J Intell Manuf
DOI 10.1007/s10845-013-0791-5

An empirical study of design-of-experiment data mining for yield-loss diagnosis for semiconductor manufacturing

Chen-Fu Chien · Kuo-Hao Chang · Wen-Chih Wang

Received: 6 January 2013 / Accepted: 13 May 2013

Abstract To maintain competitive advantages, semiconductor industry has strived for continuous technology migrations and quick response to yield excursion. As wafer fabrication has been increasingly complicated in nano technologies, many factors including recipe, process, tool, and chamber with the multicollinearity affect the yield that are hard to detect and interpret. Although design of experiment (DOE) is a cost effective approach to consider multiple factors simultaneously, it is difficult to follow the design to conduct experiments in real settings. Alternatively, data mining has been widely applied to extract potential useful patterns for manufacturing intelligence. However, because hundreds of factors must be considered simultaneously to accurately characterize the yield performance of newly released technology and tools for diagnosis, data mining requires tremendous time for analysis and often generates too many patterns that are hard to be interpreted by domain experts. To address the needs in real settings, this study aims to develop a retrospective DOE data mining that matches potential designs with a huge amount of data automatically collected in semiconductor manufacturing to enable effective and meaningful knowledge extraction from the data. DOE can detect high-order interactions and show how interconnected factors respond to a wide range of values. To validate the proposed approach, an empirical study was conducted in a semiconductor manufacturing company in Taiwan and the results demonstrated its practical viability.

Keywords Data mining · Design of experiment · Yield enhancement · Defect diagnosis · Semiconductor manufacturing

C.-F. Chien (✉) · K.-H. Chang · W.-C.Wang
Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu 30013, Taiwan
e-mail: cfchien@mx.nthu.edu.tw

Introduction

Semiconductor manufacturing is among the most complex manufacturing today, in which a lengthy and reentrant process steps are employed with advanced tools to fabricate a variety of Integrated Circuits (IC) including microprocessors, memories, digital signal processor, and application-specific logic on the wafers. Driven by Moore's law that the number of transistors fabricated on a wafer will be doubled every 12 or 24 months with lower average selling price (Moore 1965), semiconductor industry has strived for continuous technology migrations to maintain competitive advantages and thus has achieved unparalleled growth via penetrating into various domains for component substitution in the past few decades. Due to technology challenges, semiconductor manufacturing has become very capital intensive, in which building a modern 12 inch wafer fabrication facility (fab) with 40 nanometer process technology requires more than 4 Billion USD (Chien et al. 2010; Chen and Chien 2011). Now only a small number of survived companies compete on providing foundry services or selling interchangeable semiconductor products with most leading process technologies to customers worldwide. High yield and high overall wafer effectiveness (OWE) lead to low unit cost and high profit margins (Chien et al. 2013a). Therefore, fast ramp up for newly released process technology and quick response to yield excursion to reduce loss are crucial to maintain technology innovation and cost reduction per transistor (Chien et al. 2011, 2012a; Wu 2013). However, as wafer fabrication processes become increasingly complicated in nano technologies, many factors including recipe, process, tool, and chamber that are interrelated affect the yield of fabricated wafers, while spec tolerance is significantly reduced owing to physical limits. Hence it is increasingly difficult to quickly identify the

Published online: 28 May 2013

root causes and suggest the corresponding corrective actions when yield loss occurs to minimize the incurred excursion cost.

Design of Experiments (DOE) is a cost-effective approach to consider multiple factors simultaneously for performance characterization or sensitivity analysis, rather than conducting experiments via a trial-and-error manner or changing one factor at a time. Moreover, DOE allows high-order interactions to be detected and shows how interconnected factors respond to a wide range of values. Thus, DOE has been successfully applied in many fields such as manufacturing, pharmaceutical, and business. However, applying DOE is a difficult task in semiconductor manufacturing owing to the following reasons. Firstly, semiconductor industry is very capital intensive, in which equipment utilization is usually driven to maximum to get an early return on the capital investment (Chien et al. 2012b), and thus leaves little room for experiments to be conducted in production line. Secondly, since the experimentation cost in fabs is high in terms of time and money, few companies can afford to conduct comprehensive design of experiments to characterize semiconductor production system with hundreds of process steps and tools. Thirdly, wafer fabrication involves hundreds of reentrant process steps including oxidation, deposition, metallization, lithography, etching, ion implantation, photo-resist strip, cleaning, inspection and measurement (Chien and Chen 2007a,b; Wu and Chien 2008; Wu et al. 2012). Thus, DOE principles such as *blocking*, *randomization* and *replication* are difficult to follow in the complicated wafer fabrication process with complex data variety. While the aforementioned restrictions pose significant challenges to employing DOE in semiconductor manufacturing, there exists abundant data collected automatically or semi-automatically during wafer fabrication process. If well exploited, the data can provide useful information to identify or narrow the scope of potential root causes of low yield. Data mining has been widely used to extract meaningful and useful information from huge amounts of data in various domains. Because hundreds of factors must be considered simultaneously to accurately model the system's behavior, data mining requires tremendous time for analysis and interpretation. However, because many process steps are highly dependent on the previous steps in semiconductor manufacturing, complicated multicollinearity is involved in the related factors. Thus, without proper preparation in advance, data mining may become data dredging, resulting in misleading relationships that are difficult to be interpreted and used for fault diagnosis. Thus, in practice, most companies rely on domain engineers to perform fault diagnosis based their own domain knowledge in different process modules, which however tends to be biased and the whole process can take a long time.

To address the needs in real settings, this study aims to develop a novel approach that integrates DOE and data mining to enable effective information extraction from massive manufacturing data efficiently. Rather than the traditional DOE in which data is collected based on a designed experiment, the proposed approach, namely, retrospective DOE data mining, takes a backward direction to "match" potential experimental designs from the data and then employ effective analysis to extract useful information from alternative experimental designs. In particular, the proposed approach first employs Kruskal-Wallis test and multiple comparison tests to identify the key process stages and distinguish the problematic machines from the normal ones. Then, the retrospective DOE is performed to find the experimental designs existing in the dataset by matching the alternative combinations of process stages (factors) and machines (levels). Therefore, the matched experimental designs enable the application of ANOVA and regression analysis to establish the causal relationship between yield and the problematic machines across different process stages. Indeed, the proposed retrospective DOE has following benefits. Firstly, it does not require extra experimentation cost. By exploiting the data normally collected during manufacturing process to match alternative experimental designs, it allows the advantages of traditional DOE to be realized. Secondly, it empowers the use and advantages of the statistical tools such as ANOVA and regression analysis where the functional relationship among factors can be established and easily interpreted for fault diagnosis. Thirdly, the use of orthogonal design can mitigate the multicollinearity issue that exists in semiconductor manufacturing data, since the effects of main factors in orthogonal designs are independently assessed (Khuri and Cornell 1996). Fourthly, with a systematic approach, the developed framework can quickly narrow the scope of possible root causes when yield loss occurs and provide useful manufacturing intelligence for domain engineers to take needed actions. The practical viability of the proposed framework is validated by an empirical study in a leading semiconductor company in Taiwan.

