01 Jan 1996

Applications of the Upside-Down Normal Loss Function

David Drain
Missouri University of Science and Technology

A. M. Gough

Correspondence

Applications of the Upside-Down Normal Loss Function

David Drain and Andrew M. Gough

Abstract— **The Upside-Down Normal Loss Function (UDNLF) is a weighted loss function that has accurately modeled losses in a product engineering context. The function's scale parameter can be adjusted to account for the actual percentage of material failing to work at specification limits. Use of the function along with process history allows the prediction of *expected loss*—the average loss one would expect over a long period of stable process operation. Theory has been developed for the multivariate loss function (MUDNLF), which can be applied to optimize a process with many parameters—a situation in which engineering intuition is often ineffective. Computational formulae are presented for expected loss given normally distributed process parameters (correlated or uncorrelated), both in the univariate and multivariate cases.**

I. Introduction

Loss functions quantify the relationship between process performance and manufacturing yield [1]–[8]. When applied along with knowledge of process variation and unit production costs, one can derive the manufacturer's economic cost of process variation. This paper illustrates applications of the upside-down normal loss function (UDNLF)—a loss function we have found to accurately model losses in a real manufacturing situation, and which has desirable mathematical properties enabling easy prediction of average losses due to typical manufacturing variation.

II. The UDNLF

A. The UDNLF Defined

The UDNLF is one minus a scaled normal probability density function, with mean τ and variance λ^2, defined by the following formula

$$L_{\mathrm{UDN}}(x \mid \tau, \lambda) = 1 - e^{-\frac{(x-\tau)^2}{2\lambda^2}} \tag{1}$$

where

$$\tau = \text{process parameter target}$$
$$\lambda = \text{scale factor.}$$

The UDNLF is zero at the target, and asymptotically approaches one. It thus avoids a disadvantage of quadratic loss functions: unrealistic values far from the target.

The scale factor adjusts the penalty for deviation from the target: a large λ indicates that the process can tolerate relatively greater deviation from the target. λ can be empirically determined, or it can be set to some predetermined fraction of the parameter specification range. Lacking better information, a pragmatic choice is to set λ to 42.5% of the specification range. In this case, the loss when a process parameter is at a specification limit is about 50% (corresponding to step function loss in the same situation).

Manuscript received October 14, 1993; revised April 6, 1995, and July 24, 1995.

The authors are with the Intel Corporation, Chandler, AZ 85226 USA.

Publisher Item Identifier S 0894-6507(96)01142-6.

B. Example: Loss Due to Equipment Variation from Target

The etch rate of a polysilicon etcher has a target of 25 Å/s and specification limits of 22 and 28 Å/sec. Analysis of historical data established that about 50% of etched die will fail when the process is centered at either specification limit; the scale factor λ is therefore chosen as 0.425 times the specification range of 6 Å, so $\lambda = 2.55$.

C. Example: Symmetric Fit to Yield Data

A new microcontroller product exhibited low yield at hot temperature in its initial manufacturing runs. Examination of the failing die uncovered a speed path in a subcircuit of the device that would cause functional failures if slow transistors were manufactured. The length (L) of the MOSFETs' polysilicon gates was suspected to have the greatest effect on transistor speed; this was verified in an experiment which allowed polysilicon CD's to vary about their target (1.60 μm) from 1.15 μm to 1.90 μm.

A nonlinear regression model was used to fit a UDNLF to the resulting losses

$$L_{\mathrm{UDN}}(x \mid \tau, \lambda) = L_{\mathrm{Min}} + (Y_{\mathrm{Max}} - L_{\mathrm{Min}})\left[1 - e^{-\frac{(x-\tau)^2}{2\lambda^2}}\right] + \epsilon \tag{2}$$

where

$$x = \text{poly CD variable } (\mu\text{m})$$
$$\tau = \text{Process target poly CD } (\mu\text{m})$$
$$\lambda = \text{fitted shape parameter } (\mu\text{m})$$
$$L_{\mathrm{Min}} = \text{minimum expected loss (constant, die/wafer)}$$
$$Y_{\mathrm{Max}} = \text{maximum possible product wafer yield (constant, die/wafer)}$$
$$\epsilon = \text{error term of regression.}$$

τ was fixed at the process target and λ was fit by the regression.

