Journal of Applied Sciences

ISSN 1812-5654

Journal of Applied Sciences 13 (21): 4695-4701, 2013
ISSN 1812-5654 / DOI: 10.3923/jas.2013.4695.4701

A Markov Property Considered Data Generation Approach for Etchers Fault Detection Test

[1]Jungang Yang, [1]Jie Zhang, [2]Yihua Ma and [2]Zhiyu Wang
[1]Institute of Computer Integrated Manufacturing, Shanghai Jiao Tong University, Shanghai, 200240, People's Republic of China
[2]Shanghai Baosight Software Co., Ltd, Shanghai, 201203, People's Republic of China

Abstract: The objective of this research was to investigate the process features of etchers and automatically provide mass multi-characteristics data for fault detection test. Before applying a fault detection system in a modern semiconductor wafer fabrication system, a great number of experiments and tests must be carried out to ensure the effectiveness and efficiency of the system. However, it is difficult to collect and storage different types of data with different characteristics for testing. This study proposed a Markov Property considered data generation approach based on analysis of the process state changes in etchers. The markov property of etcher's state changes was demonstrated and a data generation model was built. Comparing with the existed historical data based data generation method and random data generation method, the proposed method considered not only statistical information from historical data but also the impact of the etchers' state changes. Experiments and industrial examples were used to measure the performance of the proposed method and results show that it has advantages such as simple expression, rapid and automatic data generation and easy reconfiguration, therefore is especially useful for the Fault Detection system test and simulation.

Key words: Markov property, test data generation, fault detection, etcher, semiconductor manufacturing

INTRODUCTION

While the etching process of the modern Semiconductor Wafer Fabrication System (SWFS) is challenged by smaller line width and more complex techniques (May and Sze, 2007), several measures has been applied to monitor and control the process and equipment. Fault Detection and Classification (FDC) is one of them (Imai *et al*., 2006). In order to ensure the effectiveness and efficiency of the FDC sytem and control risks, massive data tests are usually carried out before applying a FDC system. However, there is some difficulty in test data collection. First of all, it is hardly to collect and storage process data sets as much as possible. What's more, the data sets get from practical process may not contain all of faults and failures, thus can not validate some specific equipment state. Finally, the real wafer fabricating equipment is rather expensive and applying an uncertified FDC system is extremely high-risk. As a result, it is urgent to search for a etcher's state data generation method to expand the amount of data and test FDC systems.

The etcher state parameters usually follow a certain statistical distribution but its value changes randomly. A random data sequence is widely used in chip test, program test and logical test (Wang, 2007; Como *et al*., 2003; Bertolino, 2007). There are several generally used methods to output a sequence of data:

- **Stochastic methods:** Random number, random table search and finite-state machine are typical stochastic methods
- **Time series methods:** Random distribution and Markov process are common examples
- **Artificial Intelligence methods:** For example, fuzzy logic and artificial neural network

Among all of these data generation methods, Markov process was the most common applied because Markov process can be easily treatment under mathematics methods and tools. Besides, the transition matrix of a Markov process is usually calculated according to historic statistic data. A Markov Property based process state data generation method can analyze former situation and generate more actual data. Hence this study aimed to build a Markov Property considered data generation model and generate a series of test data for FDC system simulation.

Corresponding Author: Jie Zhang, Institute of Computer Integrated Manufacturing, Shanghai Jiao Tong University, Shanghai, 200240, People's Republic of China

DATA GENERATION METHOD

Markov property of etcher's state changes: A stochastic process has the Markov Property if the conditional probability distribution of future states of the process (conditional on both past and present values) depends only upon the present state, not on the sequence of events that preceded it (Feller, 2008).

Assume that $S = \{S_n, n = 0, 1, 2, 3, \dots\}$ is a discrete stochastic time process, in which n means time, Sn means the state on time n. A Markov chain is a sequence of random variables $X_1, X_2, X_3, \dots$ with the Markov Property, namely that, given the present state, the future and past states are independent. Formally:

$$\begin{aligned} &\Pr(X_{n+1} = x \mid X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) \\ &= \Pr(X_{n+1} = x \mid X_n = x_n) \end{aligned}$$

The possible values of x_i form a countable set S called the state space of the chain.