The remainder of this paper is organized as follows. Section "Fundamental" reviews related literature. Section "The approach" describes the proposed framework with an illustration for fault diagnosis in semiconductor manufacturing. Section " An empirical study" presents an empirical study to detect the root causes in a semiconductor company for validation. Section "Conclusion" concludes with discussions of contributions and future research directions.

Fundamental

Design of experiment

For any system that the underlying relationship among the factors is unknown, DOE can be employed to con-

struct the functional relationship between the response and a number of factors. Compared to the naïve one-factor-at-a-time approach, DOE can simultaneously estimate the main and secondary effect of factors with fewer experimental runs (Wu and Hamada 2000). A typical DOE framework begins with defining the problem and selecting the response variable and the related factors (Coleman et al. 1993). Based on the selected factors, an experimental design including the design type, the values of factors and the number of replications should be explicitly specified. Then, the experiments are conducted according to the experimental design. Once the experiment is finished, statistical tools such as ANOVA and regression analysis are applied to analyze the collected data for making inference about the system. The conclusion can be drawn based on a statistical framework. A more comprehensive treatment for each part of the DOE framework can be found, for example, in Montgomery (2005). It is remarkable that, although DOE is employed in R&D for semiconductor technologies (May et al. 1991; Chien and Hsu 2006, 2011), it is generally very difficult to implement DOE to analyze the semiconductor manufacturing system to uncover underlying causes for low yield.

Data mining for yield enhancement

Data mining has been efficiently used, by automatic or semi-automatic means, to analyze and explore the huge amount of data in various fields (Berry and Linoff 1997; Han and Kamber 2001; Kusiak and Kurasek 2001). Data mining techniques can be categorized into four types: association rules, clustering, classification, and prediction (Fayyad et al. 1996; Han and Kamber 2001). Association rules refer to the discovery of association between the data items. A popular application of association is market basket analysis, which discovers the purchase habits of customers by searching for the sets of items that are frequently purchased together. Clustering is the process of dividing a dataset into several groups via maximizing the intra-cluster similarity and minimizing the inter-cluster similarity. Classification is to derive a model that can determine the categorical class of an object based on its attributes. Prediction model can be developed to forecast a continuous value or future data trends.

Data mining has been applied for semiconductor manufacturing to deal with various problems such as process improvement (Braha and Shmilovici 2002, 2003), wafer bin map clustering (Hsu and Chien 2007; Jeong et al. 2008; Chien et al. 2013b; Liu and Chien 2013), yield enhancement (Chien et al. 2007; Hwang and Kuo 2007), demand forecast (Chien et al. 2010), human capital management (Chien and Chen 2007a,b, 2008), and cycle time reduction (Kuo et al. 2011; Chien et al. 2012a). However, the challenges for semiconductor manufacturing data mining is not only to integrate the interwoven big data into a data analysis platform but also to develop effective analytical solutions to deal with increasingly complex interrelations in various data and thus extract useful manufacturing intelligence to support various decisions (Chien et al. 2007). Due to the complicated multicollinearity in the related factors of semiconductor manufacturing data, data mining may become data dredging to find misleading patterns that are difficult to be interpreted and useless for fault diagnosis.

Comparison of design of experiment and data mining

While DOE and data mining share the same goal of extracting useful information and intelligence from data, their fundamental difference in data collection gives them the advantages and disadvantages when performing data analysis. The required data for data mining can be collected during the normal operations of the manufacturing process being studied and it is therefore generally not necessary to introduce dedicated processes for data collection (Harding et al. 2006). That is, data mining does not require extra experimentation cost, which is practical to many companies. Furthermore, data mining does not assume a predefined model, yet is capable of extracting and discovering potentially meaningful patterns and relationship within data. That is, data mining is discovery-driven rather than hypothesis-driven. However, data mining has some disadvantages. Firstly, some algorithms such as neural networks are a black box to analysts and the extracted results are difficult to be interpreted by domain experts. Secondly, factors causally related to yield loss can be confounded with a range of other factors and as a result, the extracted patterns can be spurious. Thirdly, using data mining techniques can be easily misled by the apparently significant results due to random noises and multicollinearity involved in hundreds of tools and processes and can thus become data dredging and lost the confidence of domain experts to use the results. Finally, semiconductor data mining is time consuming for data preparation before further analysis and thus is difficult to support real-time trouble shooting.

DOE is a hypothesis driven framework to conduct experiments in a controlled and designed manner, and hence it typically entails extra experimentation cost. DOE has the advantages of transparent analysis framework that is conceptually easy to implement and many well-established statistical tools such as ANOVA and regression analysis are applicable for analyzing collected data. Moreover, DOE allows the analysts to focus on a subset of interested factors rather than considering all the factors at one time. Finally, the DOE-based statistical analysis can eliminate multicollinearity when orthogonal designs are used, due to that the effects of factors in orthogonal designs are independently assessed.

Because data mining and DOE both have their advantages and disadvantages, the proposed approach integrates them

