Modeling of Defect-Limited Yield in Semiconductor Manufacturing

Emmanuel Yashchin
IBM Research Division

(joint work with M. Baron, A. Takken and M. Lanzerotti)

Goals

- Stochastic model explaining chip failures by defects
- Use this model to
  - forecast yield
  - evaluate “kill ratios”, effects of each defect type
  - determine the most dangerous *“layer-type-size-location-…”* combinations
  - determine ways to enhance the yield

Modeling

A defect occurred on some chip.

$$P\{chip\ survives\} = e^{-r(j)a(l)g(X)}$$

or $$P\{chip\ survives\} = e^{-\{a(l)+r(j)\}g(X)}$$

or $$P\{chip\ survives\} = e^{-a(l)+r(j)g(X)}$$

where l=level, j=defect type, X=defect size

If all defects were observed and classified...

$$\phi_c = P\{survival\} = e^{-b(m)} \prod_{\substack{Defects \\ (k,c,w,m)}} e^{-a(l_k)r(j_k)g(X_k)}$$

where

$$\left.\begin{aligned} \mathbf{a} &= a(1),\ldots,a(L) = level\ \ effects \\ \mathbf{r} &= r(1),\ldots,r(J) = defect\ type\ effects \\ \mathbf{b} &= b(1),\ldots,b(M) = effects\ of\ \ other\ causes \end{aligned}\right\} Parameters$$

$$L(\mathbf{a},\mathbf{b},\mathbf{r}) = \prod_{\substack{Lots \\ (m)}} \prod_{\substack{Wafers \\ (w,m)}} \prod_{\substack{Chips \\ (c,w,m)}} \phi^{\xi_c}(1-\phi)^{1-\xi_c},$$

But… most defects are unclassified

$$P\{\text{chip survives an unclassified defect } k\}$$

$$= \sum_{j=1}^{J} P\{j_k = j\} P\{\text{chip survives a classified defect } k\}$$

$$= \sum_{j=1}^{J} \left(\frac{\lambda_{j,l_k,m_k}}{\lambda_{l_k,m_k}} \right) e^{-a(l_k) r(j) g(X_k)}$$

$\lambda_{j,l,m}$ = frequency, expected number of type j defects per chip on level l of lot m.

Including defect sizes

$$P\left\{\text{type } j \mid size \text{ X}\right\} = \frac{P\{j\}P\{X \mid j\}}{\sum P\{j'\}P\{X \mid j'\}}$$

where

$$P\left\{size \text{ X} \mid \text{type } j\right\} = \pi_j(X)$$

the distribution of defect sizes (by defect type)

Including defect frequencies

$$L(\mathbf{a},\mathbf{b},\mathbf{r},\lambda) = P\{defects\}\,P\{types \mid defects\}\prod \phi^{\xi}(1-\phi)^{1-\xi}$$

where

$$P\{observed\ defects\} = \prod_{layers(l)} e^{-c_l\lambda_l}\frac{(c_l\lambda_l)^{N_l}}{N_l\,!}$$

$$P\{classified\ types \mid defects\} = \prod_{layers(l)}\prod_{types(j)}\left(\frac{\lambda_{jl}}{\lambda_l}\right)^{d_{jl}}$$

But… only a few levels get inspected

$$P\left\{\text{chip survives an uninspected level } l\right\}$$

$$= \prod_j P\left\{\text{all type } j \text{ defects are not killers}\right\}$$

$$= \prod_j E^{N(c,j,l)}\left\{\prod_{k=1}^{N(c,j,l)} E^{X_k} e^{-a(l)r(j)g(X_k)}\right\}$$

