions on Semiconductor Manufacturing*, vol. 3, no. 2, pp. 60-71, 1990.

Leachman, Robert C., "Competitive semiconductor manufacturing: final report on findings from benchmarking eight-inch, sub-350nm wafer fabrication lines," CSM-52, , Engineering Systems Research Center, University of California at Berkeley, Berkeley, CA 94720, 2002.

Murphy, B. T., "Cost-size optima of monolithic integrated circuits," *Proceedings of IEEE*, vol.52, pp. 1537-1545, 1964.

Stapper, C. H., "Integrated circuit yield management and yield analysis," *IEEE Transactions on Semiconductor Manufacturing*, vol. 8, no. 2, pp. 9-102, 1995.