行政院國家科學委員會專題研究計畫　期中進度報告

子計畫五：多步程製造系統演進最佳化方法及應用於晶圓廠良率提昇之研究(1/3)

計畫類別： 整合型計畫
計畫編號： NSC92-2213-E-002-100-
執行期間： 92 年 08 月 01 日至 93 年 07 月 31 日
執行單位： 國立臺灣大學工業工程學研究所

計畫主持人： 陳正剛
共同主持人： 林佑昇
計畫參與人員： 范治民 鄭伊婷 鄭佳惠

報告類型： 精簡報告

處理方式： 本計畫可公開查詢

中　華　民　國 93 年 5 月 25 日

行政院國家科學委員會多年期專題研究計畫期中報告

系統最佳化方法研究及其於半導體製造之應用子計畫一
多步程製造系統演進最佳化方法及應用於晶圓廠良率提昇之研究（1/3）
Evolutionary Process Optimization Methods for Multi-stage Manufacturing Systems with Applications to Semiconductor Yield Ramp-up

計畫編號：NSC 90-2218-E-002-046
執行期限：92 年 8 月 1 日至 93 年 7 月 31 日
主持人：陳正剛
(e-mail: achen@ntu.edu.tw)
共同主持人：林佑昇 協同主持人：范治民
執行機構及單位名稱：國立台灣大學工業工程學研究所

中文摘要

為了可以在生產過程中即時而且持續的改善生產製程，許多不同的技術已在過往陸續地被發展出來，這些技術分別由不同學術領域的學者基於不同需求所發展出來；然而遭到實務界的忽視卻是這些技術的共同的命運。一個重要的理由是，這些技術往往需要複雜的數學理論與計算，因為這些複雜的計算使得在計算機資源仍然相當昂貴的年代，無法被應用於實際的問題上。另外一個理由是，過去較簡易的企業經營模式與製造過程不需要如此複雜的技術。在本計劃的第一年度中，我們重新檢視這些技術，其中包括統計學者所發展出的演進作業程序技術與最佳化實驗設計；由數學及控制學者所發展的擾動最佳化技術；以及最近由離散事件控制學者所發展的排序最佳化技術。我們也經由對這些技術的徹底了解，而發展並提出一適用於實務界的整合演進最佳化方法。

關鍵詞：演進作業程序 最佳化實驗設計 擾動最佳化 排序最佳化

Abstract

Over the years, different techniques have been developed to address issues of continuous process improvement during production. These techniques are developed in different academic communities initiated by different concerns. However, a common fate of these techniques is the negligence by industrial practitioners. An important reason was that the techniques' unbearable mathematical sophistication had diminished their applicability during the age of expensive computing resources. Another reason was that these techniques seemed to over-shoot the needs of the relatively simple business model and manufacturing processes at that time. In the first year of this project, we re-examine these techniques, including Evolution Operations (EVOP) and Optimum Experimental Design (OED) and Ridge Analysis (RA) developed by the applied statistics community, Perturbation-based Real-Time Optimization (PRTO) by mathematics and control communities and Ordinal Optimization (OO) techniques by the discrete-event control community. After reviewing these methods, we propose an integrated, comprehensive evolutionary optimization method that can be easily applied in practice.

Keywords：Evolutionary Operations (EVOP) Planning, Optimum Experimental Design (OED), Perturbation Optimization (PO), Ordinal Optimization (OO)

1. Introduction

Fast yield ramp-up is a critical factor to shorten the time to market and thus raise the product values. Under the concurrent engineering framework, product R&D activities are aligned with manufacturability concerns in order to ensure the robust product quality and speed up the time-to-market process. Due to the increasing complexity of R&D and manufacturing processes for modern hi-tech products, such as semiconductor products and electronic appliances, a recent trend is to divide the product values roughly into three major elements: product design, product manufacturing and product service. Instead of having a single company responsible for the entire product life cycle, different companies are now in charge at different stages of the life cycle. Taiwan is a de facto manufacturing center where companies' core competences are mainly low manufacturing costs and flexible production systems. In fact, the modern manufacturing processes have become so complicated that in addition to further lowering the manufacturing cost, companies have increasingly focused their efforts on raising the product values by products' early introduction to the market. To achieve this, yield ramping up during the manufacturing stage becomes extremely important.