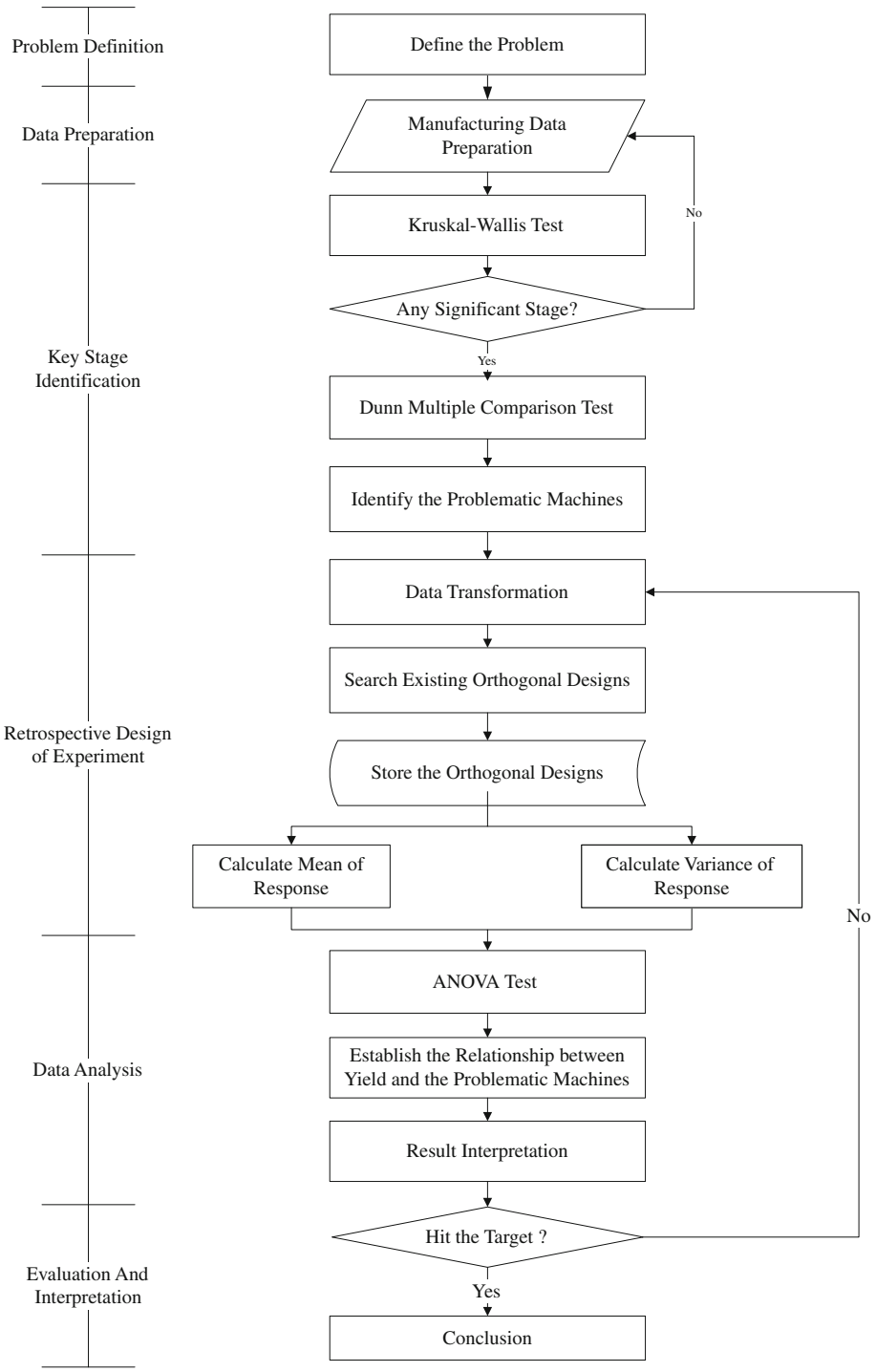

Fig. 1 The integrated framework for retrospective DOE data mining

to deal with yield-loss diagnosis problem in semiconductor manufacturing.

The approach

The proposed framework for retrospective DOE data mining consists of six stages as illustrated in Fig. 1: problem definition, data preparation, key stage identification, retrospective DOE matching, data analysis, and result evaluation and interpretation. The proposed approach starts with collaboration with domain experts to define the right problem, which entails a better understanding of the problem nature for selecting the target response variable and related factors. This study focused on the yield loss problem for identifying the assignable root causes (machines) of low yield in

semiconductor manufacturing. The yield rate and process stages are defined as the response variable and factors, respectively. Then, data preparation including data cleaning, data integration and data transforming is performed to ensure data quality to generate effective results. Then, the Kruskal–Wallis test (K–W test) (Kruskal and Wallis 1952) is employed to screen the key stages where some machines have significantly low yield rates than others. For each key process stage, the multiple comparison tests can be further used to detect the problematic machines, e.g., Holm–Bonferroni method (Holm 1979), closed test procedure (Marcus et al. 1976), and Dunn's test (Dunn 1964). In this paper, we choose Dunn's test to distinguish the problematic machines from the normal ones. The retrospective DOE is conducted to find the potential orthogonal designs, either a full factorial or fractional factorial design that can be matched in the existing data. When all candidate designs are matched, data analysis including ANOVA and regression analysis is performed to establish the functional relationship between the yield and the problematic machines and to identify the problematic machines that have the most significant impact on the yield rate. The derived results are discussed with domain experts for evaluation and interpretation. The whole process is iterative until all problematic machines that result in low yield are identified.

Problem definition

The semiconductor manufacturing process is complex and lengthy and involves hundreds of operations at different process stages. A typical semiconductor factory may include one to ten major fabrication process flows, producing 40,000 or more wafers per month. Product cycle times ranges from 40 to 60 days depending on the complexity of the technology. Fab operates non-stop 7 days a week, 24 h a day, in which a huge amount of data are automatically or semi-automatically generated and accumulated during wafer fabrication process.

This research focused on yield-loss diagnosis to identify the abnormal process stages and problematic machines. That is, the yield rate is defined as the response variable, and the process stages and machines are considered as "factor" and "level," respectively.

Data preparation

Because the collected manufacturing data usually include noisy, missing and inconsistent data, data preparation is important to produce effective and meaningful analysis results. Data preparation includes data integration, cleaning and transforming. In particular, data integration merges data from multiple sources into a coherent data store; data cleaning removes noisy data, identifies outliers, fills in missing data and corrects inconsistencies in data; data transformation consolidates data to make them appropriate for the subsequent mining process.

Key stage identification

Key stage identification is to narrow the number of process stages to reduce the time and efforts required for fault diagnosis.

Firstly, nonparametric K–W test is employed to examine the differences of the yields among the machines at the same process stage, because the distribution of the yield is not normal. If the process stage has some machines with the yields significantly different from the others, the process stage is identified as a key stage for further investigation. In particular, the following steps are employed for K–W test:

Step 1 Rank all data from all groups together. Let R_{ij} be the rank of observation j from sample i. Assign any tied values the average of the ranks they would have received.
Step 2 Calculate the test statistic:

$$H = \frac{1}{S^2}\left[\sum_{i=1}^{k}\frac{{R_i}^2}{n_i} - \frac{N(N+1)^2}{4}\right] \tag{1}$$

where N is the total number of observations, $R_i = \sum_{j=1}^{n_i} R_{ij}$ is the sum of the ranks for sample i, n_i is the number of observations from group i, and

$$S^2 = \frac{1}{N-1}\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i} R_{ij}^2 - \frac{N(N+1)^2}{4}\right]. \tag{2}$$

Step 3 If $H > \chi^2_{(\alpha,k-1)}$, there exists significant difference among samples.

