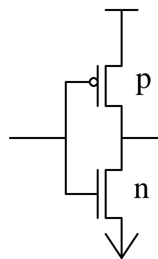
Next Time on Parametric Yield

- Current and Advanced DFM techniques
  - Worst Case Files
  - The role of process simulation (TCAD)
  - Complete Process Characterization
  - Statistical Design
- The economics of DFM

What Drives "Worst Case" Analysis?

- Basically, optimize an inverter design…
- … then, since most designs are just combinations of inverter-like gates…
- …it follows that the entire design would be OK even at the extreme points of process variation!



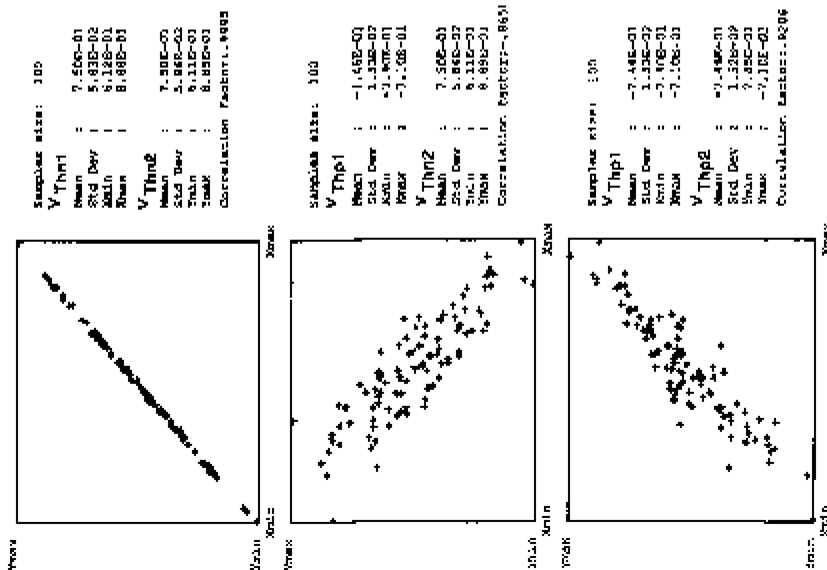
Process domain:
Vt_n, ΔL_n, ΔW_n, kp_n,
Vt_p, ΔL_p, ΔW_p, kp_p
Tox, T, Vdd
Performance domain:
τ, P
Yield Body:
$P < x\ \mu W$, $t < y$ nsec.

Simple Digital Worst Case

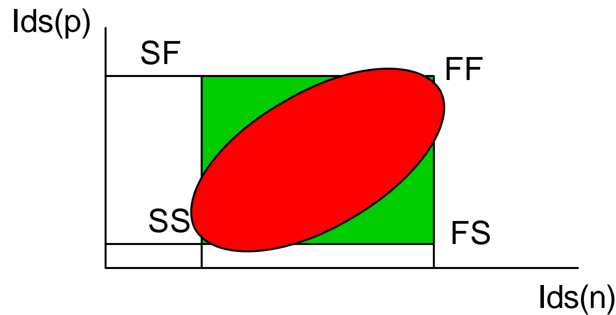
I_{dsat} sets power and speed in digital logic, so the extremes of I_{dsn} and I_{dsp} set the “performance” spread.

- Method
  - Identify typical, fast and slow N, P.
  - Extract the respective process parameters.
  - Name the cases TT, SS, FF, SF, FS, respectively.
  - Make sure performance meets specs at extremes.

Parametric Transistor Measurements



Simple Digital Worst Case (cont)



- Problems
  - It implies that the yield body is convex.
  - The Box might be unnecessarily big (over-design).
  - Idsat extremes do not always map to performance extremes.
  - Local variability is ignored (under-design).

Table Driven Digital Worst Case

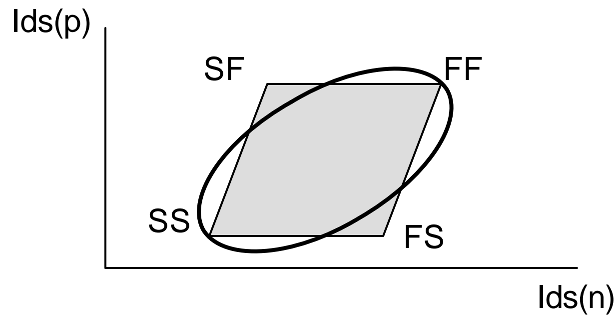
- N and P device behavior is highly correlated!

Process	Ids(n)	Ids(p)	Related?
Cox	↘	↘	Yes
DW	↘	↘	Yes
DL	↗	↗	Yes
LatDiff	↗	↗	no
msurface	↗	↗	no

Method

- Use table to relate process variability to Idsat deviations.
- Identify extremes through measurements.
- Extract the process parameters for each.
- Call them TT, SS, FF, SF, FS, respectively.
- Make sure performance goals are met by all.

Table Driven Worst Case (cont.)

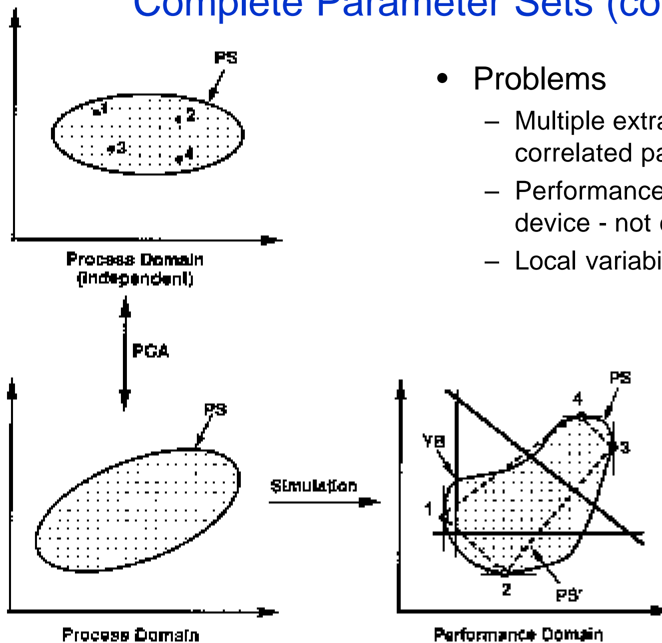


- Problems
  – It still implies that the Yield Body is convex.
  – Idsat extremes do not always map to the actual performance spread.
  – Local variability is ignored (under-design).

Working with "Complete" Parameter Sets

- Do not just focus on Idsat.
- Assume that all process parameters are varying. Independence cannot be assumed - use Principal Component Analysis (PCA) to transform the Problem.
- Method
  – Extract process parameters from a population of devices.
  – Find the correlation matrix of the process parameters.
  – Apply PCA to transform to an independent parameter space.
  – Use Constrained optimization to find the performance extremes (space is nicely convex).

Complete Parameter Sets (cont.)

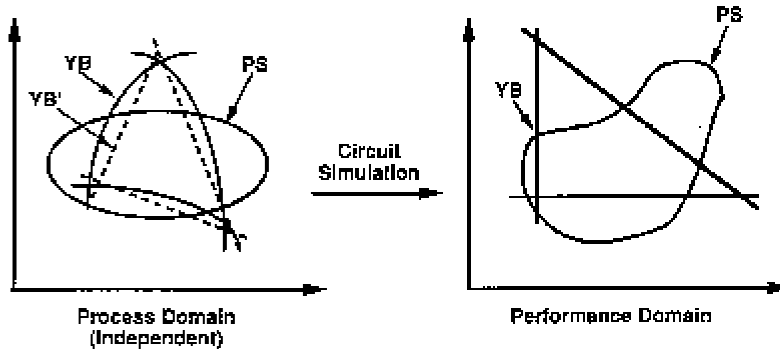


- Problems
  - Multiple extractions of correlated parameters.
  - Performances analyzed at device - not circuit level.
  - Local variability is ignored.