As mentioned, concepts of manufacturing systems have evolved over the years from divided functions within a company to integrating a company's efforts and resources and now to a new business model with companies focusing on different aspects of a product. The birth of this new business model is due largely to the very different characteristics of modern hi-tech products. Unlike conventional commodities, hi-tech products usually have a very short life cycle. To speed up the time-to-market process, companies responsible for the product's original design, usually also being the product owners, have turned their back on manufacturing issues. To shorten the time to market, they seek reliable manufacturing companies as their business partners to take care the manufacturing aspects of the product. Very often, these companies release product designs that are immature and untested and rely on their partners to refine the design during the manufacturing stage. Therefore, manufacturing companies nowadays have to provide a wide spectrum of services that include proprietary IP's for product designs and manufacturing process development. TSMC, Taiwan Semiconductor Manufacturing Company, has championed such a foundry service model and enjoyed a great success in the semiconductor sector. Many companies of various sectors in Taiwan are following such a model and intend to evolve from low-value-adding manufacturers to high-value-adding manufacturing services providers.

To become a successful manufacturing services provider, a speedy yield ramp-up during the early stage of product/process development is critical by reasons manifested earlier. Because of the increasing complexity of modern manufacturing processes, to ramp up production yield is no easy problem. A typical semiconductor fabrication process consists of more than 300 steps. Too many possible factors during the process could contribute to the low yield. Nevertheless, no systematic solutions exist as far to this problem. Even for a successful company like TSMC, the yield ramp-up process remains to be a tedious, time-consuming task of swarming process integration engineers. Even though many techniques have been developed over the years aiming to continuously improve yield or product quality during the manufacturing stage, the practitioners have mostly overlooked these techniques. It is this project's mission to take up again these techniques and to develop an effective, systematic yield-ramp-up solution. These techniques are developed by different research communities but have a common goal: *to on-line continuously optimize the process performance*. They include: evolutionary operations (EVOP), an on-line design of experiment (DOE), technique first developed by Box, an applied statistician, in 60's [1]; Optimum Experimental Design [2]; perturbation-based optimization techniques [3]-[7] by Chemical process control researchers and Mathematicians; and Ordinal Optimization methods [8][9] by discrete-event control researchers. In this concise report, we'll investigate and compare these techniques. We then propose an integrated, comprehensive evolutionary optimization method that can be easily applied in practice.

2. Comparsion of on-line optimization methods

EVOP is an approach that simultaneously investigates effects of two or three variables on the system's output performance. Statistical effects significances then serve as the bases to determine the improvement strategies. As shown in Fig. 1, two variables, temperature and flow rate, are studied. An initial 2-level factorial experimental region is first chosen. Effects are calculated and their significances are checked after data is collected. "Replicate" data are continuously collected until effects are found significant. In the example, "yield" is the system response of concern and is found higher at the lower level of the temperature and at the higher level of the flow rate. To continuously improve the yield, the next phase of experiment region is chosen. Such steps are continued until the system performance is improved to a desired level or the improvement is no longer significant.

Perturbation-based process optimization is another technique used to on-line improve system performance. The perturbation-based optimization method is to impose a perturbing waveform onto a manipulating variable and use the observed responded variation to apply a correction to the average manipulating variable, *x*, as shown in Fig. 2. The perturbing signal is a sine wave in Fig. 2. The observed output, *y*, will respond a waveform embedded with noise. Suppose that the output is continuously measured electrically. Its correct waveform can be extracted from the noise by multiplying a correlator over time: $I_t = \int_0^t yzdt$.