In particular, Table 1 provides an example for illustration. Table 2 shows the yield rates of three machines and the corresponding ranking in the parenthesis.

Table 1 Yield rate of the machines at process stage A

Stage	Machine		
	M1 (%)	M2 (%)	M3 (%)
A	9.10	8.86	9.27
	9.41	9.00	9.15
	9.07	8.92	9.02
	9.03	8.72	9.06

Table 2 Ranks of the machines at process stage A

Stage	Machine		
	M1	M2	M3
A	9.10 % (9)	8.86 % (2)	9.27 % (11)
	9.41 % (12)	9.00 % (4)	9.15 % (10)
	9.07 % (8)	8.92 % (3)	9.02 % (5)
	9.03 % (6)	8.72 % (1)	9.06 % (7)
Rank sum	35	10	33

Then, the hypothesis test is conducted as follows:

H_0: The yield rates of all three machines in stage A are equal
H_1: The yield rates of all three machines in stage A are not all equal

Based on Eqs. (1) and (2), the test statistic H and S^2 are calculated. Since $H = 7.423 > \chi^2_{0.15}(2)$, H_0 is rejected at the level of significance $\alpha = 0.15$. That is, the yield rates of the three machines at the stage A are not equal. Therefore, the process stage A is considered as a key stage that may include problematic machines causing low yield.

If there are significant differences among the machines of a specific process stage, multiple comparison tests can be applied to distinguish the problematic machines from the normal ones as follows:

Let $\overline{R_i}$ and $\overline{R_j}$ denote the average rank in the ith and jth samples, respectively. Further let $n = \sum_{i=1}^{k} n_i$. If

$$\left|\bar{R}_i - \bar{R}_j\right| \le Z_{[\frac{\alpha}{k(k-1)}]}\sqrt{\frac{n(n+1)}{12}\left(\frac{1}{n_i}+\frac{1}{n_j}\right)} \tag{3}$$

then ith sample is declared to be consistent with the jth sample with α-level type-I error. Note that the experiment-wise error rate of multiple comparison procedure is controlled at α.

Following the previous example, the sums of ranks of the three machines in stage A can be calculated as R1 = 35, R2 = 10, and R3 = 33, respectively. Based on Eq. (3), the test statistics are computed as follows:

$$\left|\overline{R_1}-\overline{R_2}\right| = |35-10| = 25 \ge Z_{\left[\frac{0.15}{3(3-1)}\right]}\sqrt{\frac{12(12+1)}{12}\left(\frac{1}{4}+\frac{1}{4}\right)} = 4.997$$

$$\left|\overline{R_1}-\overline{R_3}\right| = |35-33| = 2 \le Z_{\left[\frac{0.15}{3(3-1)}\right]}\sqrt{\frac{12(12+1)}{12}\left(\frac{1}{4}+\frac{1}{4}\right)} = 4.997$$

$$\left|\overline{R_2}-\overline{R_3}\right| = |10-33| = 23 \ge Z_{\left[\frac{0.15}{3(3-1)}\right]}\sqrt{\frac{12(12+1)}{12}\left(\frac{1}{4}+\frac{1}{4}\right)} = 4.997$$

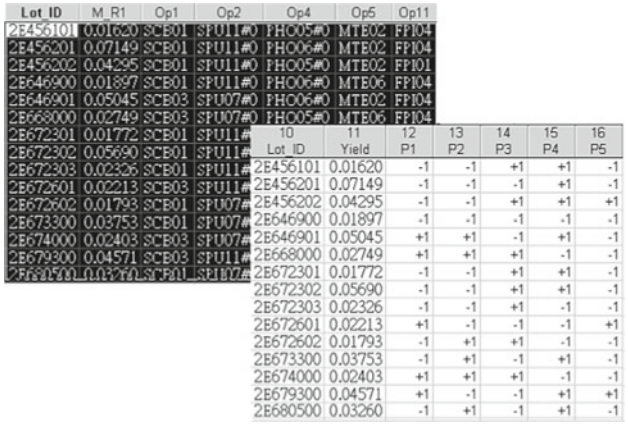

Fig. 2 Data transformation

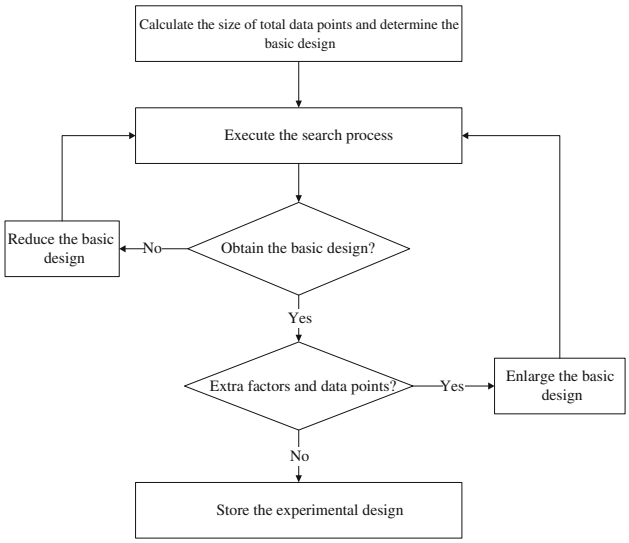

Fig. 3 Flowchart of retrospective DOE

Let the significance level $\alpha = 0.15$. It is concluded that the yield rate of M1 differs from M2 and equals M3, and the yield rate of M2 differs from M3. In other words, M1 and M3 are potentially problematic machines, while M2 is a normal machine.

Retrospective DOE Matching

The purpose of retrospective DOE is to match potential orthogonal designs in the existing data. The identified designs, in combination with statistical tools, can provide useful information to support further fault diagnosis. Recall that the multiple comparison test has distinguished the problematic machines from the normal ones. The normal machines is labeled as +1 and the problematic machines are labeled as −1 as illustrated in Fig. 2.

Then, the proposed retrospective DOE is performed as the flowchart in Fig. 3.