If the state space is finite, the transition probability distribution can be represented by a matrix, called the transition matrix, with the (I, j) th element of P equals to:

$$p_{ij} = \Pr(X_1 = j | X_0 = i)$$

The transition matrix can also be expanded as follows:

$$P = \begin{bmatrix} p_{11}(k) & p_{12}(k) & p_{13}(k) & \cdots \\ p_{21}(k) & p_{22}(k) & p_{23}(k) & \cdots \\ p_{31}(k) & p_{32}(k) & p_{33}(k) & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

The transition matrix has the following features:

$$\begin{cases} 0 \le p_{ij} \le 1 \\ \sum_{j \in S} p_{ij} = 1 \end{cases} \quad i = 1, 2, \dots; j = 1, 2, \dots$$

If the Markov chain is time-homogeneous, then the transition matrix P is the same after each step, so the k-step transition probability can be computed as the k-th power of the transition matrix, P^k.

The etching process is a complex chemical and physical process which is continuously reacts in the chambers of an etcher. In order to monitor the process, there are usually a lot of different sensors in different part of an etcher. These sensors commonly collect and send back data with a certain interval. In this situation, a continuous parameter value is scattered into a discrete time series data. Each group of data represents a certain equipment state on a specific time point. The equipment state changes at these time points. Hence the continuous process is simplified as a discrete time series of sample point

The discrete etching process has Markov Property because of the following reasons:

- Failure probability of an etcher has its own characteristic and follows a certain distribution
- Changes of state parameter value are randomly occurred
- Next equipment state depends on the current state and has no relationship with former states

Model definition: According to the characteristics of process state changes, the etch process meet the following assumptions:

- States of each chamber in an etcher are independent
- A chamber can only be a specific state at a time
- The total amount of process state is limited
- Parameters value change will not exceed the limit of certain range

The notations put below are used in this study:

S : Set of all the process state
n : No. of process state
s_i : The ith process state, $s_i \in S$, I = 1, 2, ….., n
U : Set of all the process state parameters
m : No. of process state parameters
u_j : The jth process state parameter, $u_j \in U$, j = 1, 2, ….., m
P : Transition matrix
W : Set of model state

The Markov Property based state model of etchers is as follows:

$$M = (S, U) = \{(s_i, I = 1, \dots, n), (u_j, j = 1, \dots, m)\}$$

In the worst case, a minor change of parameter u_i may cause all of the process states s_i. In this situation, the model state set is:

$$W = \{(s_1, u_1), (s_2, u_1), .., (s_n, u_m)\}$$

The total amount of states is n×m = nm.

However, in a practical situation, some of these states may not occur and the state set W can be reduced.

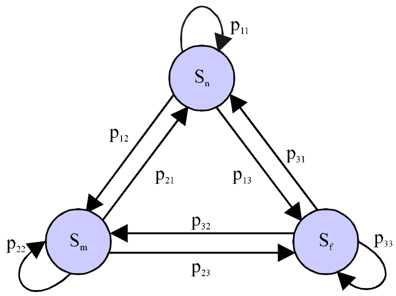

Fig. 1: A 3 states transfer map

Here, if we assume that there are 3 states:

$$S = \{\text{normal, fault, failure}\}$$

According to the relationship of the 3 states, we can get the transfer relationship of each two states.

Hence, the transition matrix can be drawn from Fig. 1:

$$P = \begin{bmatrix} s_n \to s_n & s_n \to s_m & s_n \to s_f \\ s_m \to s_n & s_m \to s_m & s_m \to s_f \\ s_f \to s_n & s_f \to s_m & s_f \to s_f \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$$

From time t-t+1 then t+2, the probable state and its transfer probability can be demonstrated in Fig. 2.