Working with "Reduced" Parameter Sets

- Assume that only few process parameters (V_{fb}, T_{ox}, L, W) and operational parameters (V_{dd}, T) are varying independently. (No PCA necessary!)
- Use circuit simulation to define extremes.
- Method
  - Develop special "statistical" device model.
  - Perform n+1 circuit simulations at extremes.
  - Establish linear approximation of yield body.
  - Integrate Yield numerically.
  - (Find Yield gradient, optimize yield)
  - (Non-linear modeling and complex experimental designs also possible)

Reduced Parameter Sets (cont.)

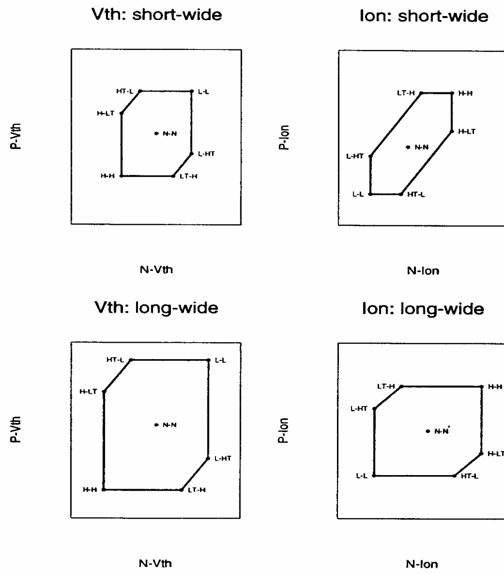


- Problems
  - Special device modeling considerations.
  - Linear approximations good for high yield only.
  - Local variability is ignored (no analog designs!)

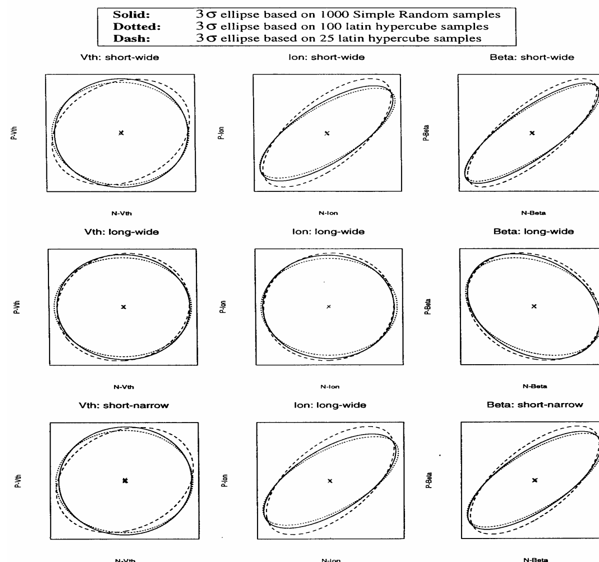
Modern Worst Case Files...

- We now fully realize that the device level worst case files are just an intermediate abstraction to capture the process variability.
- We also know that process parameters are highly correlated.
- So, state-of-the art techniques attempt to approximate the “ellipse” that typically describes the distribution of device parameters.

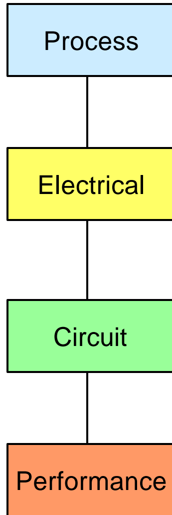
Modern Worst Case files



Carefully designed arrays of simulations can reproduce statistics of process

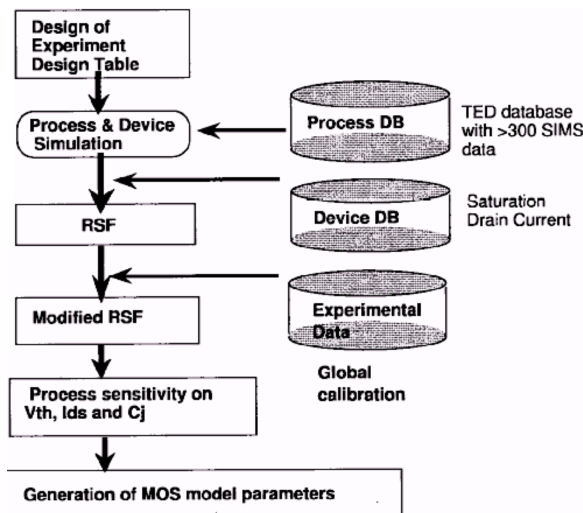


Time out! What do all these Techniques need?



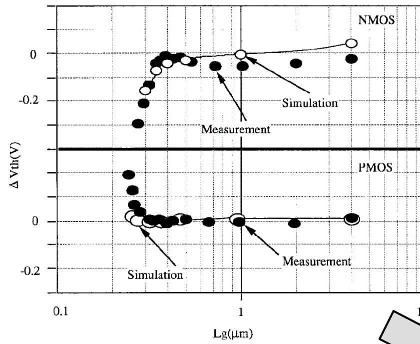
- Accurate process characterization!
- Deep understanding of how variability propagates from one level of abstraction to the next!

An Example of Using TCAD in Process Characterization

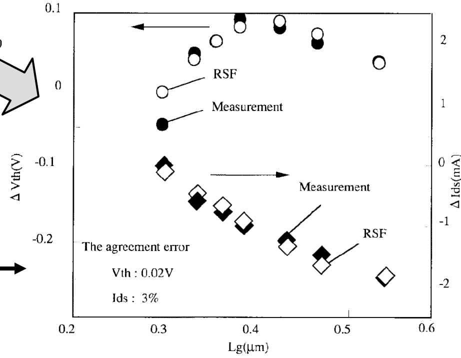


Accurate Statistical Process Variation Analysis for 0.25- m CMOS with Advanced TCAD Methodology, Hisako Sato, Hisaaki Kunitomo, Katsumi Tsuneno, Kazutaka Mori, and Hiroo Masuda, *Senior Member, IEEE, IEE TSM, Vol 11, No 4, November 1998*

The basic *Macromodeling* Idea

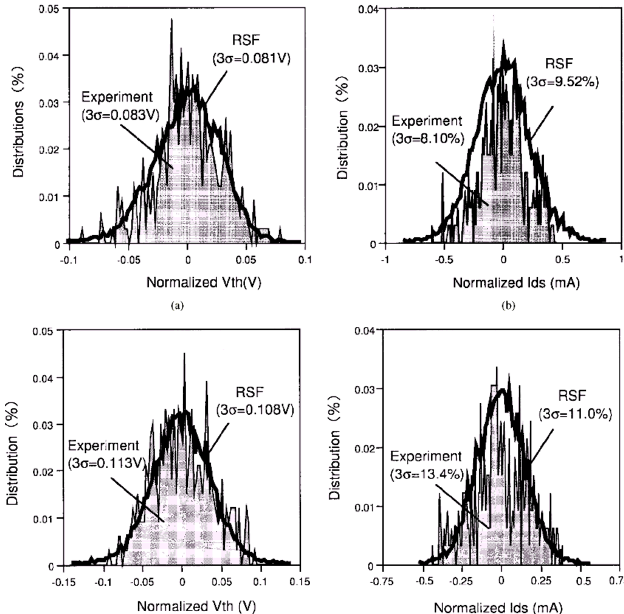


1. "tune" process/device simulator for good agreement with process.
2. Run the tuned simulator over a "designed" experiment.