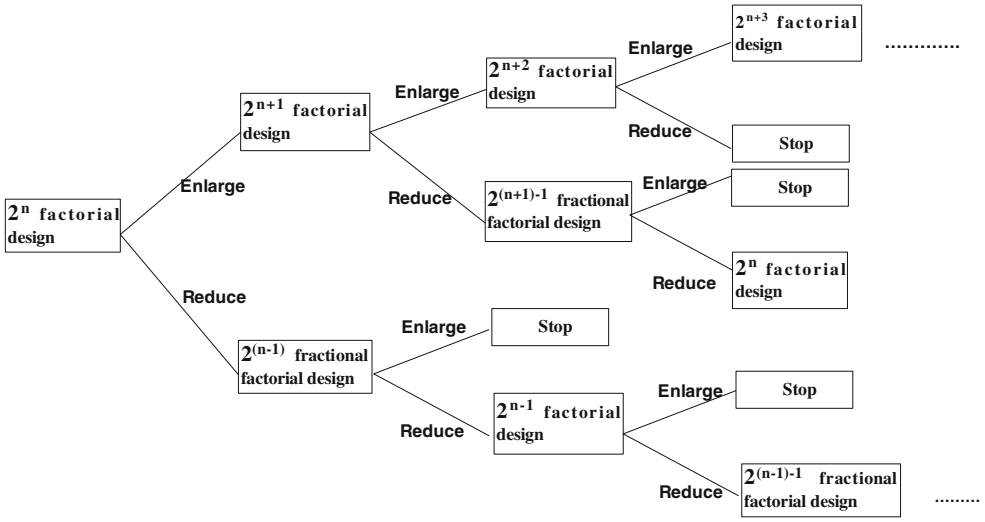

Fig. 4 Sequential search for best-match designs of experiment

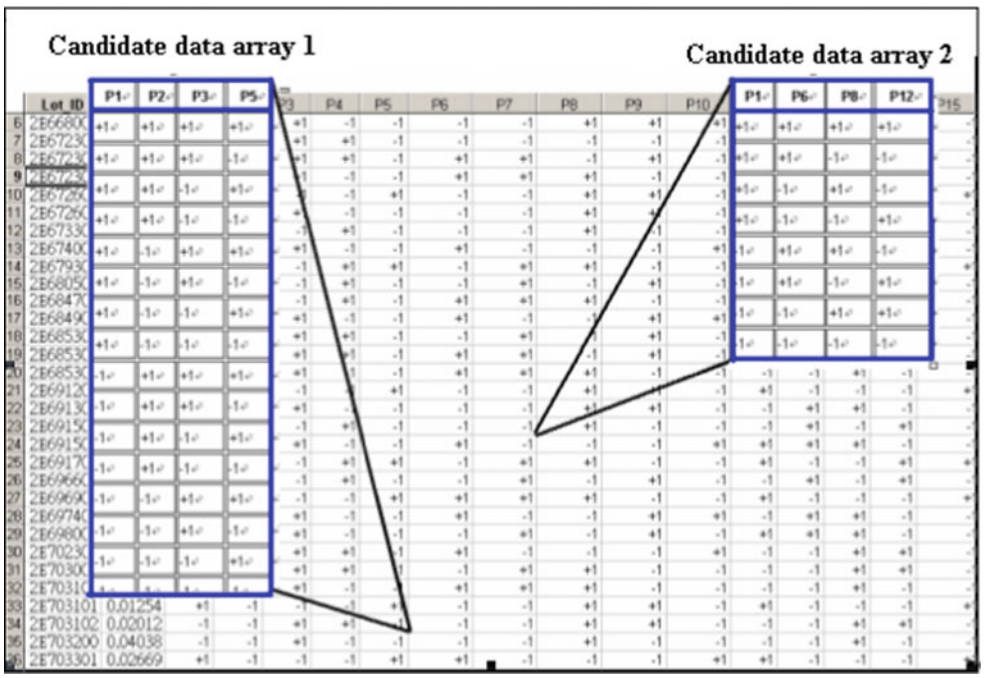

Fig. 5 Search for the candidate designs in data

Step 1: Calculate the size of total data points and determine the basic design

The basic design is determined based on the size of total data points. Specifically, the basic design is the experimental design that only requires less than or equal to one quarter of the total data points. For example, suppose there are 100 data points and 8 process stages are identified as key stages. The basic design are those that only requires 100/4 = 25 data points. Therefore, $2^4 = 16$ fractional factorial design is one of the basic designs. All experimental designs, including full and fractional factorial designs that satisfy the requirement should be considered.

Step 2: Execute the search process

The search process is illustrated as Fig. 4. Suppose that 2^n full factorial design is selected as the basic design. In particular, the search process is first focused on the 2^n full factorial design. If it is found, yet there are factors not being considered and extra data points are still available, the basic design is enlarged and the search process is re-executed. The process is continued until no larger design can be found in

the dataset. On the other hand, if 2^n full factorial design cannot be found, the basic design is reduced and the search process is re-executed. The process is also continued until at least one fractional factorial design can be found in the data set. In general, there are many candidate designs that can be matched with the existing data, as illustrated in Fig. 5.

Step 3: Store the experimental design

As soon as all designs are found, the data array is stored for subsequent data analysis in the next step.

Data analysis

Once the data arrays are matched with specific designs, ANOVA test and regression analysis are performed for identifying the problematic machines that actually result in low yield, and establishing the relationship between the yield and the problematic machines. Notably the analysis requires that the response variables be adequately described by the model $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$, where τ_i refers to the effect of treatment i and ε_{ij} are independent and normally-distributed random variables with mean zero and constant but unknown variance. These assumptions can be checked through the *normal probability plot* and the *residual plot*. If the model assumptions are suspected not satisfied, some remedy measures should be taken. For example, if the variance of response variables is not a constant, the variance stabilization approach should be used (Montgomery 2005). The model appropriateness can be evaluated by the "goodness of fit," also known as R^2, which can be loosely interpreted as the proportion of the variability in the data "explained" by the ANOVA model. Thus, if R^2 is high, it implies that the model fitting is good; otherwise, the model is of poor quality and more factors will be needed to be included in the model.

Result evaluation and interpretation

The derived results should be presented to domain engineers for evaluation and interpretation. As the developed framework provides valuable information for domain engineers to perform trouble shooting, the inputs offered by the domain engineers also improves the quality of analysis. It is worth mentioning that sometimes the root causes may not be found in the first place. The reasons can be attributed to:

(i) Process-loss effect: the semiconductor manufacturing process is very complex and lengthy; wafers typically go through more than one hundred steps. Consequently, some stages containing problematic machines may be missed and the problematic machines cannot be identified.
(ii) Data-loss effect: there are no sufficient combinations to form a proper design, or the data arrays found by retrospective DOE do not contain sufficient data points. Consequently, the information is not enough for identifying the root causes of low yield.

The incorporation of domain engineers' opinions can significantly reduce the occurrence of the above two problems. The identified key process stages should be presented to domain experts and check if all key process stages are included. If it is suspected that some other process stages that can contain problematic machines are not included, discussions with domain experts would be needed. Similarly, the identified problematic machines should be confirmed by domain experts whether they are actually the root causes of low yield and ensure that no other problematic machines are yet found. The involvement of domain engineers not only can improve the analysis quality but also assist in result interpretation.