For the microcontroller data, the function

$$L_{\mathrm{UDN}}(x \mid 1.60, \lambda) = 38 + (538 - 38)\left[1 - e^{-\frac{(x-1.60)^2}{2\lambda^2}}\right] \tag{3}$$

was fit using nonlinear regression software. The regression determined the best fit with $\lambda = 0.1851$, resulting in the function

$$L_{\mathrm{UDN}}(x \mid 1.60, 0.1851) = 38 + (538 - 38)\left[1 - e^{-\frac{(x-1.60)^2}{2(0.1851)^2}}\right]. \tag{4}$$

This function is shown graphically in Fig. 1.

D. Expected Loss

Expected loss is the average loss one would observe from a stable process over a long period of operation, so it can be a critical piece of information when making process change decisions affecting process targets or variability.

To compute expected loss, one also needs the probability density function of the manufacturing variable (x) under study. Expected loss is then computed by evaluating the integral of the product of the loss function and the probability density function.

If the actual process parameter distribution and a realistic loss function are given, expected loss can be determined by numerical

0894–6507/96$05.00 © 1996 IEEE

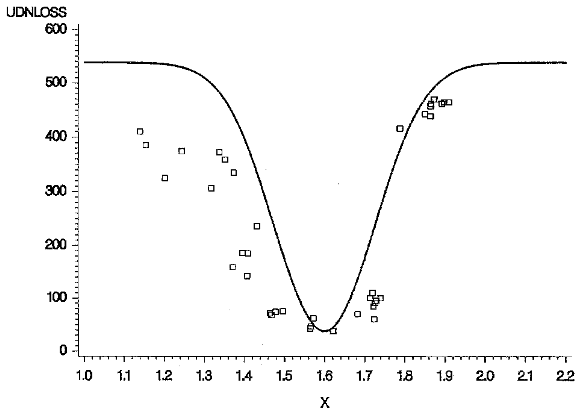

Fig. 1. Symmetric UDNLF fit to yield loss data with $\lambda = 0.1851$. Note how the loss function gradually approaches the constant value of 538, which is the number of die printed on each wafer.

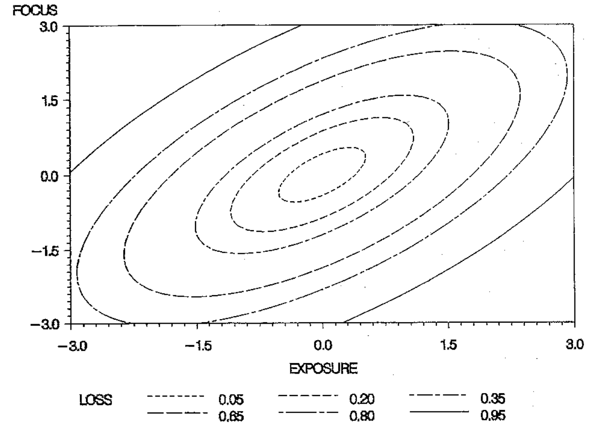

Fig. 2. Contour plot of multivariate upside down normal loss function (MUDNLF) with synergistic process parameters.

integration. However, it is usually reasonable to assume the process parameter is normally distributed with mean μ and standard deviation σ, and with this simplifying assumption, expected loss can be determined analytically as follows

$$\mathrm{EL}_{\mathrm{UDN}}(\mu, \sigma, \tau, \lambda) = 1 - \frac{\lambda}{\sqrt{\sigma^2 + \lambda^2}} e^{-\left(\frac{(\mu-\tau)^2}{2(\sigma^2+\lambda^2)}\right)}. \quad (5)$$

This formula can predict the loss due to typical manufacturing variation, assess the damage caused by a drift from the target, estimate losses due to an increase in variance, and quantify the economic consequences of process changes.

Example: In the case of the microcontroller above, substituting the scale factor and Poly CD target (1.60) and standard deviation (0.0835) into (5) results in an expected loss of 0.0885. This value is then transformed with the same linear transformation used in fitting the UDNLF to the yield data, resulting in an expected loss of 82.2 die/wafer. These results agree with actual losses, which had a median value of 81 die/wafer over one quarter's production.

III. The Multivariate UDNLF

The UDNLF can also be applied to processes with more than one important process parameter.

A. The Multivariate Upside-Down Normal Loss Function Defined

The Multivariate Upside-Down Normal Loss Function (MUDNLF) for n parameters is defined as follows

$$L_{\mathrm{MUDN}}(x \mid \tau, L) = 1 - e^{-\frac{1}{2}(x-\tau)^T L^{-1}(x-\tau)} \quad (6)$$

where $\mathbf{x}$ and τ are $n \times 1$ column vectors, and $\mathbf{L}$ is an $n \times n$ scaling matrix relating deviation from target to loss for all n parameters. As defined here, $\mathbf{L}$ must be *symmetric and positive definite*: $\mathbf{L}$ is symmetric if it remains the same when its rows and columns are interchanged ($\mathbf{L} = \mathbf{L}^T$); $\mathbf{L}$ is positive definite if $\mathbf{x}^T\mathbf{L}\mathbf{x} > 0$ for all nonzero column vectors $\mathbf{x}$. These requirements may not actually be necessary for the definition of a reasonable loss function; they were chosen because they give the function desirable mathematical properties.