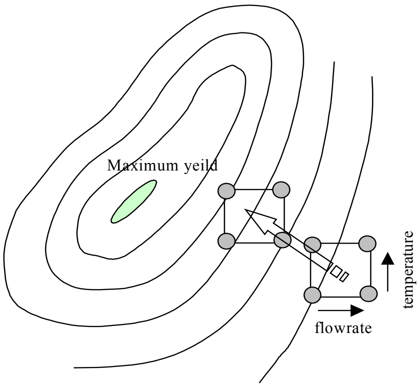

Figure 1 Evolutionary Operation (EVOP) procedure to improve process yield

Since the integral I_t is proportional to $\partial y / \partial x$, information contained in I_t can be thus used to feed back to the controller, which in turn determines the adjustment to *x*.

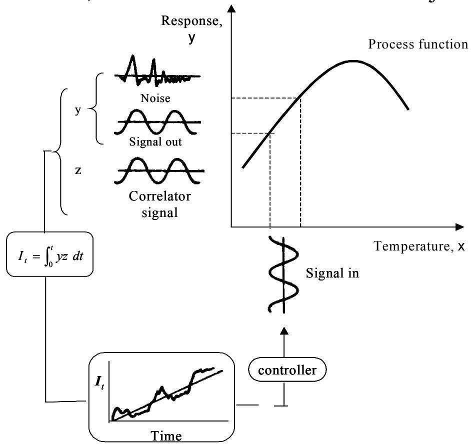

Figure 2 Perturbation-based continuous process improvement

As illustrated in the above example, the key idea of perturbation-based technique is to perturb in order to estimate $\partial y / \partial x$. Adjustment is then made based on this estimate. Over the years, this technique has been developed into become a perturbation-based real-time optimization (PRTO) techniques. In the literature, there are mainly three types of perturbation-based PRTO methods: dynamic integrated system optimization and parameter estimation, linear adaptive on-line optimization and Quadratic Adaptive On-line Optimization.

In the late 90's, discrete-event control researchers proposed an ordinal optimization method to more effectively optimize discrete-event system performance through simulation. Though the problem domain is very different from the manufacturing process improvement problem, the idea is strikingly similar. Let the system manipulating variable be *x*, which is adjustable to optimize the system output, and t_x be the sample size. The sampling cost is then the summation of t_x over possible values of x, denoted by the set X: $\sum_{x\in X} t_x$. The goal of ordinal optimization is then to choose t_x for all possible values of x such that the total sampling cost is minimized, subject to a restriction that the process performance has to be improved to a certain degree

$$\min_{t_x} \sum_X t_x$$
$$s.t. \text{ alignment probability} \geq P^*$$

where alignment probability is defined to represent the confidence level that one can be assure how close the resulted system performance is to the best possible performance and P^* is a pre-determined confidence level. Or one can formulate the optimization problem as to choose t_x such that the alignment probability is maximized, subject to a limited sampling cost:

$$\min_{t_x} [\text{alignment probability}]$$
$$s.t. \ \sum_{x\in X} t_x = T$$

where T is the sampling budget available. In this task, ordinal optimization methods will be also studied rigorously and to see how it fits in the process improvement problems.

Similar to Ordinal Optimization, the goal of optimum experimental design (OED) is to maximize the robustness of the empirical process model with a minimal number of experimental runs. The so-called variance-optimal designs, such as central composite designs and D-optimal designs, are widely used to achieve this goal. After the empirical model is built, the ridge analysis is then used to find the process optimum. For example, the temperature and plasma power of etcher affects the performance of critical dimension (CD). The empirical model of this process can be expressed as follows:

$$y = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 + b_{11}x_1^2 + b_{22}x_2^2 \qquad (1)$$

where, y represents the response, CD performance; x_1 *and* x_2 are the temperature and the plasma power. The Rridge Analysis (RA) will then determine a direction (Δx_1, Δx_2) to improve the CD performance.

We summarize comparison of above methodologies in the following table:

Table 1 Comparison of on-line optimization methods

	Physical Model	Empirical Model	Sequential Experiment	Process Perturbatio	Steady State Model	Dynamic Model
EVOP		♦	♦		♦	
RTO	♦	♦		♦		♦
OO	♦		♦			♦
OED/RA		♦	♦		♦	

Based on the comparison, we propose a methodology combining concepts of OED/RA and OO.