3. Fit polynomial (*macromodel*) to simulator results.
4. Replace costly simulator with inexpensive (but rather "local") macromodel.

Good Statistical Match Can be Achieved...



How about Statistical Process Simulation?

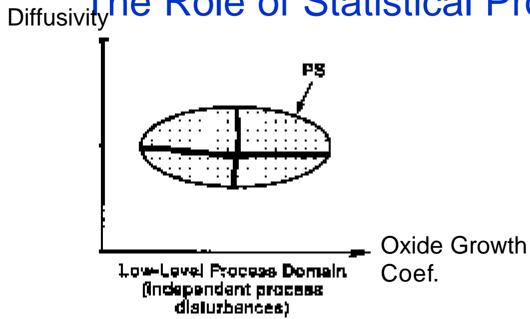
- Low level, physical process parameters can be used as *statistically independent disturbances*.
- Method
  - Identify process disturbances.
  - Infer multi-level statistics of process disturbances.
  - Use fats process and device simulation to generate device parameters.
  - Global *and* local variations can be represented.

Statistical Process Simulation

- Compact Process / Device Models.
  - Typically 1-D or compact 2-D models.
- The concept of process disturbances.
  - Low level process parameters that are, almost by definition, independent from each other.

The Role of Statistical Process Simulation

Diffusivity

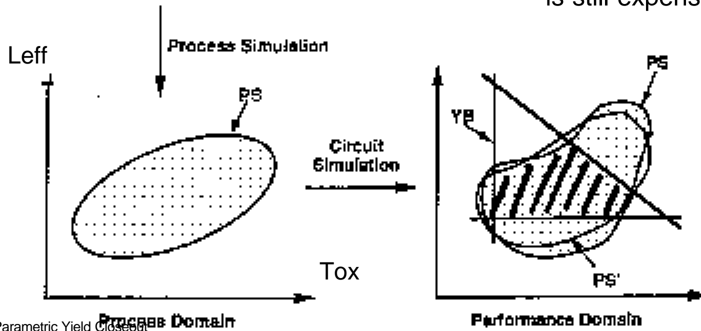


Oxide Growth Coef.

Problems

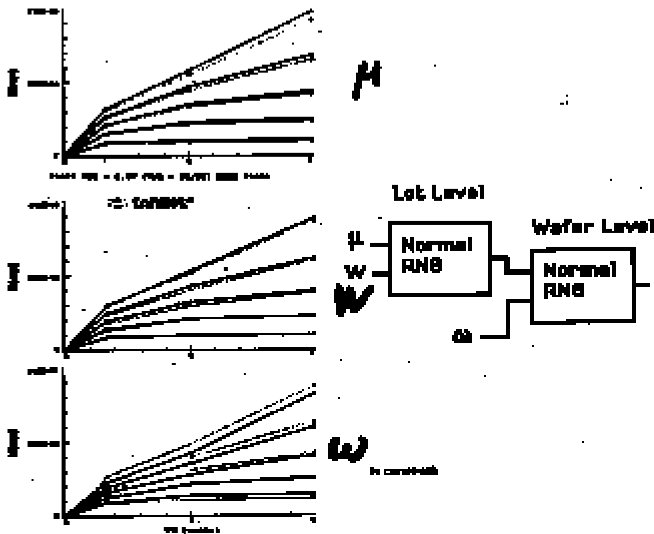
- Process Characterization is difficult
- Modeling accuracy is an issue
- Monte Carlo simulation is still expensive

Leff

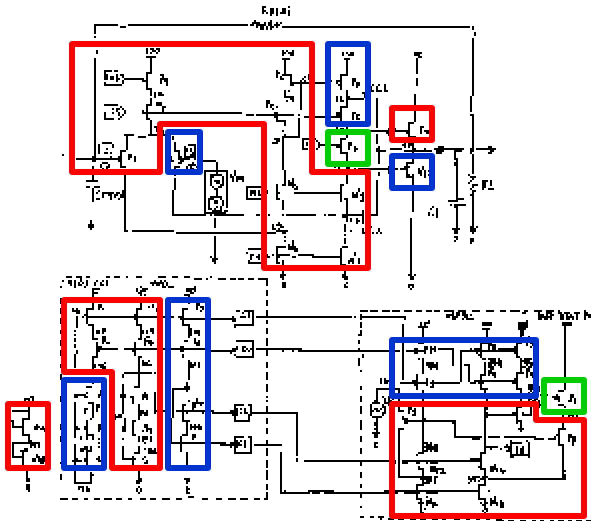


Tox

Possible to Model *Statistics* of Device Performance



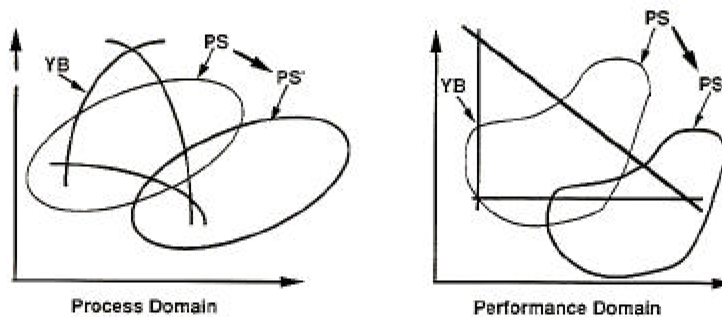
For complex Analog Circuits...



- "Matching" matters.
- It is impossible to consider matching of individual devices
- It is practical to "group" devices, and consider matching across groups.

Why do we *really* loose Yield?

- Catastrophic defects due to contamination.
- "One time" departures from Statistical Control.



Problems with earlier Techniques

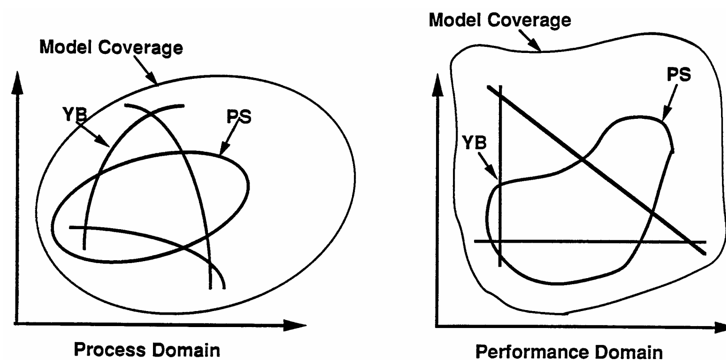
- Earlier Techniques assume:
- The entire fab line is “under SPC”.
- Critical steps can be “characterized” via measurements.

- *VLSI Technologies, however, are often too short-lived and too complex to achieve SPC status!*

A Binning Prediction Problem

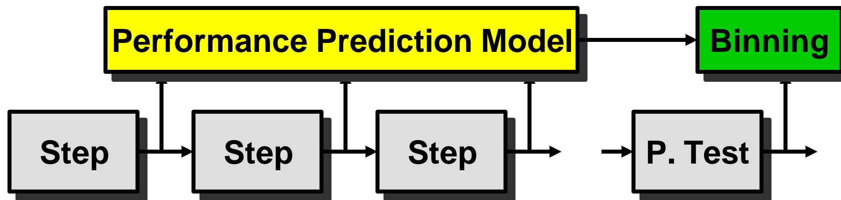
Create a model that relates yield to easy-to-measure, in-line and electrical parameters.

Base model on process physics - so that is is valid even when the process is out of SPC.