An empirical study

In this section, we conduct an empirical study in collaboration with a leading semiconductor company in Taiwan to validate the proposed approach.

Problem definition

This semiconductor company has been a world-leading flash memory manufacturer that produces a wide range of high-performance non-volatile memory ICs and micro-controller ICs used in communication systems, computers, and high-end electronic consumer systems.

The product in this empirical study is Mask ROM. The problem background is that the function parameter (Fail_bin1) is unusually high during the period 5/14–6/17, as indicated with red circle in Fig. 6. Because high Fail_bin1 implies low yield rate, it is important to find the root causes, removing them as quickly as possible so as to minimize the incurred yield loss. In this study, we define Fail_bin1 as the response variable, process stage as factor and machine as level.

Data preparation

The case company has built an Engineering Data Analysis System (EDAS) to collect and manage the manufacturing data. All related manufacturing data during the period 5/14–6/17, including Fail_bin1, process stages, machines, lot id etc, is selected from database and integrated into a coherent data set. Totally, there are 200 data points and each data point contains the information of 146 process stages and the machines. Careful data cleaning is performed to eliminate incorrect and inaccurate data such as spelling errors.

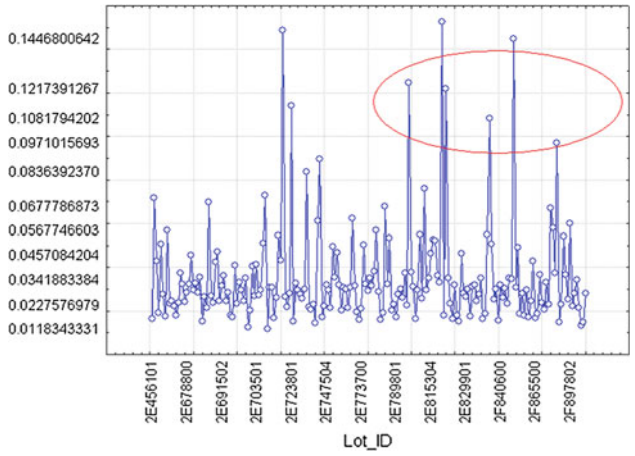

Fig. 6 Scatter plot of Fail_bin1

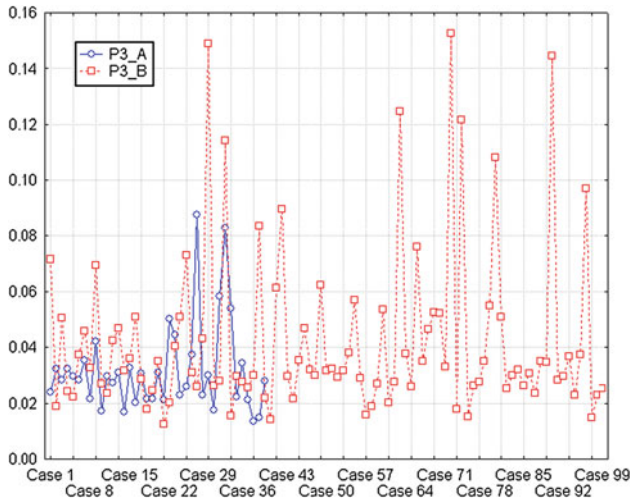

Fig. 7 Scatter plot of P3_A and P3_B

Key stage identification

Firstly, the K–W test is applied to identify the key process stages. Based on the discussions with domain experts, the process stage with p value less than 0.15 is identified as a key stage. In this study, 13 process stages, labeled as P1 to P13, are identified as the key process stages. Further, multiple comparison test is conducted to distinguish the normal machines from the problematic ones. We take P3 as an example. In P3, there are three machines, P3-A, P3-B, P3-C. As shown in the scatter plot of Fig. 7, where the blue line represents machine P3-A and the red dot represents machine P3-B. Clearly, P3-B has much higher Fail_bin 1 than that of P3-A. Based on the multiple comparison test listed in Table 3, it is found that the difference of Fail_bin1s between P3-A and P3-B is the most significant with a significance level of 0.005. Thus, we derive P3-B as the problematic machine and P3-A as the normal machine.

Table 3 Multiple comparison test

Lots mean	Machine Id		
	P3-A (0.02931)	P3-B (0.04209)	P3-C (0.028376)
P3-A		**0.002389**	0.981415
P3-B	**0.002389**		0.011039
P3-C	0.981415	0.011039	

Retrospective design of experiment

This step finds the useful experimental designs existing in the large database. To facilitate the search process, the normal machines is labeled as +1 and the problematic machine as −1. Since there are a total of 200 data points, all designs that only require 200/4 = 50 or fewer design points are searched. Following the proposed retrospective DOE, several $2^3, 2^4, 2^5$ full factorial designs, and $2^{4-1}, 2^{5-1}, 2^{6-1}$ fractional factorial designs can be matched in the dataset. These data arrays are stored for subsequent data analysis. For illustration, only the data analysis performed on the data array of 2^3 full factorial design with three factors P2, P3, and P4 is presented.

Data analysis

This study applies ANOVA to identify the problematic machines, and detect whether there exist interactions between problematic machines. Table 4 shows the results of the ANOVA test with a significance level of 0.05, where P3, P4, are significant factors at α-level = 0.05, coinciding with the K–W test. Moreover, the interaction between P3 and P4 are found to be significant. The R^2 of the fitted model is 0.6. In practice, different process steps may have interrelated impacts on yield. For example, lithography and etching are two fabrication processes for determining critical dimension (CD) for wafer fabrication. The CD measured by the metrology after the lithography process is called the developed critical dimension (DCD). The CD measured by the metrology after the etching process is called the etched critical dimension (ECD). The ECD represents the linewidth of fabricated patterns of each layer on the wafer, in which the variation of both DCD and ECD will directly affect process yield and product quality.

This finding is further confirmed by the interaction plot shown as Fig. 8. Clearly, there is an obvious interaction between P3 and P4. It means the combination of these two processes have impacts on the yield performance. For example, the pattern development and etching process are interrelated and thus may create excursion due to the combination of specific tools used in these two steps.