3. Evolutionary Process Optimization through

In the typical semiconductor fabrication, factors in an earlier process step affect the response of a later process step by affecting the effects of factors at that later process step. As shown in Fig 3, the temperature setting of PECVD indirectly affects the post-etching CD by affecting the effects of the plasma power and temperature of the etcher on the CD and consequently on the final yield.

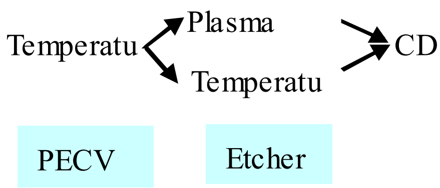

Figure 3 Effects of variables on CD

The empirical model for such a multi-step process can be written as

$$y = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 + b_{11}x_1^2 + b_{22}x_2^2 + b_{13}x_1x_3 + b_{23}x_2x_3 \qquad (2)$$

where x_3 is the PECVD temperature and the additional terms, x_1x_3 and x_2x_3, describe the indirect effect of PECVD temperature on the CD through the etcher's two major factors: temperature and plasma power. In this research, we propose an evolutionary experimental design to better estimate the process improvement direction for such multi-step processes.

We first derive a measure to evaluate the confidence region of the improvement path. With the confidence region evaluation measure, we propose an R-optimum sequential design procedure. Finally, we'll show how the proposed experimental design outperforms the conventional designs.

The issue now is to assure the accuracy of the estimated improvement path attained by the ridge analysis. As shown in Fig 4, ridge points on different radius form the ridge path, i.e. the improvement path of the empirical model. The objective is to statistically find the confidence interval of these ridge points.

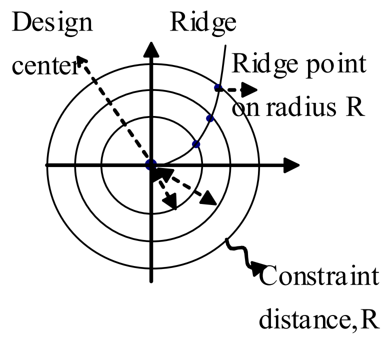

Figure 4 Ridge points on various radius R

Using the Lagrange multiplier, μ, the ridge point (x_{s1}, x_{s2}, x_{s3}) at a given perimeter can be found by (3):

$$
\begin{aligned}
(b_{11}-\mu)x_{s1} + \tfrac{1}{2}b_{12}x_{s2} + \tfrac{1}{2}b_{13}x_{s3} &= \tfrac{1}{2}b_1 \\
\tfrac{1}{2}b_{12}x_{s1} + (b_{22}-\mu)x_{s2} + \tfrac{1}{2}b_{23}x_{s3} &= \tfrac{1}{2}b_2 \\
\tfrac{1}{2}b_{13}x_{s1} + \tfrac{1}{2}b_{23}x_{s2} + (b_{33}-\mu)x_{s3} &= 0
\end{aligned}
\qquad (3)
$$

(3) can be rewritten as $(\boldsymbol{B}-\boldsymbol{I})\boldsymbol{x}=(1/2)\boldsymbol{b}$.

Using an inequality established by [12], we eventually show that confidence region volume of the ridge point is proportional to

$$\left|\sqrt{\Lambda^{-1}}\right| = \frac{1}{\sqrt{\lambda_1}}\frac{1}{\sqrt{\lambda_2}}\frac{1}{\sqrt{\lambda_3}} \qquad (4)$$

where Λ is the diagonalized matrix of $\boldsymbol{J}$ and λ_1, λ_2, and λ_3, are eigenvalues of $\boldsymbol{J}$. And

$$\boldsymbol{J}=(\boldsymbol{B}-\mu\boldsymbol{I})^{\mathrm{T}}\boldsymbol{V}^{1}(\boldsymbol{B}-\mu\boldsymbol{I})$$

where $\boldsymbol{V}$ is the covariance matrix of ridge point (x_{s1}, x_{s2}, x_{s3}).