The Binning Prediction (cont.)

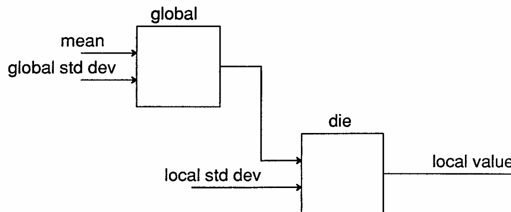
- Use this model in conjunction with measurements to predict functional yield (speed binning) early on.
- This will give early feedback about the line and the product in it.
- Some of the costly performance testing might be avoided.



The Generic Process Model

Mean and Std. Dev. (σ) for NMOS Process Parameters

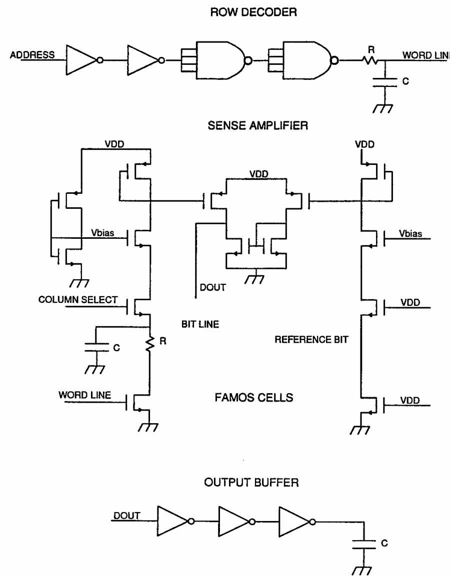
Parameter	Mean	Global σ	Local σ	Units
ΔL	0.30	0.10	0.025	microns
ΔW	0.15	0.05	0.0	microns
T_{ox}	35.0	1.5	0.38	nm
N_{subs}	5.5 x 10^{15}	1.67 x 10^{15}	0.42 x 10^{15}	cm^{-3}



NMOS/PMOS Global Parameter Correlations

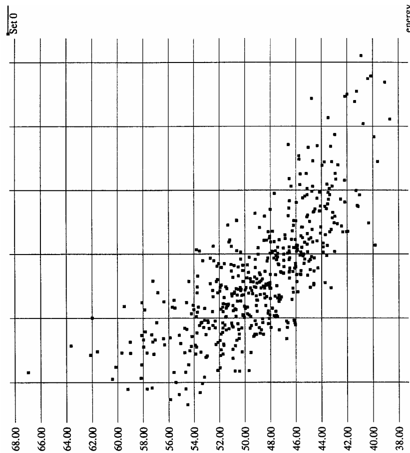
Parameters	N-ΔL	N-ΔW	N-T_{ox}	N-N_{subs}
P-ΔL	0.75	0.0	0.0	0.0
P-ΔW	0.0	1.0	0.0	0.0
P-T_{ox}	0.0	0.0	1.0	0.0
P-N_{subs}	0.0	0.0	0.0	0.0

The Specific IC Product Model

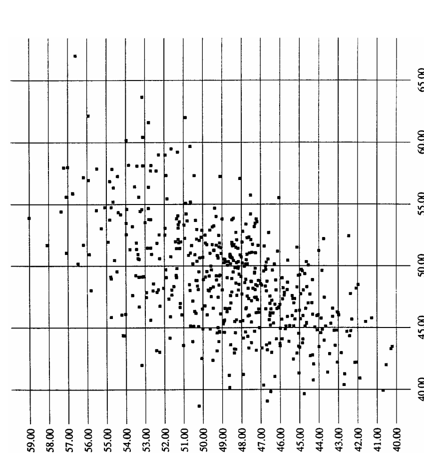


Simulated Statistical Distribution of Performances

Energy vs. Read 0 Access time



Read 0 vs. Read 1 Access Time



Yield Body Representation

- Using a linear model, each side of the yield body is a hyperplane of the form:

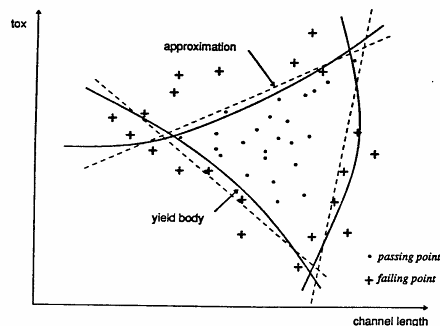
$$Performance = A_0 + A_1(N_{ld}) + A_2(N_{tox}) + \ldots$$

- The adequacy of the model can be examined by plotting the distance of good and bad parts from the hyperplanes:

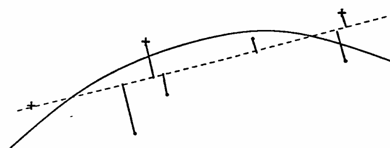
$$d_p = AX_p$$

- Projecting constraints into two dimensional planes gives insight into how to improve the model.

Linear Approximation of the Yield Body



Detail of one Constraint



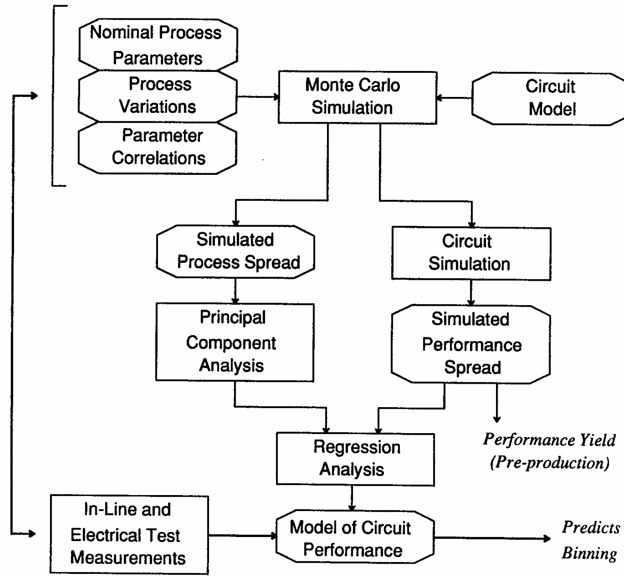
Test Patterns

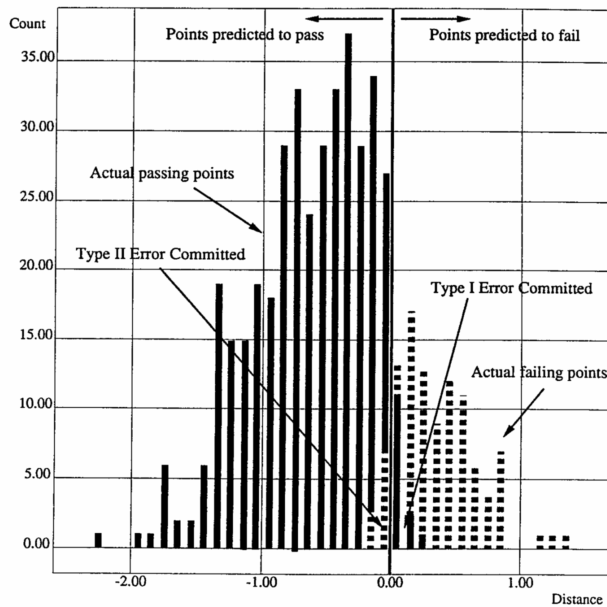
- We need to characterize both interconnect and active areas of the chip. Test patterns include:
  - Transistor arrays (three sizes of each type).
  - Capacitors for three typical oxide thicknesses.
  - Polysilicon structures for word line resistance.