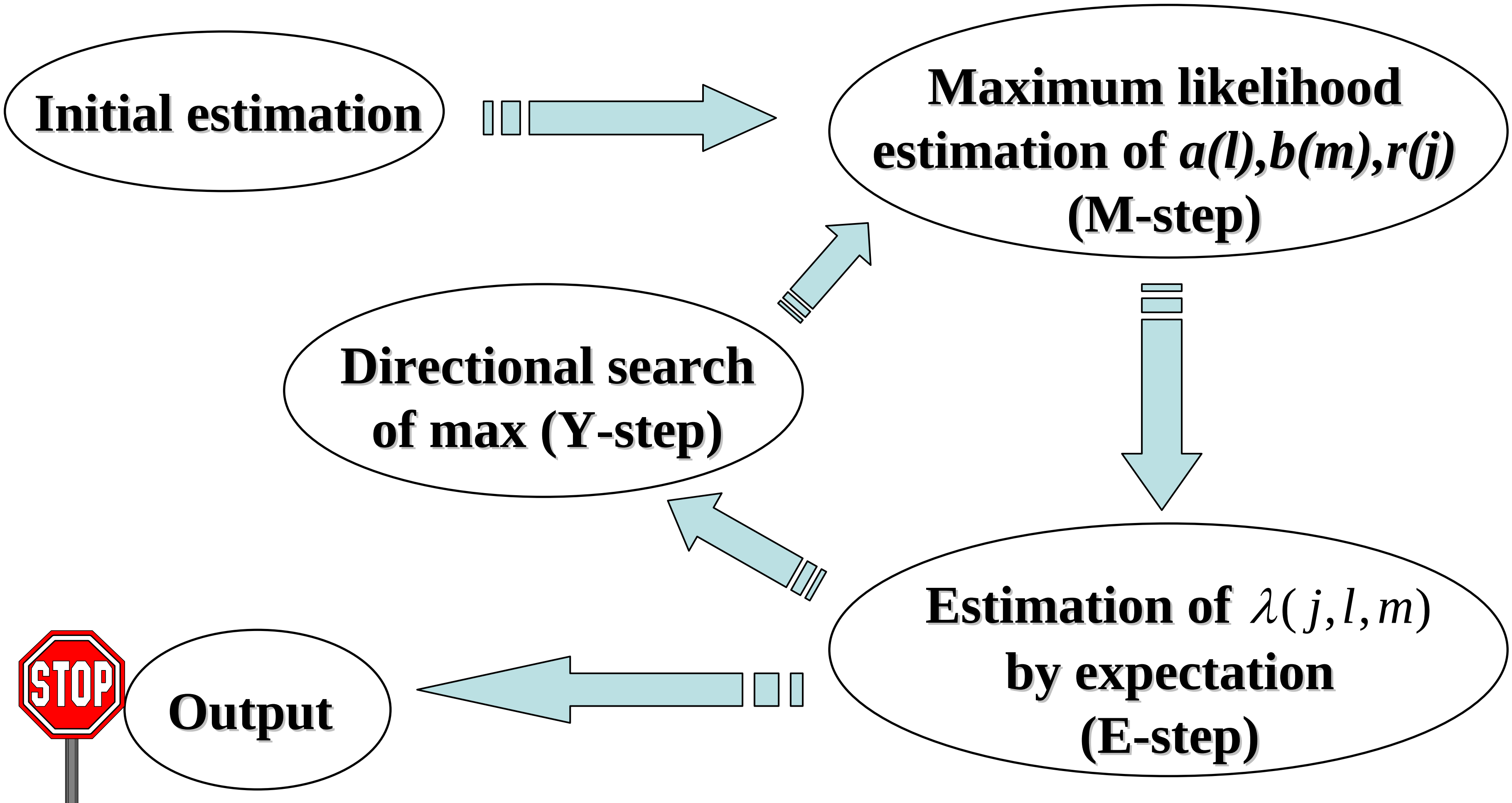
$$= \exp\left\{\sum_j \lambda_{jlm}\left(1 - E_j^X e^{-a(l)r(j)g(X)}\right)\right\}$$

Putting together all the covariates:

- classified defects
- unclassified defects
- uninspected layers
- other causes
- other terms related to data collection

$$L(parameters \mid data) = ...$$

EM algorithm (modified in many ways)



Estimation

E-step:

Take *E(missing data | observed data)*

$$\lambda_{jlm}^{[new]} = E\{\#of\ (j,l)\text{-}defects\ per\ chip\}$$

$$= \frac{\left(\begin{array}{l} \#classified + (\#unclassified)E\{proportion\ of\ j \mid X, \xi\} \\ +E\{\#defects\ of\ type\ j\ on\ uninspected\ layers \mid \xi\} \end{array} \right)}{\#chips}$$

Applications

1. Yield forecast

$$\varphi = P\{\text{good chip}\} = E\{\ \#\ \text{good chips} \mid 1\ \text{chip}\ \}$$

Hence,

$$\sum(\varphi) = E\{\ \#\ \text{good chips} \mid \text{all chips}\ \} = E\{\ \text{yield}\ \}$$

Goodness of fit: on a training set,

- compare *E(yield)* with the actual yield
- compare *E(yield)* among good and among bad chips

How good can it get?

Ideally, should we predict
Yield=100% on good chips
Yield=0% on bad chips ?