Table 4 The ANOVA test

Factor	SS	*df*	MS	F	*P* value
P2	0.000681	1	0.000681	4.1778	0.071
P3	0.008236	1	0.008236	50.5084	**0.0000**
P4	0.008109	1	0.008109	49.7322	**0.0000**
P2*P3	0.000754	1	0.000754	4.6212	0.0604
P2*P4	0.000010	1	0.000010	0.11070	0.7461
P3*P4	0.001082	1	0.001082	6.6383	**0.0296**
Error	0.001467	9	0.000163		
Total	0.020347	15			

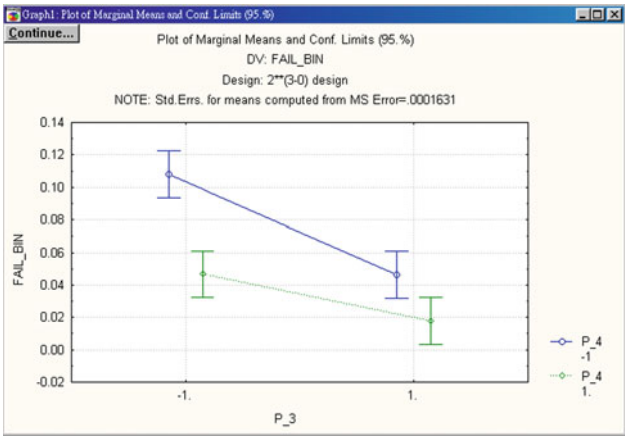

Fig. 8 Interaction plot between P3 and P4

The model assumptions are validated by conducting residual analysis. The normal probability plot (Fig. 9) and the residual plot (Fig. 10) both show that there is no serious violation for the assumptions. Because the $R^2 = 0.6$ is not high, this suggests that there might be some other factors that can affect the response variable but are yet found. Therefore, we proceed to use other data arrays found in the previous step and perform data analysis to identify other significant factors. The procedure is iterated several times and finally it is concluded that the process stages that may contain problematic machines are P3, P4, and P7.

Evaluation and interpretation

The final results are presented to domain experts for evaluation and interpretation. The proposed framework revealed that P3, P4 and P7 are process stages that can contain problematic machines. Because the discussion with domain experts eliminates P7, the attention is placed on the process stages P3 and P4. In previous steps, P3-A and P4-B are problematic machines in process stages P3 and P4 respectively. After checking the manufacturing data, it is found that the Fail_bin1 of wafer lots passing through P3-A and P4-B is indeed much higher than those that do not pass through

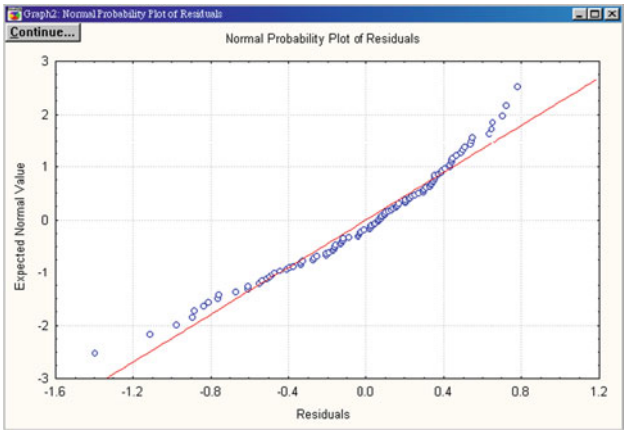

Fig. 9 Normal probability plot of the residuals

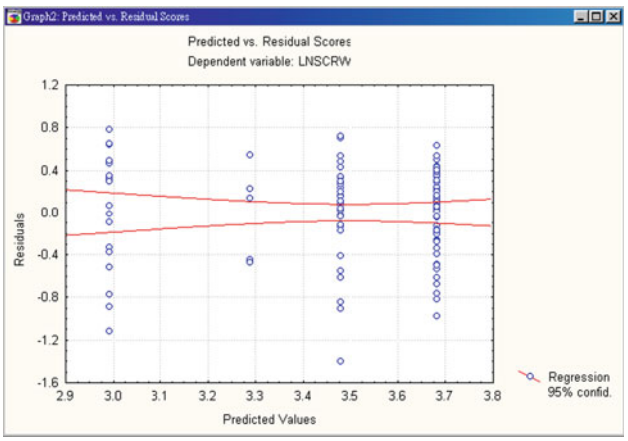

Fig. 10 Residual plot

these two machines. Therefore, P3-A and P4-B are confirmed as the root causes that result in low yield. Indeed, domain experts also agree this conclusion. Based on the results, the domain engineers are able to take some corrective actions to handle the problems and reduce the yield-loss cost.

Conclusion

This study proposed a novel approach that integrates DOE and data mining to enable effective yield-loss diagnosis in semiconductor manufacturing. Compared to the existing practice that relies on domain engineers' personal experience in different process modules, the proposed approach is more systematic and holistic that can narrow the scope of possible root causes and thus reduce the efforts required for fault diagnosis. This approach also requires less time than conventional data mining that may become data dredging in semiconductor manufacturing and can generate useful results that are easier for interpretation to support yield enhancement. The developed approach does not require extra experimentation

cost and can fully utilize the existing data to employ appropriate DOE-enabled analysis to extract meaningful patterns for potential usage. To validate the proposed framework, an empirical study was conducted in a leading semiconductor manufacturing company in Taiwan and the results showed that the proposed approach can effectively identify the root causes of low yield and provide useful manufacturing intelligence to assist domain engineers in removing assignable causes. Furthermore, since the developed approach does not require extra experimentation or specific designs to be followed in fab, it is suitable to be implemented in semiconductor manufacturing where the experimentation cost is high and the environment is changing and thus difficult to control the designed setting for effective DOE. This is of high practical value for many industries, especially for those where the experimentation cost is high or the designed experiments are difficult to be conducted in real setting. Indeed, the proposed approach has been employed as a function of the engineering data analysis platform in the case company. Future research can be done to develop effective production control mechanism to maximize potential manufacturing intelligence extracted from regular operations to achieve both productivity and quality objectives in modern production systems.

Acknowledgments This research is partially supported by National Science Council, Taiwan (NSC99-2221-E-007-047-MY3; NSC102-2622-E-007-013), National Tsing Hua University under the Toward World-Class University Project (101N2073E1), and Macronix International Ltd. (93A0309J8) in Taiwan.

References

Berry, M., & Linoff, G. (1997). *Data mining techniques for marketing, sales and customer support*. New York, NY: Wiley.

Braha, D., & Shmilovici, A. (2002). Data mining for improving a cleaning process in the semiconductor industry. *IEEE Transactions on Semiconductor Manufacturing*, *15*(1), 91–101.

Braha, D., & Shmilovici, A. (2003). On the use of decision tree induction for discovery of interactions in a photolithographic process. *IEEE Transactions on Semiconducter Manufacturing*, *16*(4), 644–652.

Chen, W., & Chien, C.-F. (2011). Measuring relative performance of wafer fabrication operations: A case study. *Journal of Intelligent Manufacturing*, *22*(3), 447–457.