In-Line and Electrical Measurements

- In order to use this model as a yield predictor, we have to “feed” to it easy to collect measurements.
- Standard test patterns are in use to collect Tox, Vt, kp, ΔL, ΔW, doping levels, poly resistivities.

Overall Binning Prediction Methodology





Issues in IC Manufacturability

- Importance and Causes of Yield Loss
- A Formulation of the DFM Problem
- Circuit Design for Manufacturability
- Yield Modeling Methods
- CIM and DFM

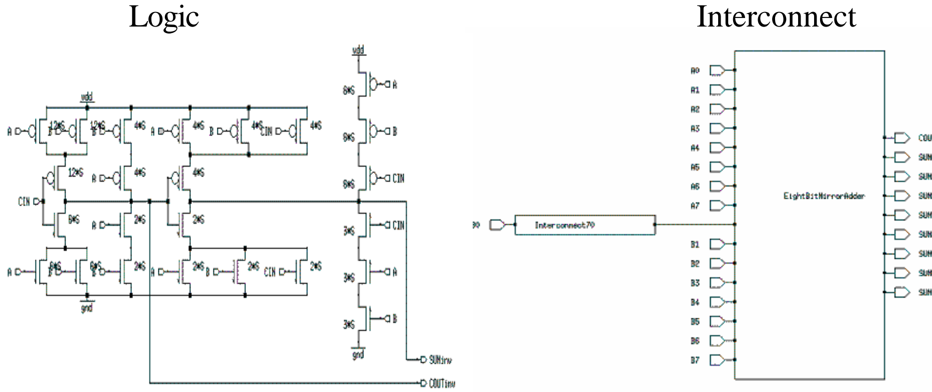
Projected Parametric Variation in Future Nodes

Technology	180nm		130nm		100nm		70nm	
Device	NMOS	PMOS	NMOS	PMOS	NMOS	PMOS	NMOS	PMOS
L_{eff}(nm)	100 $\mp$ 16.7%	100 $\mp$ 16.7%	65 $\mp$ 16.7%	65 $\mp$ 16.7%	45 $\mp$ 16.7%	45 $\mp$ 16.7%	28 $\mp$ 16.7%	28 $\mp$ 16.7%
T_{ox}(Å)	22 $\mp$ 4%	22 $\mp$ 4%	16 $\mp$ 4%	16 $\mp$ 4%	13 $\mp$ 4%	13 $\mp$ 4%	10 $\mp$ 4%	10 $\mp$ 4%
R_{dsw}($\Omega - \mu m$)	250 $\mp$ 10%	450 $\mp$ 10%	200 $\mp$ 10%	400 $\mp$ 10%	180 $\mp$ 10%	300 $\mp$ 10%	150 $\mp$ 10%	280 $\mp$ 10%
X_j (nm)	30		24		20		13	
V_{th}(V)	0.214 $\mp$ 30%	$-$0.327 $\mp$ 30%	0.232 $\mp$ 30%	$-$0.273 $\mp$ 30%	0.217 $\mp$ 30%	$-$0.253 $\mp$ 30%	0.169 $\mp$ 30%	$-$0.218 $\mp$ 30%
Interconnect	Local	Global	Local	Global	Local	Global	Local	Global
E	3.5 $\mp$ 3%		3.2 $\mp$ 5%		2.8 $\mp$ 5%		2.2 $\mp$ 5%	
w(nm)	250 $\mp$ 20%	525 $\mp$ 20%	175 $\mp$ 20%	335 $\mp$ 20%	123 $\mp$ 20%	237 $\mp$ 20%	85 $\mp$ 20%	160 $\mp$ 20%
s(nm)	250 $\mp$ 20%	525 $\mp$ 20%	175 $\mp$ 20%	335 $\mp$ 20%	123 $\mp$ 20%	237 $\mp$ 20%	85 $\mp$ 20%	160 $\mp$ 20%
t(nm)	500 $\mp$ 10%	1050 $\mp$ 10%	280 $\mp$ 10%	670 $\mp$ 10%	197 $\mp$ 10%	498 $\mp$ 10%	145 $\mp$ 10%	325 $\mp$ 15%
h(nm)	500 $\mp$ 15%	1050 $\mp$ 15%	280 $\mp$ 15%	670 $\mp$ 15%	197 $\mp$ 15%	498 $\mp$ 15%	145 $\mp$ 15%	325 $\mp$ 15%
ρ (Ω-m)	2.2e-8 $\mp$ 30%		2.2e-8 $\mp$ 30%		2.2e-8 $\mp$ 30%		2.2e-8 $\mp$ 30%	
Rvia(Ω)	23 $\mp$ 20%		25 $\mp$ 20%		27 $\mp$ 20%		29 $\mp$ 30%	
Length(μm)	62.5	5000	55.56	3333	50	2500	45.45	2000
Wn(μm)	1.26	15	1.01	9.75	0.76	6.75	0.48	4.2
Dynamic								
V_{dd}(V)	1.8 $\mp$ 10%		1.2 $\mp$ 10%		1.0 $\mp$ 10%		0.9 $\mp$ 10%	
Tr(ps)	160		125		110		87	
Temp (^oC)	25		25		25		25	

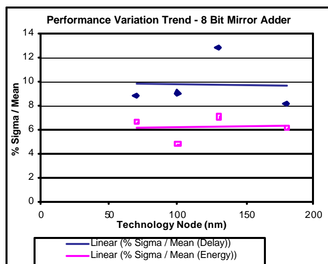
Projected Parametric Variation in Future Nodes Test Circuits used to facilitate Analysis