Modeling steps:

Step 1: Determine the state space. A sample mean-standard deviation classification is used to divide the interval into several ranges:

- Calculate the mean of the data:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

- Calculate the standard deviation of the data:

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

- Divide the interval as follows:

$$\begin{cases} [-\infty, \bar{x} - \alpha s) \\ [\bar{x} - \alpha s, \bar{x} + \alpha s), \qquad \alpha \in [1, 1.5] \\ [\bar{x} + \alpha s, \infty) \end{cases}$$

Step 2: Calculate the transition matrix

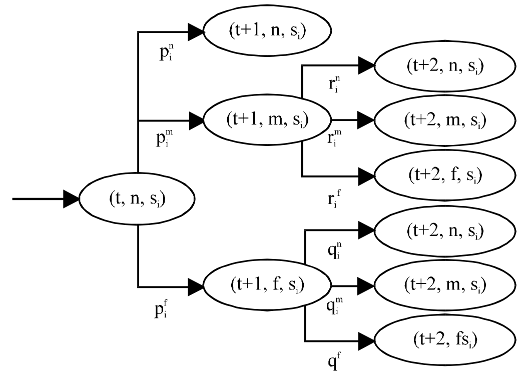

Fig. 2: Transfer of state space probability

- Count the amount of all the states
- Calculate amount f_{ij} of states which has changed from current state to next state
- Calculate the state transition probability:

$$P_{ij} = \frac{f_{ij}}{\sum_{i=1}^{n} f_{ij}}$$

And the transition matrix is as follows:

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}$$

Step 3: Generate state probability vector according to the definition of Markov Chain:

$$v_{i+1} = v_i P$$

Step 4: Generate parameter values

- Make a judgment of current generated state v_i.

$$v_i = [v_{i1}, v_{i2}, \ldots, v_{in}$$

Here, a random number r is introduced to judge which state it will locate:

$$r_j = \text{random}\ (j),\ j = 1, 2, \ldots, m,\ 0 \le r_j \le 1$$

$$\begin{cases} r_j < v_{i1} \\ v_{i1} \le r_j < v_{i1} + v_{i2} \\ \cdots \\ \sum_{k=1}^{m-1} v_{ik} \le r_j < \sum_{k=1}^{m} v_{ik} = 1 \end{cases} \quad \text{then} \quad \begin{matrix} r_j \in \text{state1} \\ r_j \in \text{state2} \\ \cdots \\ r_j \in \text{state3} \end{matrix}$$

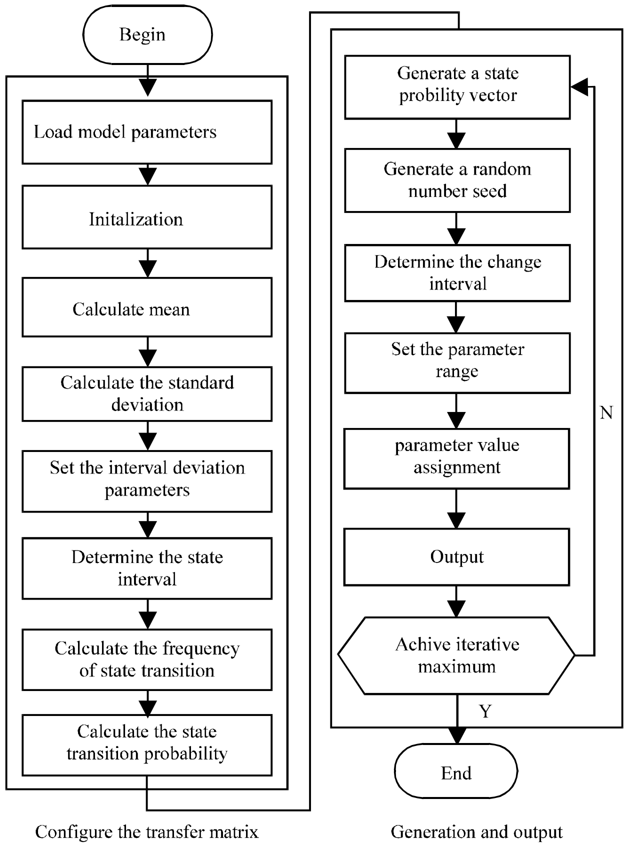

Fig. 3: Flow chart of the data generation model

- Restore parameter values in the corresponding target state based on their range and distribution
- Output state parameter values

The flow chart of the proposed data generation model follows here.

CASE STUDY

This section presents numerical and actual experiment studies to demonstrate the effectiveness of the proposed approach. Prototype system is developed by C No. under. Net Framework 2.0 in Microsoft Visual Studio 2008. The program is cooperating with Microsoft Excel 2007 as data treatment. All experiments are carried out on a computer with a configuration of 2.26 GHz P8400 Intel Core 2 Duo CPU and 4 GB memory.