Base on the derivation above, $|\Lambda^{-1}|$ is used as a new experimental design criterion to enhance the confidence of improvement path. With the criterion, we propose a ridge-optimum sequential design, referred to as R-optimum sequential design.

Back to the example of the semiconductor etching process, an initial empirical model as equation (1) is built for the etching process by a 3^2 factorial experimental. An example of the initial ridge path estimated from the initial model is shown in Figure 5.

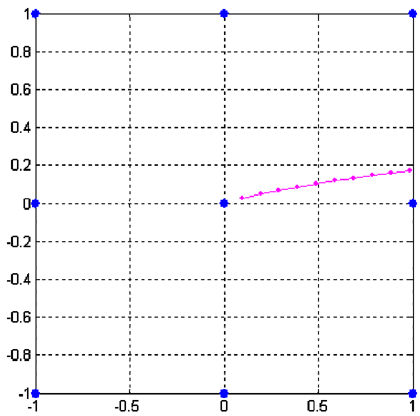

Figure 5 Initial ridge path for etching

Based on the initial ridge path, the initial design is augmented by adding one design point to minimize the confidence region of the initial ridge point. With the augmented design point, a new empirical model is built and a new ridge point is found. A new design point is then added again to minimize the confidence region of the new ridge point. To minimize the ridge-point confidence region, $|\Lambda^{-1}|$is used as the measure. The design point that minimizes $|\Lambda^{-1}|$ will be selected as the new experimental design point. Figure 6 illustrates this R-optimum design procedure.

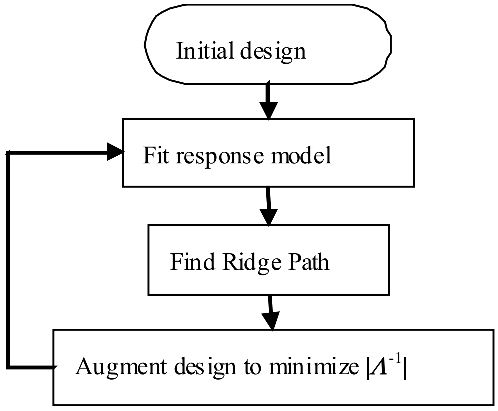

Figure 6 R-optimum sequential design

Considering the previous example, the initial ridge path in Figure 5 is now shown as the gray stars in the 3-dimentional space in Figure 7. The design is then augmented first by adding a new design point (green point in Figure 7). Model is then extended to the form of equation (2) in order to estimate the indirect effect of the PECVD temperature. A new ridge path is then found (green line in Figure 7) based on the first augmented model. Red, blue and purple paths in Figure 7 are three more paths found evolutionarily by the R-optimum design procedure.

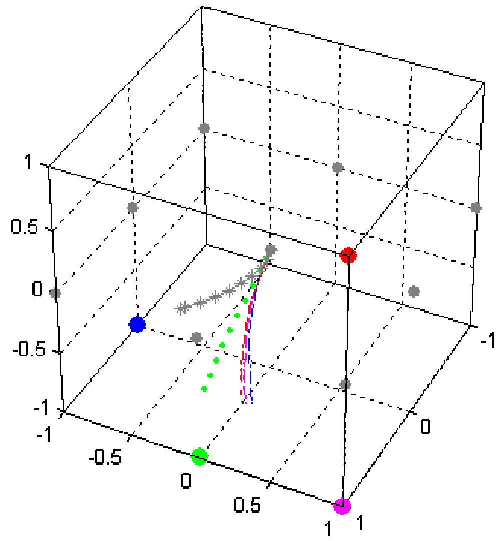

Figure 7 Evolutionary ridge paths by R-optimum experimental design

4. Validation and Concluding Remarks

405 hypothetical models are used to evaluate the performance of R-optimum sequential design. The design region is restricted to [-1, 1]. With initial points fixed to 3^2 factorial design points as the example above, we intend to augment the model with four more design points. The R-optimal design continuously augments the design one point at a time. The proposed design is compared to two types of D-optimal designs: one-by-one augmentation design and 4-in-once augmentation design. The augmentation procedure follows a basic exchange algorithm [1] for both R- and D-optimum designs. Exchangeable candidates are the 3^3

factorial points. For each step of the augmentation, candidates that can gain the information most (D-optimum design) or reduce $|\Lambda^{-1}|$ most (R-optimum design) would be selected.