Off-diagonal elements of $\mathbf{L}$ are used to account for interaction effects of process parameters—those cases where the deviation of two factors simultaneously produces a different effect than would be expected from the individual factor effects alone.

The elements of $\mathbf{L}$ are easily interpreted when written in a form similar to the covariance matrix of the multivariate normal distribution

$$\begin{bmatrix} \lambda_1^2 & \lambda_1\lambda_2\xi \\ \lambda_1\lambda_2\xi & \lambda_2^2 \end{bmatrix}. \quad (7)$$

Parameters on the diagonal scale losses due to individual parameters; parameters off the diagonal indicate synergy or antagonism of loss when both parameters vary from target. Positive off-diagonal elements (ξ positive) mean that losses are lower than would be expected when the two factors vary simultaneously in the same direction.

As an example, consider a hypothetical lithography process in which focus and exposure are represented by the MUDNLF with $\mathbf{L}$ as follows

$$L = \begin{bmatrix} 2.89 & 1.8023 \\ 1.8023 & 2.6602 \end{bmatrix} = \begin{bmatrix} \lambda_1^2 & \lambda_1\lambda_2\xi \\ \lambda_1\lambda_2\xi & \lambda_2^2 \end{bmatrix}. \quad (8)$$

This choice of $\mathbf{L}$ corresponds to a case where the loss due to variation of focus (alone) at the specification limits is 50%. Specification limits for focus are ± 2.0 μm, so

$$\lambda_1 = 0.425 \cdot 4 = 1.70.$$

The similar loss for exposure is 90% at the specification limits of ± 3.5 mJ, so

$$\lambda_2 = 0.233 \cdot 7 = 1.6310.$$

The choice of $\xi = 0.65$ indicates that losses are less when the two parameters vary in the same direction than when they vary in different directions. This MUDNLF takes the shape shown in the contour plot of Fig. 2.

For a real process, the loss definition matrix could be based on empirical observations, experiments, or simulations.

B. The Multivariate Normal Distribution

As in the one-parameter case, it is often reasonable to assume process parameters have a normal distribution; however, since multiple parameters are involved, one must apply the multivariate normal distribution. The multivariate normal probability density function has the following form

$$f(x) = \frac{e^{-\frac{1}{2}(x-\mu)^T M^{-1}(x-\mu)}}{(2\pi)^{\frac{n}{2}} |M|^{\frac{1}{2}}} \quad (9)$$

where

$\mathbf{x}$ = a $n \times 1$ column vector of variables

$\mathbf{M}$ = a (positive definite, symmetric) covariance matrix

μ = an $n \times 1$ column vector of means

n = the number of variables

C. Expected Loss With MUDNLF

The expected loss from a MUDNLF defined by $\mathbf{L}$, and a multivariate normal process parameter distribution defined by μ and $\mathbf{M}$, with target τ, is given by

$$\begin{aligned}\mathrm{EL}_{\mathrm{MUDN}}(\mu, M, \tau, L) &= 1 - \frac{|(L^{-1}+M^{-1})|^{-\frac{1}{2}}}{|M|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mu^T M^{-1}\mu + \tau^T L^{-1}\tau)} \\ &\quad \times e^{\frac{1}{2}(\mu^T M^{-1} + \tau^T L^{-1})(L^{-1}+M^{-1})^{-1}(M^{-1}\mu + L^{-1}\tau)}. \qquad (10)\end{aligned}$$

This closed-form solution for expected loss has even greater utility in the multivariate case than it does for univariate loss functions because engineering intuition is often ineffective for multivariate problems.

IV. Conclusion

We found that actual losses can be predicted by the UDNLF. The loss function is easily adaptable to multivariate cases, even when process variables are correlated and losses are the result of synergy or antagonism between those variables. Loss functions can be used in process design and optimization by aligning losses with process parameter distributions in a way which minimizes expected loss.

Acknowledgment

The authors thank to John Ramberg of the University of Arizona for planting the seed for this work, and to David Martinich, Russell Miller, Matthew Ploor, Bob Sankman, and Ann Tiao for reviewing the manuscript.