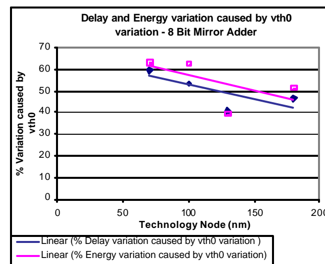
Logic

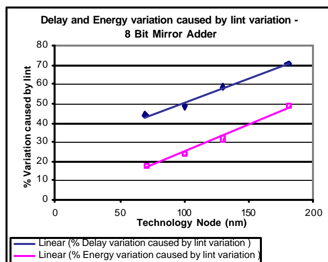
Interconnect

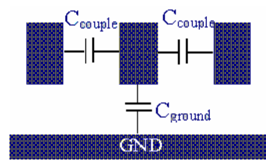


Projected Parametric Variation in Future Nodes

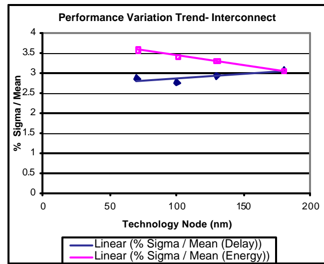


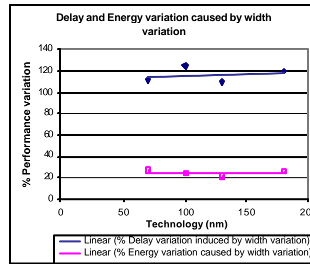


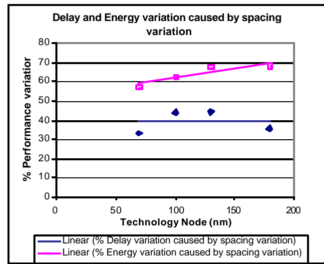


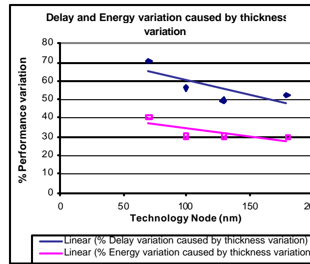


Projected Parametric Variation in Future Nodes

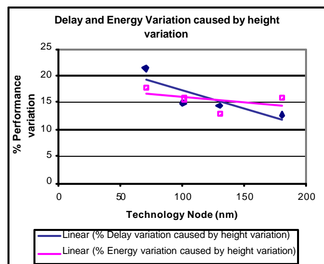


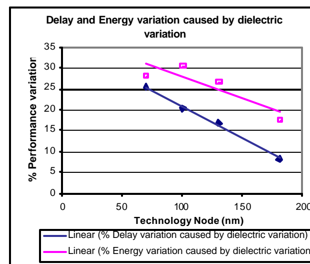


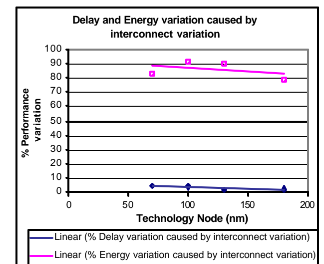


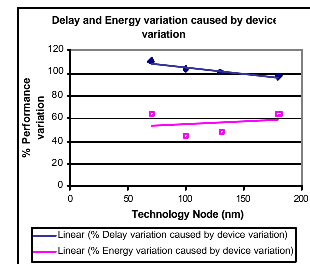


Projected Parametric Variation in Future Nodes

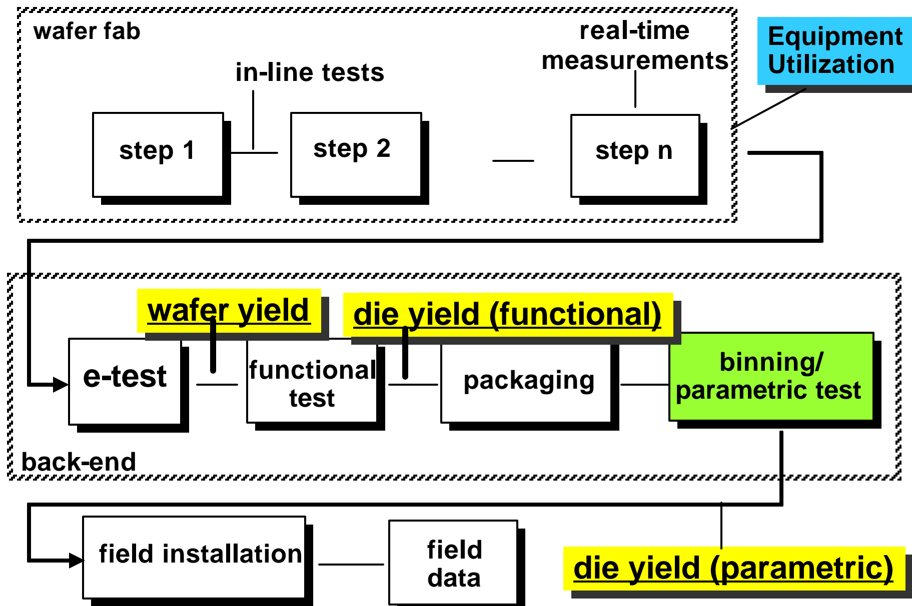








Let's Summarize...



IC production suffers from routine and assignable *variability*.

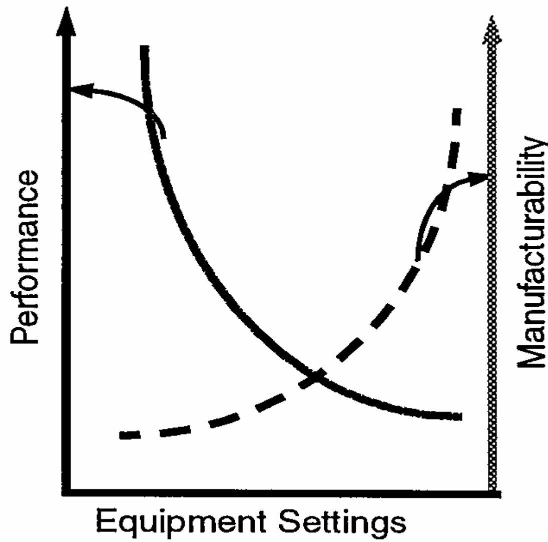
- Human errors, equipment failures
- Processing instabilities
- Material non-uniformities
- Substrate inhomogeneites
- Lithography spots

Variability causes *deformations*

- Geometrical
  - Lateral
  - Vertical
  - Spot defects
- Electrical
  - Global
  - Local

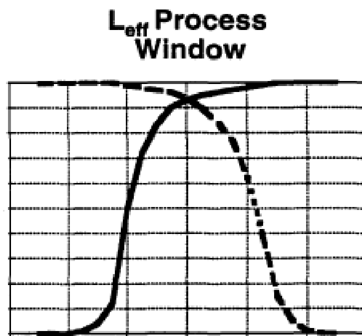
Deformations have *deterministic* and *random* components, are *global* and/or *local*, can be *independent* or can *interact*.

Performance vs Yield

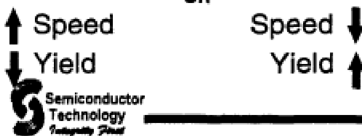


How does reduced variability help?

L_{eff} Process Window



<-- L_{eff} -->



Tighter lot CD distribution allowed target shift to shorter L_{eff} without yield hit

Speed Bin	% Pre-ACDC	% Post-ACDC
1	1.6	0.0
2	8.8	0.0
3	39.6	3.1
4	44.2	91.1
5	5.8	5.8
Mean ASP:	$ $	$ $

$ 2M / wk / 1K starts

MOTOROLA

600k APC investment, recovered in two days...

Design Yield Prediction

First we define functional (timing and other functionality constraints) and parametric (speed, power etc.) measures:

$$Q_F = \{q_1^F, q_2^F, \ldots, q_{n_F}^F\}^T$$
$$Q_P = \{q_1^P, q_2^P, \ldots, q_{n_F}^P\}^T$$

Then we define the space that contains both types of measures:

$$S_Q = S_Q^F \times S_Q^P$$

The set of acceptable IC performances is given as:

$$A_Q = \{Q \in S_Q \mid \textit{each } q_j \textit{ acceptable} \; \forall \; j=1,2,\ldots, n_Q\}$$
$$n_Q = n_F + n_P$$

Design Yield Prediction (cont.)

All performances are determined from the state variables of the process (geometric and electrical parameters after wafer fabrication):

$$A_W = \{X_w \in S_W \mid Q(X_w) \in A_Q\}$$

This is the acceptability region defined in the *state variable space.* The Yield is defined as:

$$Y = \int g(x_W)\, f_W(x_W)\, dx_W$$

$f_W(x_W)$ is the *jpdf* of X_W and $g(x_W) = \begin{cases} 1, \; \forall \; x_W \in A_W \\ 0, \; \textit{otherwise} \end{cases}$

$$Y = \int_{A_W} f_W(x_W)\, dx_W$$

Design Yield Prediction (cont.)

Let us now assume that *X w* really depends on the parameter sets defined as *C* (Controls), *L* (Layout), and *D* (disturbances). Given fixed values for *C* and *L*, then:

$$Y = \int_{A_D(C^0, L^0)} f_D(\delta)\, d\delta$$

Let us further assume that *C, L, D* are *separable* (i.e. some affect performance and some functionality, but none affects both). Then:

$$Y'_{FUN} = \int_{A_D^F(C^0, L^0)} f_D'(\delta')\, d\delta'$$

and

$$Y'_{PAR} = \int_{A_D^P(C^0, L^0)} f_D''(\delta'')\, d\delta''$$

Manufacturing Yield vs. Design Yield

$$Y_M = \frac{N_F}{N} = \frac{N_W}{N}\frac{N_F}{N_W}\frac{N_P}{N_F} = Y_W Y_{PT} Y_{FT}$$

Because of test coverage limitations:

$$Y_M \geq Y$$

Because of imperfect separation:

$$Y \leq Y_{PAR}' \cdot Y_{FUN}'$$

Finally, since we do not probe the actual circuit, we have:

$$Y_{PRO} \approx Y_{PAR}$$

Yield and Production Cost

The objective is not to maximize yield but to minimize cost.