No, even under the ***ideal*** fit (when $\{\varphi_i\}$ are known)

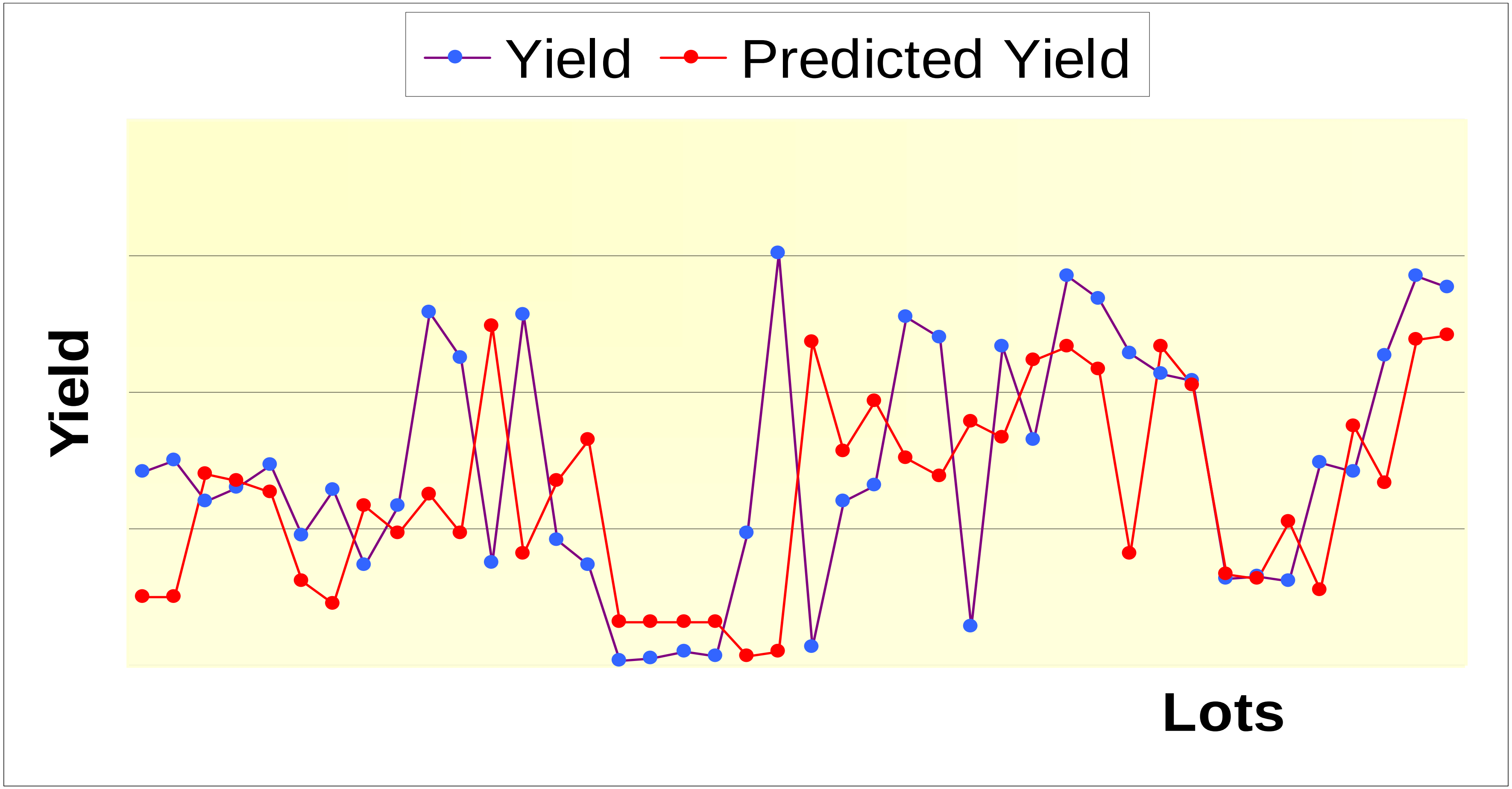
Assuming φ_i are iid from F,

$$\text{Predicted Yield (good chips)} \rightarrow E_F\varphi^2 / E_F\varphi$$

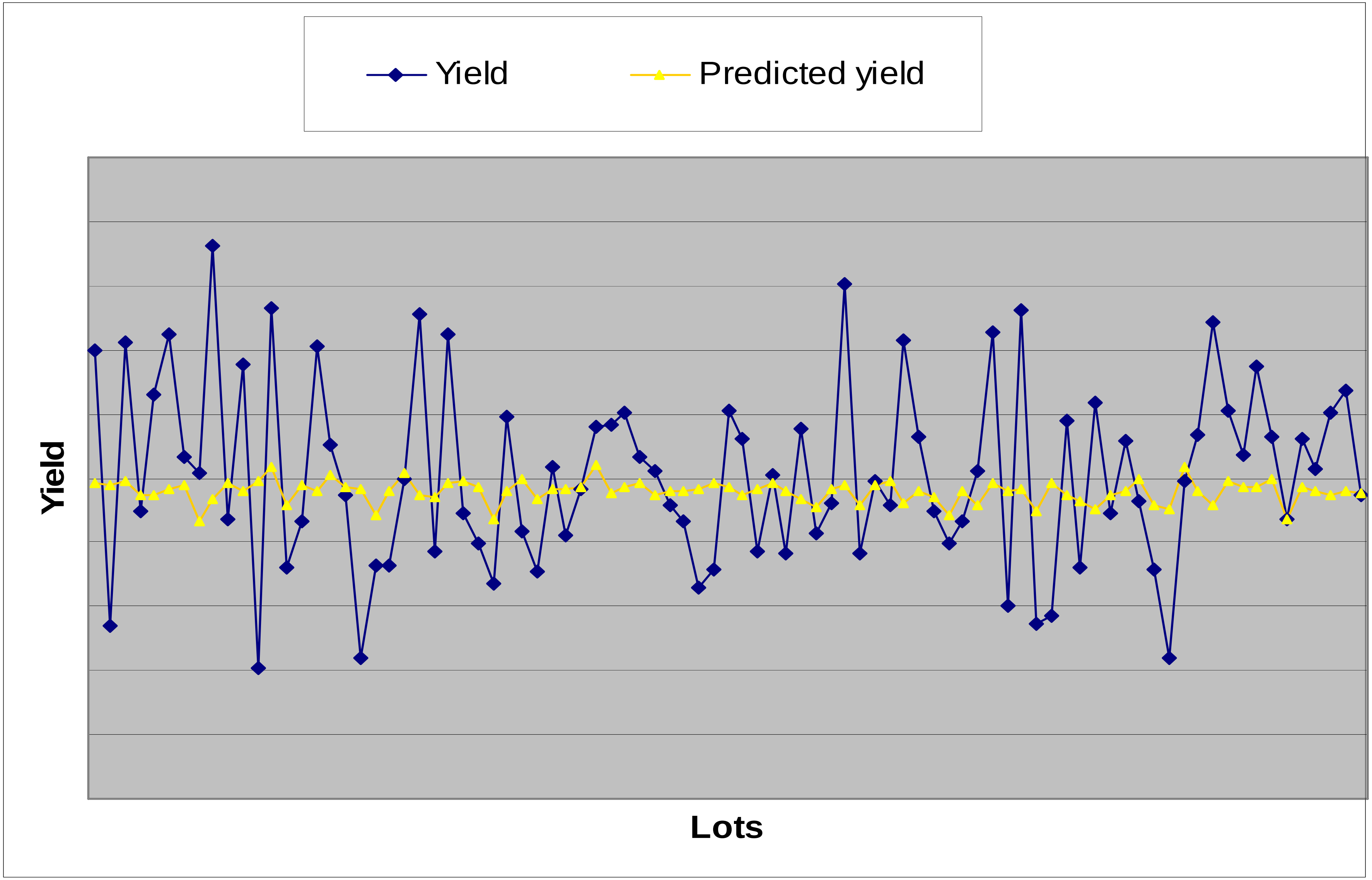
$$\text{Predicted Yield (bad chips)} \rightarrow E_F\varphi(1-\varphi) / (1-E_F\varphi)$$

$$Difference \rightarrow \frac{Var_F\varphi}{E_F\varphi(1-E_F\varphi)}$$

Example 1: many levels inspected



Example 2: few levels inspected



Applications

2. Determining leading causes of failures

- **Effect of each factor**

$$P\{\text{defect of type } j \text{ on level } l \text{ is fatal}\} = 1 - E^X e^{r(j)a(l)X}$$

(kill ratio)

$$P\{\text{defect of type } j \text{ is fatal}\} = 1 - E^X E^{l|X} e^{r(j)a(l)X}$$

$$P\{\text{defect on level } l \text{ is fatal}\} = 1 - E^X E^{j|X} e^{r(j)a(l)X}$$

- **Causes of failures**

$$P\{\text{a defect of type } j \text{ on level } l \text{ was fatal} \mid \text{chip failed}\}$$

$$P\{\text{a defect of type } j \text{ was fatal} \mid \text{chip failed}\}$$

$$P\{\text{a defect on level } l \text{ was fatal} \mid \text{chip failed}\}$$

Also compute

$$P\{\text{a defect of type } j \text{ and nothing else was fatal} \mid \text{chip failed}\}$$

etc.

3. Leading causes of failures on current data

- **Contribution of each level = yield differential it makes**

$$= \sum(\varphi) - \sum(\varphi \mid \text{ level } l \text{ is uninspected})$$

or

$$\frac{\sum(\varphi) - \sum(\varphi \mid \text{ level } l \text{ is uninspected})}{\sum(\varphi)} \cdot 100\%$$

4. Sensitivity to changes

If the frequency λ_{jl} changes by $x\%$, then the yield changes by $y\%$,

$$y = \left(\frac{\partial \log \varphi}{\partial \log \lambda_{jl}} \right) x = \left(-\lambda_{jl} [1 - E^X e^{r(j)a(l)X}] \right) x$$

Similarly: elasticities for λ_j, λ_l, and defect sizes

Result - yield improvement strategies