Illustrative examples: In this part a single parameter data generation example is given to illustrate the capabilities of the proposed data generation method.

Assume that there are 3 process states:

$$S = \{normal, fault, failure\}$$

The initialize state is normal, which is:

$$v_0 = [1, 0, 0]$$

The transition matrix is given as:

$$P = \begin{bmatrix} 0.95 & 0.04 & 0.01 \\ 0.45 & 0.45 & 0.1 \\ 0.2 & 0.55 & 0.25 \end{bmatrix}$$

The interval of data generation is 1 Hz and the time span is 1000 sec.

The distribution of normal, fault and failure states are U (110, 120), U (120, 124) and U (124, 126).

The results are shown in Fig. 4 and Table 1.

Industrial examples: In this part, an industrial example is used to demonstrate the performance of the proposed data generation method. The original data set is collected from an Aluminum stack etch process performed on a commercial scale Lam 9600 plasma etch

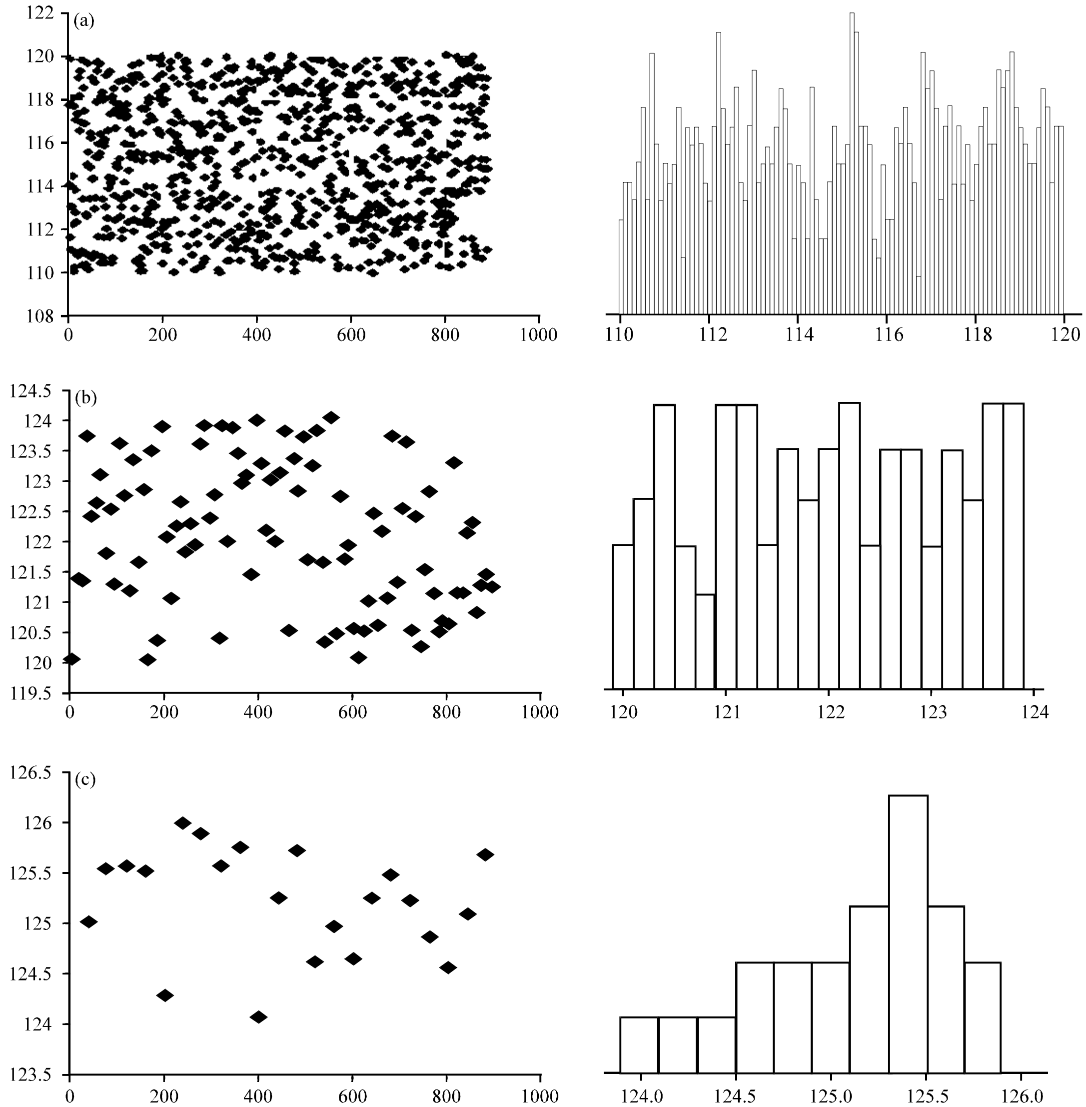

Fig. 4 (a-c): Distribution and histogram of generated data (a) Normal state data, (b) Fault state data and (c) Failure state data

tool at Texas Instrument, Inc. (Eigenvector Res. Inc., 1999; Wise *et al*., 1999). The data is stored by the machine state sensor system at 1-s intervals during etching. The data consists of 108 normal wafers and 21 fault wafers, Each wafer's data set contents about 100 group of records and the total number of all the records is more than 75, 000. A more detailed description on the data can be found in Eigenvector Res. Inc. (1999).