Wafer Cost: $C_1 = (N - N_W)(c_P^* + c_W)$

Probe Cost: $C_2 = (N_W - N_P)(c_{PT} + c_P + c_W)$

Final Test Cost: $C_3 = N_P(c_{FT} + c_A + c_{PT} + c_P + c_W)$

Total Cost per chip sold: $C = C_1 + C_2 + C_3$

$$\frac{C}{N_F} = \frac{N_P}{N_F}(c_{FT} + c_A + c_{PT} + c_P + c_W) + \frac{N_W}{N_F}\left(1 - \frac{N_P}{N_W}\right)(c_{PT} + c_P + c_W) + \frac{N}{N_F}\left(1 - \frac{N_W}{N_F}\right)(c_P^* + c_W)$$

Yield and Production Cost (cont)

Finally, total cost per chip sold:

$$\frac{C}{N_F} = (c_{FT} + c_A + c_{PT} + c_P + c_W) + \frac{Y_P}{Y_M}(1 - Y_{FT})(c_{FT} + c_A + c_{PT} + c_P + c_W) + \frac{Y_W}{Y_M}(1 - Y_{PT})(c_{PT} + c_P + c_W) + \frac{1}{Y_M}(1 - Y_W)(c_P^* + c_W)$$

This means that the yield maximization problem and the cost minimization problem are not equivalent!

VLSI Design for Manufacturing: Yield Enhancement
Director, Maly and Strojwas
Kluwer Academic Publishers 1989

In Conclusion

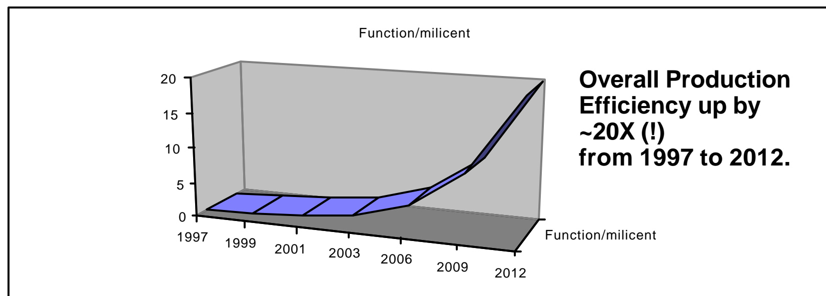
- Yield is good and Variability is bad.
  - metrology
  - statistical process control
  - run-to-run and real-time control
- Must manipulate process steps to accommodate circuits and designs.
  - modeling
  - design of experiments
- Must keep cost of manufacturing low.
  - Automation of product flow
  - Automation of information management

Calculating IC Cost vs Defect Density

An example...

The 1997 Roadmap (see transistor cost)

Year	1997	1999	2001	2003	2006	2009	2012
Feature nm	250	180	150	130	100	70	50
Area mm^2	300	340	385	430	520	620	750
Density cm^{-2}	3.7M	6.2M	10M	18M	39M	84M	180M
Cost mc/tr	**3000**	**1735**	**1000**	**580**	**255**	**110**	**50**
technology	248	248	193?	157?	14	14	14
wafer size	200	300	300	300	300	450	450



Negative Binomial (a widely accepted yield model)

If f(D) follows a Gamma distribution, then:

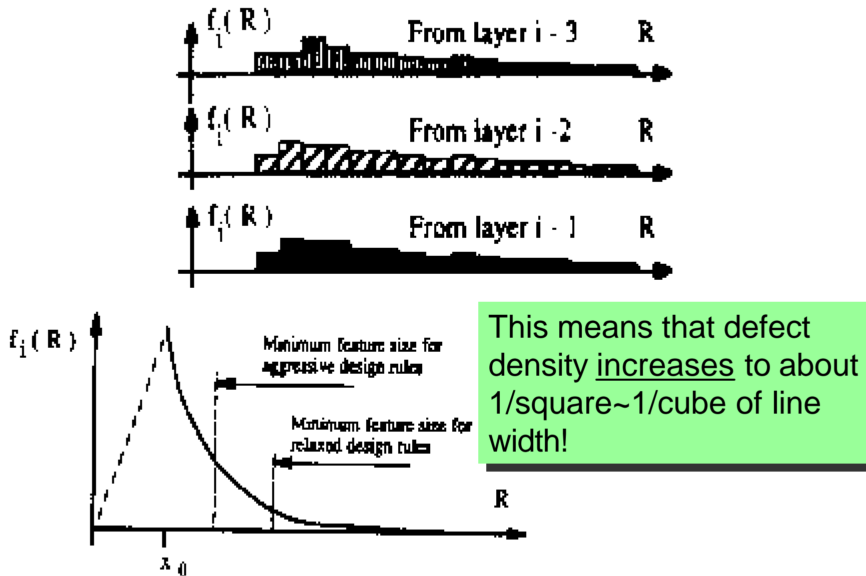
$$Y = \left[1 + \frac{A\,D}{\alpha} \right]^{-\alpha} \qquad (\alpha \sim 0.3 - 3)$$

And if clustering becomes an issue, then:

$$Y = Y_o \left[1 + \frac{A\,D}{\alpha} \right]^{-\alpha}$$

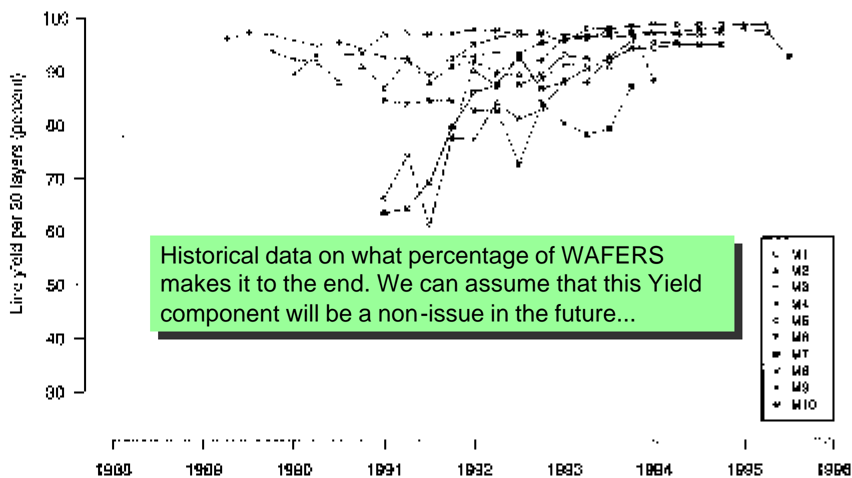
where Y_o is the “gross cluster yield”.