The distance between the estimated ridge path and the true path is used to evaluate the performance of experimental designs. The closer the distance, the better the performance. Fig 7 shows the average distance of the three experimental designs. It can be seen that the R-optimum design outperforms both D-optimum designs.

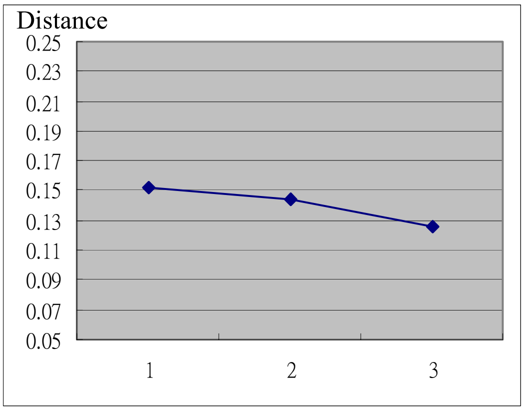

Fig 7**.** Average distance from the true ridge path of 405 models. 1: 4-in-once D-optimum.2: one-by-one augmentation D-optimum and. 3: R-optimum design

In this research, it has been shown that the novel R-optimum sequential design outperforms the conventional D-optimal designs in terms of improvement path estimation. The proposed design procedure can be applied to any multi-step processes and can be very efficient for the semiconductor yield ramp-up process.

References

1. Box, G. E. and N. R. Draper, *Evolutionary Operation – A Statistical Method for Process Improvement*, John Wiley & Sons, New York, 1969.
2. Atkinson, A. C. and A. N. Donev, *Optimum Experimental Design*, Clarendon Press, Oxford, 1992.
3. Bensoussan, A., *Perturbation Methods in Optimal Control*, John Wiley & Sons, Chichester, 1988.
4. Garcia, C. and M. Morari, "Optimal Operation of Integrated Processing Systems: Part I. Open-Loop On-Line Optimizating Control," *AIChE J.,* 27, 960, 1981.
5. Golden, M. P. and B. E. Ydstie, "Adaptive Extremum Control Using Approximate Process Models," *AIChE J.*, 35, 7, 1157, 1989.
6. McFarlane, R. C. and D. W. Bacon, "Empirical Strategies for Open-Loop On-line Optimization," *The Canadian Journal of Chemical Engineering*, 67, 665, 1989.
7. Becerra, V. M. and P. D. Roberts, "Dynamic integrated system optimization and parameter estimation for discrete time optimal control of nonlinear systems," *Int. J. Control*, 63, 2, 257-281, 1996.
8. Ho, Y. C. and R. S. Sreenivas, and P. Vakili, "Ordinal Optimization of DEDS," *J. Discrete Event Dynamic Systems*, vol. 2, no. 2, pp. 61-88, 1992.
9. Chen, H. C. and E. Yucesan, "Computing Efforts Allocation for Ordinal Optimization and Discrete Event Simulation," *IEEE Trans. Automat. Contr.*, vol. 45, no. 5, pp. 960-964, 2000.
10. Ho, Y. C. and X-R Cao, *Perturbation Analysis of Discrete Event Dynamic Systems*, *Kluwer Academic Publishers*, Boston, 1991.
11. Suri, R. and K. Otto, "Manufacturing System Robustness Through Integrated Modeling," *Journal of Mechanical Design*, 2001.
12. Box, G. E. and Hunter, J.S., (1954). "A Confidence Region for the Solution of a Set of Simultaneous Equations with an Application to Experimental Design" *Biometrika*, Volume 41, pp.190-199