In this chosen data set, 6 major variables are selected due to the relevancy to process and final product state (Wise *et al*., 1999). These variables are: BCl_3 Flow, Cl_2 Flow, Chamber Pressure, RF Bottom Power, Helium Press and TCP Top Power. Fetures of sample data can be found in Table 2.

As there are 6 types of faults, the total number of states is 7, including one normal state and 6 fault state:

$$S = \{s_1(\text{normal}), s_2(u_1\text{fault}), s_3(u_2\text{fault}), s_4(u_3\text{fault}), s_5(u_4\text{fault}), s_6(u_5\text{fault}), s_7(u_6\text{fault}), s_8(u_7\text{fault})\}$$

Then we can obtain the model state matrix W:

$$W = \begin{bmatrix} (1, s_1) & (1, s_2) & \cdots & (1, s_7) \\ (2, s_1) & (2, s_2) & \cdots & (2, s_7) \\ \vdots & \vdots & \cdots & \vdots \\ (n, s_1) & (n, s_2) & \cdots & (n, s_7) \end{bmatrix}$$

There are total 7n states.

Table 2: Features of actual and generated fault data

		Actual fault data			Generated fault data		
No.	Paragraph name	Distribution	Freq.	PCT. (%)	Distribution	Freq.	PCT. (%)
U_1	BCl_3 flow	N (751.689, 0.268)	3	2.326	N (761.7, 0.3)	49	2.45
U_2	Helium press	N (100.636, 0.602)	1	0.775	N (90.6, 0.6)	20	1.00
U_3	RF Btm. Power	N (133.292, 3.134)	4	3.101	N (121.3, 3.1)	64	3.02
U_4	Cl_2 flow	N (753.230, 0.568)	3	2.326	N (743.2, 0.6)	41	2.05
U_5	Chamber pressure	N (1203.361, 104.476)	4	3.101	N (1253.4, 104.5)	71	3.55
U_6	TCP top power	N (349.405, 26.018)	6	4.651	N (379.4, 26.0)	118	5.09
	Total	/	21	16.028	/	363	18.15

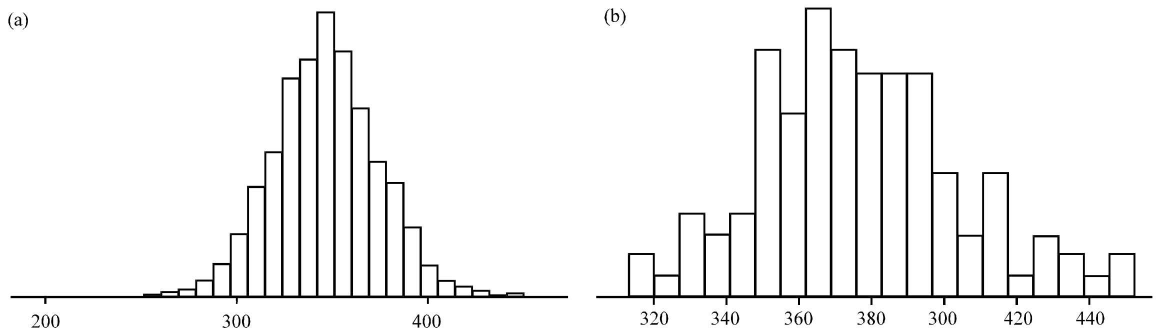

Fig. 5 (a-b) : Histogram of generated data (a) Normal data (b) Fault data

Suppose that each fault has 80% chances to be fixed, then we can construct the transition matrix P:

$$P = \begin{bmatrix} 0.83721 & 0.02326 & 0.00775 & 0.03101 & 0.02326 & 0.03101 & 0.04650 \\ 0.8 & 0.2 & 0 & 0 & 0 & 0 & 0 \\ 0.8 & 0 & 0.2 & 0 & 0 & 0 & 0 \\ 0.8 & 0 & 0 & 0.2 & 0 & 0 & 0 \\ 0.8 & 0 & 0 & 0 & 0.2 & 0 & 0 \\ 0.8 & 0 & 0 & 0 & 0 & 0.2 & 0 \\ 0.8 & 0 & 0 & 0 & 0 & 0 & 0.2 \end{bmatrix}$$

The initialize state is normal, which is:

$$v_0 = [1, 0, 0, 0, 0, 0, 0]$$

The interval of data generation is 1 Hz and the time span is 2000 sec.

The distribution of generated fault data mainly follows the actual fault data, except a minor add/minus on average value.

The actual and generated fault data features are both shown in Table 2.

CONCLUSION

In this study, a state data generation method considering with Markov Property is developed to provider mass amount data for etcher's fault detection test. Because the transition matrix is calculated according to the historical equipment data, the proposed equipment state data generation process considers about the former experience of equipment state and makes a proper prediction for the future trends, thus the data collected from this generation process is more accurate and actual than random generated data. Also, an industry experiment shows that the generated data is similar to the real behavior of etchers. This data generation method can be widely used for different types of etchers and different manufacturing process of wafers, hence can test the fault detection application in a simulation environment. This can great help to the test and improvement of FDC systems. In this study we highlight the simulation test data generation for FDC and try to set up a "simulation-improvement-application" cycle with the purpose of rapid development of FDC systems. With further extension and improving of the proposed approach, it should consider combining the data generation module with fault detection module in a universal platform and generate test example data to test FDC method.

ACKNOWLEDGMENTS

This study was supported by the National Science and Technology Major Project of China under Grant No. 2011ZX02501 and Specialized Research Fund for the Doctoral Program of Higher Education under Grant No. 20120073110036.

.REFERENCES

Bertolino, A., 2007. Software testing research: Achievements, challenges, dreams. Proceedings of the Future of Software Engineering, May 23-25, 2007, Minneapolis, MN., pp: 85-103.

Como, F., G. Cumani and M.S. Reorda, 2003. Automatic test program generation for pipelined processors. Proceedings of the 18th Annual ACM Symposiumon on Applied Computing, March 9-12,2003, Melbourne, Florida, USA., pp: 736-740.

Eigenvector Res. Inc., 1999. Metal etch data for fault detection evaluation. http://www.eigenvector.com/data/Etch/

Feller, W., 2008. An Introduction to Probability Theory and its Applications. Vol. 2, John Wiley and Sons, New York.

Imai, S., N. Sato, M. Kitabata and S. Yasuda, 2006. Fab-wide equipment monitoring and FDC system. Proceedings of the IEEE International Symposium on Semiconductor Manufacturing, September 25-27, 2006, Tokyo, pp: 114-117.

May, G.S. and S.M. Sze, 2007. Fundamentals of Semiconductor Fabrication. Wiley, New York, ISBN: 9780471232797.

Wang, L., 2007. Models and Methods For Maintenance Decisions: Theory And Applications. Zhejiang University, Zhejiang, China.

Wise, B.M., N.B. Gallagher, S.W. Butler, D.D. White and G.G. Barna, 1999. A comparison of principal component analysis, multiway principal component analysis, trilinear decomposition and parallel factor analysis for fault detection in a semiconductor etch process. J. Chemometrics, 13: 379-396.