Typical Defect Size Distribution



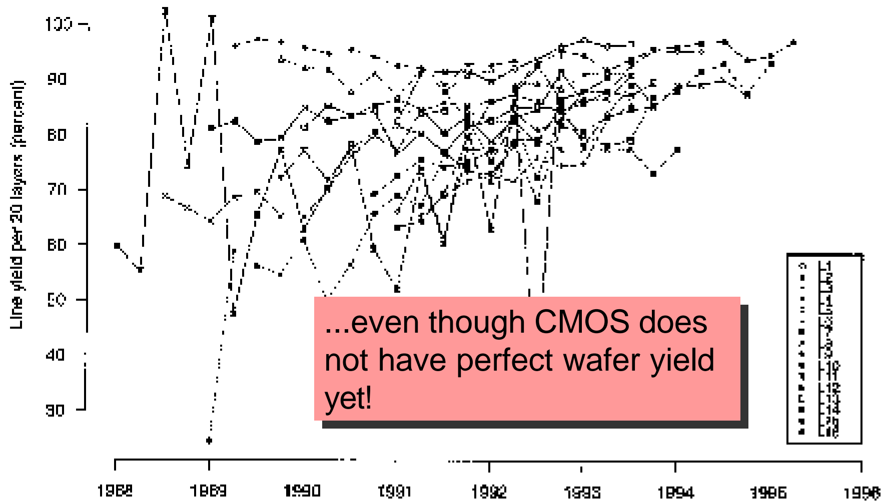
This means that defect density increases to about 1/square~1/cube of line width!

Line Yield, Memory

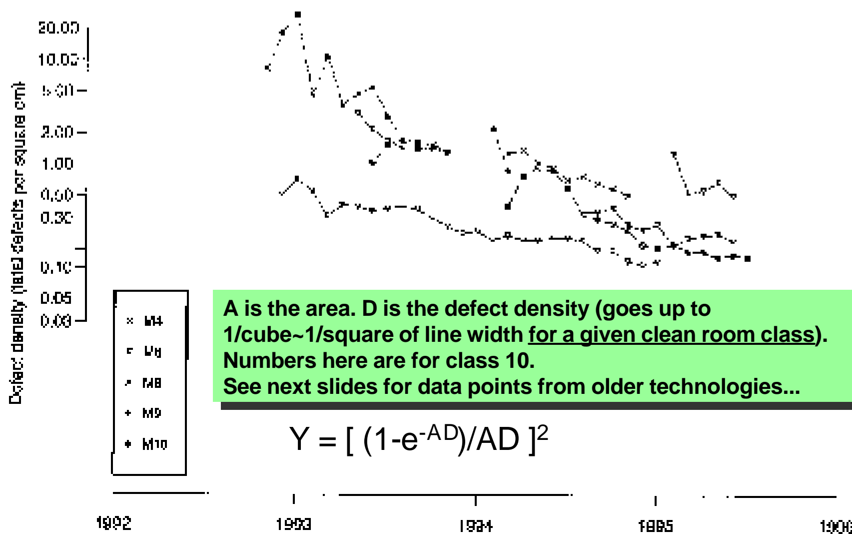


Historical data on what percentage of WAFERS makes it to the end. We can assume that this Yield component will be a non-issue in the future...

Line Yield, CMOS Logic

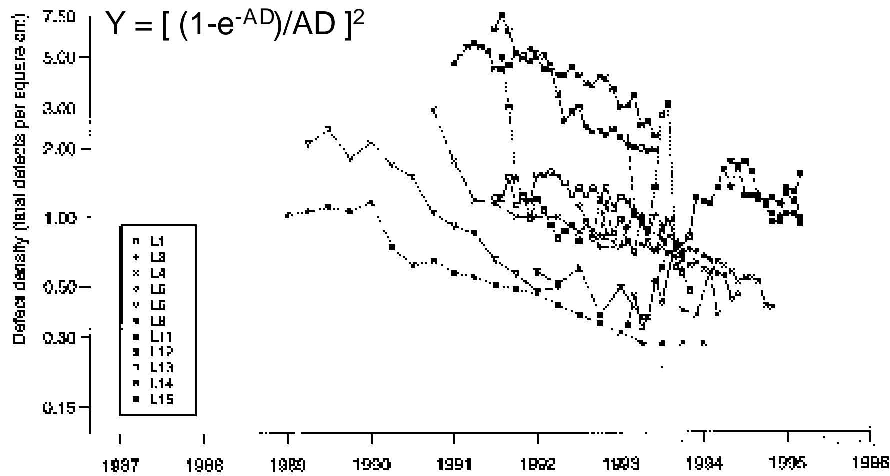


Memory Defect Density, 0.45-0.6μm

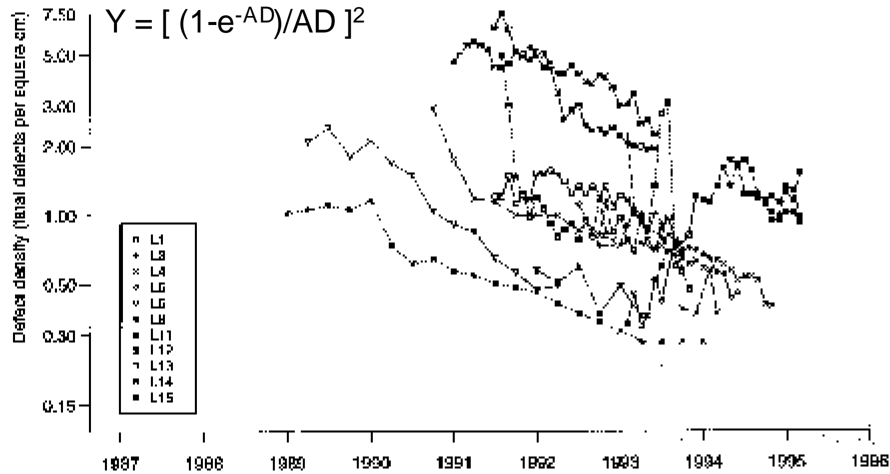


$$Y = [\ (1-e^{-AD})/AD\ ]^2$$

Logic Defect Density, 0.7-0.9μm CMOS



Logic Defect Density, 0.7-0.9μm CMOS



Bringing the Puzzle Pieces together...

Year	D	CD	Area	GD/W	Killer Def	Wafer Size	Density	Yield	Cent/Die	10^-6c/Tr
1997	0.000600	0.250	300.000	88.096	0.000600	200.000	3.700	0.842	11351	3068
1999	0.000500	0.180	340.000	137.856	0.001340	300.000	6.200	0.663	10881	1755
2001	0.000140	0.150	385.000	145.052	0.000648	300.000	10.000	0.790	10341	1034
2003	0.000050	0.130	430.000	141.792	0.000356	300.000	18.000	0.863	10579	588
2006	0.000002	0.100	520.000	133.684	0.000031	300.000	39.000	0.984	11220	288
2009	0.000004	0.070	620.000	229.707	0.000182	450.000	84.000	0.896	9795	117
2012	0.000001	0.050	750.000	193.394	0.000125	450.000	180.000	0.912	11634	65

Note how clean we must be!
Note trade-offs between wafer size and defects!
20x improvement will not be easy...

Fall 2003 EE290H Tentative Weekly Schedule

IC Yield & Performance

1. Functional Yield of ICs and DFM.
2. Parametric Yield of ICs.
3. Yield Learning and Equipment Utilization.

Process Modeling

4. Statistical Estimation and Hypothesis Testing. ←
5. Analysis of Variance.
6. Two-level factorials and Fractional factorial Experiments.

Process Control

7. System Identification.
8. Parameter Estimation.
9. Statistical Process Control. ***Distribution of projects. (week 9)***
10. Run-to-run control.
11. Real-time control. ***Quiz on Yield, Modeling and Control (week 12)***

Metrology

12. Off-line metrology - CD-SEM, Ellipsometry, Scatterometry
13. In-situ metrology - temperature, reflectometry, spectroscopy

Manufacturing Enterprise

14. The Computer-Integrated Manufacturing Infrastructure

15. *Presentations of project results.*