Chien, C.-F., & Chen, C. (2007a). A novel timetabling algorithm for a furnace process for semiconductor fabrication with constrained waiting and frequency-based setups. *OR Spectrum*, *29*(3), 391–419.

Chien, C.-F., & Chen, L. (2007b). Using rough set theory to recruit and retain high-potential talents for semiconductor manufacturing. *IEEE Transactions on Semiconductor Manufacturing*, *20*(4), 528–541.

Chien, C.-F., & Chen, L. (2008). Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry. *Expert Systems with Applications*, *34*(1), 280–290.

Chien, C.-F., Chen, Y., & Peng, J. (2010). Manufacturing intelligence for semiconductor demand forecast based on technology diffusion and product life cycle. *International Journal of Production Economics*, *128*(2), 496–509.

Chien, C.-F., Dauzere-Peres, S., Ehm, H., Fowler, J. W., Jiang, Z., Krishnaswamy, S., et al. (2011). Modeling and analysis of semiconductor manufacturing in a shrinking world: Challenges and successes. *European Journal of Industrial Engineering*, *5*(3), 254–271.

Chien, C.-F., & Hsu, C. (2006). A novel method for determining machine subgroups and backups with an empirical study for semiconductor manufacturing. *Journal of Intelligent Manufacturing*, *17*(4), 429–439.

Chien, C.-F., & Hsu, C. (2011). UNISON analysis to model and reduce step-and-scan overlay errors for semiconductor manufacturing. *Journal of Intelligent Manufacturing*, *22*(3), 399–412.

Chien, C.-F., Hsu, C., & Chang, K. (2013a). Overall wafer effectiveness (OWE): A novel industry standard for semiconductor ecosystem as a whole. *Computers & Industrial Engineering, 65*(1), 117–127.

Chien, C.-F., Hsu, C., & Hsiao, C. (2012a). Manufacturing intelligence to forecast and reduce semiconductor cycle time. *Journal of Intelligent Manufacturing*, *23*(6), 2281–2294.

Chien, C.-F., Hsu, S., & Chen, Y. (2013b). A system for online detection and classification of wafer bin map defect patterns for manufacturing intelligence. *International Journal of Production Research*, *51*(8), 2324–2338.

Chien, C.-F., Wang, W., & Cheng, J. (2007). Data mining for yield enhancement in semiconductor manufacturing and an empirical study. *Expert System with Applications*, *33*(1), 192–198.

Chien, C.-F., Wu, C., & Chiang, Y. (2012b). Coordinated capacity migration and expansion planning for semiconductor manufacturing under demand uncertainties. *International Journal of Production Economics, 135*(2), 860–869.

Chien, C.-F., & Zheng, J.-N. (2012). Mini-max regret strategy for robust capacity expansion decisions in semiconductor manufacturing. *Journal of Intelligent Manufacturing*, *23*(6), 2151–2159.

Coleman, D. E., Montgomery, D. C., Gunter, B. H., Hahn, G. J., Haaland, P. D., O'Connell, M. A., et al. (1993). A systematic approach to planning for a designed industrial experiment. *Technometrics*, *35*, 1–27.

Dunn, O. J. (1964). Multiple comparisons using rank sums. *Technometrics*, *6*(3), 241–252.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD process for extracting useful knowledge from volumes of data. *Communication of the ACM*, *39*(11), 27–34.

Han, J., & Kamber, M. (2001). *Data mining: Concepts and techniques*. San Francisco, CA: Morgan Kaufmann Publishers.

Harding, J. A., Shahbaz, M., Srinivas, & Kusiak, A. (2006). Data mining: A review. *Journal of Manufacturing Science and Engineering*, *128*(4), 969–976.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.

Hsu, S., & Chien, C.-F. (2007). Hybrid data mining approach for pattern extraction from wafer bin map to improve yield in semiconductor manufacturing. *International Journal of Production Economics*, *107*(1), 88–103.

Hwang, J. Y., & Kuo, W. (2007). Model-based clustering for integrated circuit yield enhancement. *European Journal of Operational Research*, *178*(1), 143–153.

Jeong, Y., Kim, S., & Jeong, M. (2008). Automatic identification of defect patterns in semiconductor wafer maps using spatial correlogram and dynamic time warping. *IEEE Transactions on Semiconducter Manufacturing*, *21*(4), 625–637.

Khuri, A. I., & Cornell, J. A. (1996). *Response surfaces designs and analyses*. New York: Marcel Dekker.

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks on one-criterion variance analysis. *Journal of the American Statistical Association*, *47*(260), 583–621.

Kuo, C., Chien, C.-F., & Chen, C. (2011). Manufacturing intelligence to exploit the value of production and tool data to reduce cycle time. *IEEE Transactions on Automation Science and Engineering*, *8*(1), 103–111.

Kusiak, A., & Kurasek, C. (2001). Data mining of printed-circuit board defects. *IEEE Transactions on Robotics and Automation*, *17*(2), 191–196.

Liu, C.-W., & Chien, C.-F. (2013). An intelligent system for wafer bin map defect diagnosis: An empirical study for semiconductor manufacturing. *Engineering Applications of Artificial Intelligence, 26*(5–6), 1479–1486.

Marcus, R., Peritz, E., & Gabriel, K. R. (1976). On closed testing procedures with special reference to ordered analysis of variance. *Biometrika*, *63*, 655–660.

May, G. S., Huang, J., & Spanos, C. J. (1991). Statistical experiment design in plasma etch modeling. *IEEE Transactions on Semiconductor Manufacturing*, *4*(2), 83–98.

Montgomery, D. C. (2005). *Design and analysis of experiments* (6th ed.). New York, NY: Wiley.

Moore, G. E. (1965). Cramming more components onto integrated circuits. *Electronics*, *38*(8), 114–117.

Wu, C. F. J., & Hamada, M. (2000). *Experiments: Planning, analysis, and parameter design optimization*. New York, NY: Wiley.

Wu, J., & Chien, C.-F. (2008). Modeling strategic semiconductor assembly outsourcing decisions based on empirical settings. *OR Spectrum*, *30*(3), 401–430.

Wu, J.-Z. (2013). Inventory write-down prediction for semiconductor manufacturing considering inventory age, accounting principle, and product structure with real settings. *Computers & Industrial Engineering*, 65(1), 128–136.

Wu, J.-Z., Hao, X.-C., Chien, C.-F., & Gen, M. (2012). A novel bi-vector encoding genetic algorithm for the simultaneous multiple resources scheduling problem. *Journal of Intelligent Manufacturing*, *23*(6), 2255–2270.