References

[1] H. Tsai, "Probabilistic tolerance design for a subsystem under burr distribution using taguchi's loss functions," *Commun. Statistics—Theory and Methods*, vol. 19, no. 12, pp, 4679–4696, 1990.

[2] G. Taguchi, *System of Experimental Design*, vols. 1 and 2. White Plains, NY: Kraus International, 1987.

[3] R. V. Leon, A. C. Shoemaker, and R. N. Kacker, "Performance measures independent of adjustment," *Technometrics*, vol. 29, No. 3, pp. 253–265, Aug. 1987.

[4] S. Maghsoodloo, "The exact relation of taguchi's signal-to-noise ratio to his quality loss function," *J. Quality Technol.*, vol. 22, no. 1, pp. 57–67, Jan. 1990.

[5] J. Turner, "Is an out-of-spec product really out of spec?," *Quality Progress*, pp. 57–59, Dec. 1990.

[6] Y. Hosomatsu, "Asymmetric loss function and optimal policy under uncertainty: A simple proof," *Decision Sci.*, vol. 11, pp. 577–585, 1980.

[7] C. Mohebbi and L. Hayre, "Multivariate control charts: A loss function approach," *Sequential Anal.*, vol. 8, no. 3, pp. 253–268, 1989.

[8] L. Weiss, "Estimation with a gaussian gain function," *Statistics & Decisions*, supp. issue no. 1, pp. 47–59, 1984.

Minimum Inventory Variability Schedule With Applications in Semiconductor Fabrication

Shu Li, Tom Tang, and Donald W. Collins

***Abstract*— A typical semiconductor wafer fab contains many different products and processes, some with small quantities, competing for resources. Each product flow can contain hundreds of processing steps demanding production time of the same resource many times during the flow. When this re-entry requirement is compounded with multiple product flows, short interval scheduling becomes important. Scheduling to reduce variations and to balance the whole wafer production line becomes a very complex issue.**

We investigate in this paper a new scheduling policy called minimum inventory variability scheduling (MIVS). This scheduling policy can significantly reduce the mean and variance of cycle-time in semiconductor fabs.

The conclusions are based on the real world implementation in two major semiconductor fabs since 1990, and a simulation study of a much simplified hypothetical re-entrant network to capture the nature of semiconductor manufacturing. A discrete event simulation model was used to compare MIVS with five different popular dispatching policies (FIFO, SNQ, LNQ, RAN, and CYC) practiced in wafer fabrication environments. The results gained on two factory floors and the simulation model indicate that dispatching policies have a significant impact on performance. The simulation results show that the MIVS dispatching policy demonstrated a percentage improvement over all other tested dispatching policies.

I. Scheduling in Semiconductor Fabrication

A typical semiconductor fabrication flow for a single product is a highly re-entrant process. Each product flow requires the same equipment resource (e.g., photolithography, resist clean, diffusion, LPCVD, ion implant, sputtering, CVD, and PECVD) many times before completion of its production cycle. Furthermore, the problem is much more severe in pilot production lines, as many different flows and technologies (e.g., CMOS, BiCMOS, Bipolar, TMOS, smartMOS, RF, and semiconductor sensors) with small quantities compound the resource sharing problems. In fact, many factories use simple scheduling rules such as FIFO or due date first scheduling disciplines. However, experience and some common sense may disagree with these traditional scheduling rules. For example, a machine (M_1) supplies two types of jobs (J_2 and J_3) to two different downstream machines (M_2 and M_3), respectively. Job J_2 is processed by machines M_1 and M_2 and job J_3 is processed by machines M_1 and M_3. Suppose M_1 has both J_2 and J_3 available to choose from at a given time. J_2 is ahead of J_3 in the queue. But at this particular time, M_2 is down and has an inventory of J_2's in its queue. On the other hand, M_3 is available. It makes sense to run J_3 first. It is exactly this observation in real fabs that motivated us to investigate different scheduling policies.

The result of this paper was first applied in a large semiconductor fab in 1990, and subsequently applied to two R&D pilot wafer fabs in 1992. This paper is a more detailed treatment of early publications by Li [5], [6], and Tang [11]. The materials were also covered by Li

Manuscript received February 25, 1994; revised August 15, 1995.

S. Li is with Allied Signal Aerospace Company, Tempe, AZ 85284 USA.

T. Tang is with United Parcel Service International Inc., Taipei 111, Taiwan, R.O.C.

D. W. Collins is with the College of Engineering and Applied Sciences, Department of Manufacturing and Industrial Technology, Arizona State University, Tempe, AZ 85287 USA.

Publisher Item Identifier S 0894-6507(96)01141-4.